



---

## **E-QALY: face validity testing the items for a new generic preference-based measure**

### AUTHOR(S)

Jill Carlton, Janice Connell, Tessa Peasgood, Clara Mukuria, Kristina Ludwig, Ole Marten, Simone Kreimeier, Lidia Engel, Zhihao Yang, Maria Belizan, Rocio Rodriguez, Andrea Monteiro, Maja Kuharic, Brendan Mulhern, A Simon Pickard, Nan Luo, Federico Augustovski, Wolfgang Greiner

### PUBLICATION DATE

01-10-2020

### HANDLE

[10536/DRO/DU:30148478](https://hdl.handle.net/10536/DRO/DU:30148478)

Downloaded from Deakin University's Figshare repository

Deakin University CRICOS Provider Code: 00113B

## 27th Annual Conference of the International Society for Quality of Life Research

© Springer Nature Switzerland AG 2020

### Cutting Edge Research Plenary Session

#### (1) Phase III randomized controlled trial of eRAPID (electronic patient self-reporting of Adverse-events: Patient Information and advice)—An eHealth intervention during chemotherapy

Galina Velikova, MD PhD FRCP, Leeds Institute of Medical Research at St James's, University of Leeds, Leeds Cancer Centre, Leeds, United Kingdom; Kate Absolom, Leeds Institute of Medical Research at St James's, University of Leeds, Leeds Cancer Centre, St James's University Hospital, Leeds, United Kingdom; Lorraine Warrington, Leeds Institute of Medical Research at St James's, University of Leeds Cancer Centre, St James's University Hospital, Leeds, United Kingdom; Eleanor Hudson, Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, United Kingdom; Jenny Hewison, Leeds Institute of Health Sciences, University of Leeds, Leeds, United Kingdom; Carolyn Morris, Patient representative, Independent Cancer Patients Voices, Brighton, United Kingdom; Patricia Holch, Leeds Institute of Medical Research at St James's, University of Leeds, Leeds Cancer Centre, St James's University Hospital, Leeds, United Kingdom; Robert Carter, Leeds Institute of Medical Research at St James's, University of Leeds, Leeds Cancer Centre, St James's University Hospital, Leeds, United Kingdom; Andrea Gibson, Leeds Institute of Medical Research at St James's, University of Leeds, Leeds Cancer Centre, St James's University Hospital, Leeds, United Kingdom; Marie Holmes, Leeds Institute of Medical Research at St James's, University of Leeds, Leeds Cancer Centre, St James's University Hospital, Leeds, United Kingdom; Zoe Rogers, Leeds Institute of Medical Research at St James's, University of Leeds, Leeds Cancer Centre, St James's University Hospital, Leeds, United Kingdom; Lucy McParland, Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, United Kingdom; Claire Hulme, University of Exeter, St Luke's Campus, Exeter, United Kingdom; Julia Brown, Leeds Institute of Clinical Trials Research, University of Leeds, Leeds, United Kingdom

**Aims:** eRAPID is an online eHealth system for patients to report symptoms during cancer treatment. It provides automated severity-dependent advice to patients guiding self-management or medical

contact. Patient reports are transferred to the electronic patient records for immediate use in patient care. This trial evaluated the impact of eRAPID on symptom control, clinical care, patient self-efficacy, and quality of life (QOL). **Methods:** A prospective randomized parallel two-arm trial included colorectal, breast or gynecological patients commencing chemotherapy at a UK cancer center. Participants were randomized to usual care plus eRAPID or usual care (UC). eRAPID patients completed weekly symptom reports for 18 weeks. **Primary outcome:** symptom control at 18 weeks (Functional-Assessment-Cancer-Therapy-Physical Wellbeing, FACT-PWB). **Secondary outcomes:** cost-effectiveness, clinical process measures (admissions/chemotherapy delivery), patient self-efficacy (6-item Self-Efficacy Scale) and global QOL (FACT-G). Multivariable mixed-effects repeated measures models were employed. Clinical relevance of any differences was evaluated with responder analysis. **Trial registration:** ISRCTN88520246. **Results:** Between Jan 2015 and June 2018, 508/690 eligible patients (73.6%) consented and were randomized (256 eRAPID:252 UC). No significant effect of eRAPID on FACT-PWB was found at 18 weeks (mean difference 0.20; 95% CI - 0.81, 1.20;  $p = 0.699$ ). There was a benefit at 6 and 12 weeks (1.08; 95% CI 0.12, 2.05;  $p = 0.028$  & 1.01; 95% CI 0.05, 1.98;  $p = 0.039$ ). **Responder analysis:** eRAPID patients experienced less clinically relevant symptom deterioration at 6 and 12 weeks (deterioration at 12 weeks was 47.5% in eRAPID patients vs 56.3% UC). There were no differences for admissions/chemotherapy delivery. eRAPID patients reported better self-efficacy ( $p = 0.007$ ) at 18 weeks, and better QOL EQ5D-VAS scores at 12 ( $p = 0.030$ ) and 18 ( $p = 0.009$ ) weeks. **Intervention fidelity:** Patient adherence to weekly symptom reporting varied between 71.8% in week 1 to 58.1% week 18 (average 64.7%). 3314 online reports were completed, median per patient 14.0 (range 0–117); emergency alerts were activated in 29/3314 cases (0.9%), self-management advice 2714/3314 (81.9%). **Post hoc analyses** showed high patient adherence was associated with clinician's use of the data, high baseline FACT-PWB and older age. High adherence patients had better FACT-PWB scores at 12 weeks. **Conclusion:** eRAPID improved symptom control early at 6 and 12 weeks during chemotherapy and supported patient self-efficacy, without increasing hospital workload. Engaging both patients and clinicians is important for the intervention success.

## (2) Providing meaningful interpretation of a performance outcome measure by co-calibration with a patient-reported outcome: illustration with widely used multiple sclerosis measures

Antoine Regnault, PhD, Modus Outcomes, Lyon, France; Juliette Meunier, MSc, Modus Outcomes, Lyon, France; Anna Ciesluk, MPH, Modus Outcomes, Cambridge, Massachusetts, United States; Bing Zhu, Biogen, Cambridge, Massachusetts, United States; Wenting Cheng, Biogen, Cambridge, Massachusetts, United States; Li Zhu, Biogen, Cambridge, Massachusetts, United States; Chunlei Ke, Biogen, Cambridge, Massachusetts, United States, Patrick Marquis, Modus Outcomes, Cambridge, Massachusetts, United States

**Aims:** Performance Outcome (PerfO) measures are based on tasks performed by patients in a controlled environment, making their meaningful interpretation challenging to establish. This research illustrates how co-calibrating a PerfO measure with a patient-reported outcome (PRO) measure of the same target concept can provide useful guidance for its meaningful interpretation. **Methods:** Co-calibrating the PerfO measure with a PRO measure in a Rasch measurement theory framework involves two steps: (1) Running the Rasch model with the PRO items only to create the reference metric; and (2) Adding the PerfO measure to the Rasch model with fixed parameters for the PRO items. The PerfO measure results can then be interpreted using the content of the PRO items. Two widely used PerfO measures in multiple sclerosis (MS) of walking and manual ability, the Timed 25-Foot Walk test (T25FW) and the 9-Hole Peg Test (9HPT) respectively, were co-calibrated with two PRO measures of closely related concepts, the MS walking scale—12 items (MSWS-12) and the ABILHAND, using the data of 2,043 subjects from 5 clinical trials in MS. **Results:** PRO metrics were successfully obtained for walking (MSWS-12) and manual ability (ABILHAND). T25FW and 9HPT measurement results were transformed into categorical variables and were included in the Rasch model of the MSWS-12 and ABILHAND respectively. The probabilistic relationship between the PerfO measures and the PRO metric (item characteristic curve) was used to express the PRO item estimates in the unit of the PerfOs. This provided a way to characterize how different PerfO results correspond to different typical responses to the PRO items. Some weaknesses of the PRO metrics (gaps in the coverage of the walking ability continuum by MSWS-12 items; absence of ABILHAND items to discriminate subjects with good manual ability) hindered the ability to generate meaningful interpretation on the full range of possible PerfO measurement results. **Conclusion:** Co-calibration of PerfO and PRO measures provides a way to translate the PerfO measure results in terms of concrete, meaningful PRO terms. This approach requires PRO instruments with good measurement performance, especially in terms of breadth of coverage of the targeted concept.

## (3) Value Based Healthcare—a Welsh (UK) approach to delivering what matters to patients and clinicians

Sarah Puntoni, Cardiff and Vale UHB, Cardiff, United Kingdom; Sally Lewis, Aneurin Bevan University Health Board and Welsh Government, Cardiff, United Kingdom; Kathleen Withers, CEDAR, Cardiff and Vale UHB, Cardiff, United Kingdom; Amanda Willacott, Cardiff and Vale UHB, Cardiff, United Kingdom; Sally Cox, NHS Wales Informatics Service, Cardiff, United Kingdom

**Aims:** The Welsh Value Based Healthcare (VBHC) Programme aims to improve health outcomes by involving citizens in their care by asking them to complete PROMs. Combined with other HealthData, this will allow us to create a data-driven system which can provide timely information to citizens, clinical teams and organizations to inform decision-making in a way that is evidence based and financially sustainable. VBHC seeks to bring data to life for clinicians in the clinical environment. Providing access to PROMs data in

the electronic health record in graphic format for easy interpretation and use; developing data dashboards combining multiple data sources (see Fig. 1) to understand what matters to people and brings greatest value to the system. **Methods:** A National PROM platform has been developed with approx. 800 forms collected weekly across 30 nationally agreed pathways. The national digital architecture has been opened through programming developments, information governance frameworks and data standards to allow data to flow freely. **Results:** Growing access to PROMs and other data is providing the springboard to accelerate service transformation and testing new models of care. Electronic PROM collection is now active in all health boards and data can successfully be linked for use in decision-making. Data visualization tools are being developed and rolled out to make it easier for clinicians to view and use the data during consultations via the electronic patient record. Data dashboards allow different data sources to be viewed in one place providing local and national views and access to patient identifiable level data. **Conclusion:** Using PROMs as a key enabler, the program is transforming how the health service communicates with patients to facilitate shared decision-making and a co-produced approach to service delivery. VBHC is evidence of how a national approach to at-pace electronic PROMs collection that is clinically driven and allows easy access to data can help achieve service transformation. System change is already visible, with PROMs based virtual follow-up and triaging of referrals, Value Based procurement and a more mature discussion around variation and effectiveness already evident.

Figure 1



Figure 2



## (4) A meta-analysis of response shift effects in patient-reported outcomes studies

Tolu Sajobi, PhD, University of Calgary, Calgary, Alberta, Canada; Amelie Anota, PhD, University of Besancon, Besancon, France; Oluwaghohunmi Awosoga, PhD, University of Lethbridge, Lethbridge, Alberta, Canada; Oluwaseyi Lawal, MSc, University of Calgary, Calgary, Alberta, Canada; Carlijn Oomen, University of Amsterdam, Amsterdam, Netherlands; Anita Brobbey, MSc, University of Calgary, Calgary, Alberta, Canada; Mathilde G. E Verdam, PhD, University of Leiden, Leiden, Netherlands; Richard Sawatzky, PhD, Trinity Western University, Langley, British Columbia, Canada; Lisa Lix, PhD, University of Manitoba, Winnipeg, Manitoba, Canada

**Aims:** Response shift (RS), a change in the meaning of an individual's self-evaluation, is known to influence the assessment of longitudinal

change in patient-reported outcomes (PRO) studies. While an earlier meta-analysis reported small RS effects in studies that adopted then-test, a recent scoping review has reported increased methodological and empirical investigations of RS in PRO studies. This study aims to conduct a more comprehensive meta-analysis on reported RS effects in PRO studies that adopted then-test or latent-variable methods. We will quantify RS effects sizes in PRO studies and investigate the factors that explain variations in the reported effect sizes. **Methods:** This current study is part of an international collaborative initiative focused on synthesizing current empirical and theoretical research. PUBMED, CINHAHL, EMBASE, Medline, ProQuest, PsycINFO, and Web of Science were searched to identify English-language published articles that used the then-test, structural equation modeling (SEM), or item response theory (IRT) to test for RS until June 2019. Data on type of PRO measure, PRO domains, RS method, type of RS detected, study design, study quality, and reported RS effect sizes were extracted from the included articles. A random-effects meta-regression was used to evaluate the impact of these characteristics on variations in reported RS effect sizes. **Results:** Of the 1350 articles identified, 121 articles met the inclusion criteria. While 63 (52.1%) articles adopted the then-test, 58 (47.9%) articles adopted the SEM, and 3 (2.5%) articles adopted the IRT to test RS. For articles that adopted then-test, the reported recalibration RS effect sizes ranged between  $-1.16$  and  $2.11$ . Of the 58 articles that adopted SEM articles, only 28 reported RS effect sizes, which ranged between  $-0.48$  and  $0.88$  across the five domains. Meta-regression results will be available and presented at ISOQOL conference **Conclusion:** Our study will highlight the variation in type and magnitude of RS effects and will shed more light on important factors that explain this variation. Knowledge about when to expect RS can inform future planning of longitudinal PRO studies, guide the selection of PROs for specific populations, and alert clinicians to the potential for RS in certain PRO domains.

## Symposium Sessions

### **Symposium 1: Extending the QALY (E-QALY): developing a new international preference-based outcome measure for health and quality of life**

**Moderator:** John Brazier, PhD, School of Health and Related Research, The University of Sheffield, Sheffield, United Kingdom

**Discussant:** Madeleine King, Cancer Australia Chair in Cancer Quality of Life, School of Psychology, University of Sydney, Sydney, New South Wales, Australia

Common preference-based outcome measures (EQ-5D, SF-6D etc.) value changes in health-related quality of life. As such they have limited ability to capture the impacts on carers' quality of life, or important outcomes from interventions that also impact on broader non-health aspects of quality of life, e.g., social care for the elderly. Measures can be used that are developed specifically for carers (e.g., CarerQoL) or for those receiving social care (e.g., ASCOT), however, this is problematic when valuing the overall impact of an intervention in economic evaluation since it is not possible to combine outcomes across measures and risks double counting those impacts.

The extending the QALY project, funded by the UK MRC and EuroQoL Foundation aimed to develop a broad generic measure of quality of life for use in economic evaluation across health care, social care and public health. The study had 5 development stages: (1) establishing the domains; (2) generating a long list of items; (3) testing the face validity of the items; (4) psychometric testing of the items; (5) valuation of selected items. Work was undertaken in six countries: England, Argentina, Australia, China, Germany, and USA. Extensive consultation with stakeholders, experts, and the public was used to refine the domain structure and identify high performing items to measure each selected domain.

The symposium will discuss the opportunities and challenges of developing a new international preference-based measure. John will provide the rationale for developing the measure and an overview of the project. Clara will present findings from Stages 1 and 2 where the domains and items for the measure were identified. Zhihao will focus on the face validation results with a particular focus on China (Stage 3). Ole will discuss the psychometric testing with particular focus on Germany (Stage 4). Donna will present the findings from the valuation stage. Finally, Madeleine King will start the discussion by providing her view on the approach taken and the resulting measure.

## Individual presentations

### **Extending the QALY (E-QALY): developing a new generic preference-based measure for use with patients, carers and those needing care—an overview**

**John Brazier, PhD,** School of Health and Related Research, The University of Sheffield, Sheffield, United Kingdom

**Aims:** The aim of the E-QALY project was to develop a broad generic measure of quality of life valued on the zero to one scale necessary to calculate Quality Adjusted Life Years (QALYs) for use in economic evaluation across health, social care and public health. **Methods:** The project brings together stakeholders in industry, policy-makers (across health and social care) and members of the public in five stages. (1) Establishing domains for the instrument which drew on existing qualitative literature. (2) Generating a long list of potential items for each domain/subdomain based on item generation criteria. (3) Testing the face validity of the candidate items. (4) Psychometric testing of items that had face validity. Stages 3 and 4 were undertaken in 6 countries: Australia, Argentina, China, Germany, USA and UK with patients, social care users, carers and the general public. Psychometric and face validity results were used to inform the choice of items for the final measure and classification system. (5) Valuation was undertaken using an adapted version of EuroQoL group portable valuation technology (EQ-PVT) which includes time trade-off (TTO) and discrete choice experiments (DCE). **Results:** The qualitative review identified 27 domains that were considered important. Approximately 100 items were identified or generated and taken forward to face validity and 66 were included in the psychometric survey. Cross-country comparisons were useful in informing which items could be used. Consultation on the results from Stages 3 and 4 resulted in a 25 item measure, with 9 items taken forward into a classifier for valuation. The modified EQ-PVT generated utility weights can be used for economic evaluation. **Conclusion:** The new measure covers aspects that are important for patients, social care users and carers and the items have been tested in different countries. All the stages involved bringing together different stakeholder views on methods to results in order to develop a measure that was fit for purpose. Challenges included incorporating results from different countries. Further validation work is needed on responsiveness to change.

### **E-QALY: identifying dimensions and items for a generic preference-based measure**

**Clara Mukuria,** University of Sheffield, Sheffield, United Kingdom; **Janice Connell,** University of Sheffield, Sheffield, United Kingdom; **Jill Carlton,** University of Sheffield, Sheffield, United Kingdom; **Tessa Peasgood,** University of Sheffield, Sheffield, United Kingdom; **Alison Scope,** University of Sheffield, Sheffield, United Kingdom; **Mark Clowes,** University of Sheffield, Sheffield, United Kingdom; **Karen Jones,** University of Kent, Kent, United Kingdom; **Stacey Rand,** University of Kent, Kent, United Kingdom; **John Brazier,** University of Sheffield, Sheffield, United Kingdom

**Aims:** To identify the themes to inform the content of a new generic measure that can be used in economic evaluation and in routine use across health, social care and public health, based on the views of users and beneficiaries of these services. **Methods:** A targeted qualitative review of the literature with a systematic search was undertaken to identify views of patients, informal carers and those who receive care support (i.e., social care) as well as qualitative evidence on the development of measures used in these populations. Framework analysis and synthesis was undertaken based on a conceptual model. Criteria were developed to inform the selection of themes and subsequent items. Items were identified from existing measures, item banks and new ones were also generated. **Results:** The search yielded 137 qualitative reviews on QoL for health conditions were identified while there were 15 for informal care and 17 for social care. Targeted extraction and synthesis resulted in seven high-level themes/domains (with subthemes/domains): feelings and emotions (sadness, anxiety, enjoyment, vulnerability etc.); cognition; activity (leisure, activities of daily living, communication, mobility, vision, meaningful activity); self-identity (self-esteem, dignity/respect); relationships and social connections (both positive and negative aspects); ‘coping, autonomy and control’; and physical sensations (pain, fatigue, discomfort). A long list of items ( $n \sim 100$ ) covering the themes and subthemes was identified. **Conclusion:** The breadth of conditions covered allowed for core generic aspects that went beyond health to be identified. Apart from physical sensations, most of the other themes and subthemes were relevant across almost all of the sample health conditions, for social care, and for informal carers. Items were generated for all subthemes but for some, it was challenging to identify/create items that could be asked of relevant groups and that met our criteria for appropriate items to use in a generic preference-based measure.

#### **E-QALY: psychometric testing and selecting the items for a new generic preference-based measure**

Tessa Peasgood, School of Health and Related Research, University of Sheffield, Sheffield, United Kingdom; Clara Mukuria, School of Health and Related Research, University of Sheffield, UK, Sheffield, United Kingdom; **Ole Marten, M.Sc.**, School of Public Health, Department of Health Economics and Health Care Management, Bielefeld University, Bielefeld, Germany; Simone Kreimeier, MPH, Bielefeld University, Department for Health Economics and Health Care Management, Bielefeld, Germany; Lidia Engel, Deakin Health Economics, School of Health and Social Development, Deakin University, Australia, Melbourne, Australia; Zhihao Yang, Health Services Management Department, Guizhou Medical University, China, Guiyang, China; Luz Gibbons, Institute for Clinical Effectiveness and Health Policy (IECS-CONICET), Buenos Aires, Argentina, Buenos Aires, Argentina; Andrea Monteiro, Department of Pharmacy Systems, Outcomes and Policy, College of Pharmacy, University of Illinois at Chicago, Illinois, USA, Chicago, Illinois, United States; Maja Kuharic, Department of Pharmacy Systems, Outcomes and Policy, College of Pharmacy, University of Illinois at Chicago, Illinois, USA, Chicago, United States; Brendan Mulhern, Centre for Health Economics Research and Evaluation, University of Technology Sydney, Australia, Sydney, Australia; Simon Pickard, Department of Pharmacy Systems, Outcomes and Policy, College of Pharmacy, University of Illinois at Chicago, Illinois, USA, Chicago, United States; Nan Luo, Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore; Federico Augustovski, Institute for Clinical Effectiveness and Health Policy (IECS-CONICET), Buenos Aires, Argentina, Buenos Aires, Argentina; Maria Belizan, Institute for Clinical Effectiveness and Health Policy (IECS-CONICET), Buenos Aires, Argentina, Buenos Aires, Argentina; Wolfgang Greiner, Bielefeld University,

Department for Health Economics and Health Care Management, Bielefeld, Germany; John Brazier, School of Health and Related Research, University of Sheffield, UK, Sheffield, United Kingdom

**Aims:** The aim was to quantitatively test the items from a psychometric perspective to further inform the selection of items for both a profile measure and a health state classification system for valuation. **Methods:** Self-administered paper and online surveys were conducted in six countries (Argentina, Australia, China, United Kingdom, Germany and the USA) with individuals with various physical and mental health conditions, carers and social care users to test the psychometric properties of the proposed domains and 66 draft items using a standardized protocol. Analysis included examination of response distribution, sensitivity to known group differences, testing of dimensionality by means of exploring item correlations, exploratory and confirmatory factor analysis as well as exploration of item response theory models. A consultation exercise was used to select items for the final measure. **Results:** Survey data from  $\sim 4500$  participants across all countries were available. German results ( $n = 496$ ) suggest that 4 items performed poorly while 11 items performed well over all psychometric analyses. Evidence on the remaining items was less conclusive with no clear pattern of superior or inferior psychometric item performance over items in the same (sub-)domain as items’ strengths and weaknesses across analyses methods varied. Internationally, a few items on self-care (‘activity’) seemed problematic for all language versions. There was variation in performance across country results for items on control (‘autonomy’) and items on happiness, hope and safety (‘feelings’). Items on the physical domain like pain, or sleep, concentration (‘cognition’) and meaningful activities (‘activity’) indicated a consistently good performance. Dimensionality was the same across most of the countries with the exception of China where anxiety, safety, anger, happiness, hope, coping, self-worth were combined. Consultation resulted in the final 25-item measure and 9-item classification system. **Conclusion:** Evidence from different countries provided in-depth knowledge on cross-cultural item performance although synthesizing the findings was challenging. The overall psychometric performance of individual items was very similar but the individual country data allowed item prioritization which was used to inform the final measure.

#### **E-QALY: face validity testing the items for a new generic preference-based measure**

Jill Carlton, University of Sheffield, Sheffield, United Kingdom; Janice Connell, University of Sheffield, Sheffield, United Kingdom; Tessa Peasgood, University of Sheffield, Sheffield, United Kingdom; Clara Mukuria, University of Sheffield, Sheffield, United Kingdom; Kristina Ludwig, University of Bielefeld, Bielefeld, Germany; Ole Marten, University of Bielefeld, Bielefeld, Germany; Simone Kreimeier, University of Bielefeld, Bielefeld, Germany; Lidia Engel, Deakin University, Melbourne, Australia; **Zhihao Yang, PhD**, Jinan University, Guangzhou, China; Maria Belizan, Institute for Clinical Effectiveness and Health Policy, Buenos Aires, Argentina; Rocio Rodriguez, Institute for Clinical Effectiveness and Health Policy, Buenos Aires, Argentina; Andrea Monteiro, University of Illinois at Chicago, Chicago, Illinois, United States; Maja Kuharic, University of Illinois at Chicago, Chicago, Illinois, United States; Brendan Mulhern, University of Technology, Sydney, Sydney, Australia; A. Simon Pickard, University of Illinois at Chicago, Chicago, Illinois, United States; Nan Luo, National University of Singapore, Singapore, Singapore; Federico Augustovski, Institute for Clinical Effectiveness and Health Policy, Buenos Aires, Argentina; Wolfgang Greiner, University of Bielefeld, Bielefeld, Germany

**Aims:** The aim was to qualitatively examine the content and face validity of items for the new measure. **Methods:** Semi-structured

cognitive interviews were conducted in six countries (Argentina, Australia, China, England, Germany and the USA) with individuals with various physical and mental health conditions, carers and social care users to test the content and face validity of the proposed domains and draft items using a standardized protocol. Different response options (frequency, severity, difficulty, agreement) were also tested. All items were translated into Spanish, Chinese and German by a single company who undertook forward and back translation with input from the respective teams in Argentina, China and Germany. Results: Interviews were undertaken with 170 participants across the six countries. Results in China ( $n = 30$ ) suggested that 5 items needed major modifications or to be dropped as they did not work well in China's context. Other items needed further information or clarification in order to allow accurate comprehension and completion. Some examples included in questions for added context were not considered helpful in China. Internationally, participants preferred simpler layouts of questions but wanted more information on context. There was no clear preference for response options. Questions related to dignity and communication had varied interpretations that were not always consistent with the target conceptual theme. Questions related to self-care were difficult for participants, particularly for carers. Some instructions such as those regarding recall periods were frequently ignored. Conclusion: Drawing on information from several countries identified a set of items that were suitable for taking forward to the psychometric survey although there were some challenges in terms of cultural relevance. The face validity interviews also allowed for some modifications to be undertaken both in terms of content and layout prior to the psychometric survey. Results from this stage were useful in prioritizing items for the final measure.

#### **E-QALY: valuing the items for a new generic preference-based measure using time trade-off and discrete choice experiments**

Clara Mukuria, PhD, University of Sheffield, Sheffield, United Kingdom; Tessa Peasgood, PhD, University of Melbourne, Melbourne, Australia; John Brazier, PhD, University of Sheffield, Sheffield, United Kingdom; **Donna Rowen, PhD**, University of Sheffield, Sheffield, United Kingdom; Richard Norman, PhD, Curtin University, Perth, Australia

**Aims:** A new classification system was developed that covers nine aspects of health and quality of life. The aim was to generate utility weights for the classification system to enable the measure to be used to generate utility values and Quality Adjusted Life Years (QALYs). **Methods:** EuroQoL Portable Valuation Technology (EQ-PVT) which uses time trade-off (TTO) and discrete choice experiments (DCE) to value EQ-5D-5L using a standardized protocol was modified for the new measure. Two stages of piloting were undertaken prior to the main valuation to ensure appropriateness of the protocol. First, qualitative interviews ( $n = 15$ ) were used to test whether participants could undertake the valuation of the new measure. Participants saw both the new measure and EQ-5D-5L states and were asked to compare valuation using the different descriptions. Their views on the task and the measure were recorded and used to modify EQ-PVT as appropriate. Second, computer-assisted personal interviews ( $n = 50$ ) were undertaken and analyzed to test the modified EQ-PVT protocol. The main study used this modified protocol ( $n = 500$ ) using a representative sample by age and gender from England. **Results:** Results from the pilot indicated that participants were able to undertake both TTO and DCE using the new measure. There were mixed views about the benefit of combining health and quality of life states: some participants thought the additional information was useful in helping to imagine what life would be like, while others felt that added information was overwhelming and made the tasks difficult. Pain, activities and depression were important while some participants

considered coping to be an overall assessment of the states. Data collection and assessment is on-going, and the presentation will summarize all results. **Conclusion:** The existing standardized protocol used for EQ-5D-5L has been piloted, modified and successfully applied to value the new measure. Results from the main valuation study provide utility weights that can be applied to generate utility values for use to generate QALYs in cost-utility analyses.

#### **Symposium 2: Do you know what I mean? Interpreting PRO/HRQoL scores at group and individual level**

**Moderator:** Corneel Coens, EORTC HQ, Brussels, Belgium

**Discussants:** Galina Velikova, PhD, MD, Leeds Institute of Medical Research at St James's, University of Leeds, Leeds Cancer Centre, Leeds, United Kingdom; Madeleine King, PhD, University of Sydney, Sydney, Australia; Submitted on behalf of EORTC Quality of Life Group

The use of patient-reported outcomes (PRO) to measure health-related quality of life (HRQoL) in medical research and practice has steadily increased over the years. With this growth, a wealth of data has become available but also an increasing need to draw meaningful conclusions from the observed results. Compared to many other endpoints, such as survival rates, treatment compliance or laboratory results, PROs may not have an inherent meaning attached, or that meaning may differ between different questionnaires or according to the context.

The methodology for evaluating HRQoL scores has expanded rapidly in response to this demand. On the one hand, interpretation can occur at the level of individual patient scores, either in isolation or over time. Or else, comparisons between groups may be the outcome of interest, requiring a different reading of the estimated differences. Increasingly more complex methods to appraise relevant thresholds and/or minimal important differences have led to more evidence-based interpretation guidance but it has not necessarily simplified, on the contrary. The era of simplistic rule-of-thumbs or one-size-fits-all cut points has been replaced by a catalogue of options. Understanding what interpretational rule to apply in which context is now more challenging than ever.

This symposium will consist of three presentations followed by a discussion session. The three presentations will focus on minimal important differences (both individual and group level), responder thresholds and reference values respectively. Each presenter will first explain the underlying methodology used to obtain their results and then will critically assess its purpose and required context. The focus will be on addressing what conclusions can or can not be drawn for each method. Examples from the field of oncology, all using the EORTC QLQ-C30 questionnaire, will facilitate comparisons between the different approaches. The discussants will then summarize the presented information and review each method's applicability to specific cases where decision-making based on PRO data is required. This will highlight the various advantages but also limitations of these interpretational techniques. The symposium will end with room for questions and discussion with the audience.

#### **Individual presentations**

##### **Interpretation of patient-reported outcome data at group level versus individual level; can we use the same clinical meaningful thresholds for both scenarios?**

**Jamme Musoro, PhD**, EORTC HQ, Brussels, Belgium; Corneel Coens, MSc, EORTC HQ, Brussels, Belgium; Madeline Pe, PhD, EORTC HQ, Brussels, Belgium; Andrew Bottomley, PhD, EORTC

HQ, Brussels, Belgium; Submitted on behalf of EORTC Quality of Life Group

**Aims:** PRO data are increasingly used to assess effects of a disease and its treatment across groups of patients (group level), as well as for monitoring and managing individual patients (individual level). Although the interpretation of PRO data at group and individual level is recognized as different concepts, with each having their own applications in clinical trials, the literature is not aligned on the topic of method selection for estimating group-level versus individual-level thresholds. Methods that were developed for group-level interpretation have commonly been applied to individual-level data. We will examine the principal differences between both methods and their application. **Methods:** We will review common available methods for estimating clinical meaningful thresholds for group-level interpretation of PRO results (anchor-based and distribution-based). We also will examine how group-level thresholds compare to those obtained via the receiver operating characteristic (ROC) curve method; often considered more appropriate for defining individual-level thresholds since estimates for the sensitivity and specificity of specific values are obtainable to assess individual misclassification errors. **Results:** Two important caveats apply to setting thresholds for use at individual level. First, not all group-level threshold values will translate into a score that is achievable for an individual because every scale of a PRO measure has a limited number of observable values. For example, single-item scales from the QLQ-C30 have only 4 possible values (0, 33, 66, and 100), resulting in a discrete range of change scores, while the multi-item scales have many more possible values and therefore more continuous-like change scores. For single-item scales in particular, it may be necessary to select values on either side of the group-level threshold for individual thresholds, with selection of either the higher or lower value depending on clinical context. The second caveat is that individual thresholds must be set above bounds of measurement error to avoid false positive changes that might trigger unjustified clinical actions. **Conclusion:** Group-level thresholds can be a useful starting point for defining cut-offs for individual-level changes that are clinically meaningful, but this should be done with caution. Recognizing the principal differences between each method will avoid unintended consequences.

#### **Development and use of thresholds for clinical importance to facilitate interpretation of scores from patient-reported outcome measures**

**Johannes Giesinger, PhD**, Medical University of Innsbruck, Innsbruck, Austria; Submitted on behalf of EORTC Quality of Life Group

**Aims:** Patient-reported outcome (PRO) measures have mostly been employed as outcome measures in cancer clinical trials, but more recently these measures are also used for patient monitoring in daily clinical practice. Routine PRO monitoring and screening allow the timely identification of symptoms and functional impairments and can provide important information on the impact of specific interventions over time. Routine PRO monitoring has demonstrated important clinical benefits including better symptom management and improved survival rates. However, difficulties with interpreting scores on the abstract metrics of PRO measures are a major barrier. The aim of this presentation is to explain and illustrate the development of thresholds for use in PRO monitoring as applied to the QLQ-C30 questionnaire. **Methods:** A number of approaches have been used to establish thresholds for PRO measures, including thresholds based on the wording of response categories, on score distributions in specific patient populations, or on external criteria such as need for care or disease prognosis. Reliable and relevant dichotomization at the

individual patient level allows for threshold estimation using Receiver Operator Characteristic (ROC) analysis. This approach yields sensitivity and specificity of specific thresholds, which provides important information regarding over- and/or underidentification of clinically important problems. **Results:** Recently, such thresholds for clinical importance were established for the EORTC QLQ-C30 relying on composite external criteria that reflect different aspects of clinical importance such as perceived burden, daily limitations and need for help. The obtained estimates were integrated into software for routine PRO monitoring in daily practice to improve the screening process by concise graphical presentation of individual patient responses (e.g., use of color-coding or reference lines). This made PRO scores more actionable and allowed to link these results to clinical decision-making. The integration of the above-mentioned thresholds for the EORTC QLQ-C30 was successfully applied in a hematological outpatient unit at the Medical University of Innsbruck. **Conclusion:** Routine PRO monitoring and screening is feasible but requires appropriate actionable thresholds. A clear understanding of the methodological basics is crucial for the understanding of their meaning, which depends on the underlying assumptions.

#### **Norm data for the EORTC QLQ-C30 health-related quality of life questionnaire based on 15,386 persons from the European general population, Canada, and the USA**

**Sandra Nolte, PhD**, ICON plc, Berlin, Germany; Gregor Liegl, PhD, MD, Charité - Universitätsmedizin Berlin, Berlin, Germany; Morten Petersen, PhD, Bispebjerg Hospital, Copenhagen, Denmark; Neil Aaronson, PhD, The Netherlands Cancer Institute, Amsterdam, Netherlands; Anna Costantini, PhD, MD, Sant'Andrea Hospital Sapienza, University of Rome, Rome, Italy; Peter Fayers, PhD, University of Aberdeen, Aberdeen, United Kingdom; Mogens Groenvold, PhD, University of Copenhagen, Copenhagen, Denmark; Bernhard Holzner, PhD, Medical University Innsbruck, Innsbruck, Austria; Colin Johnson, PhD, MD, University of Southampton, Southampton, United Kingdom; Georg Kemmler, PhD, Medical University Innsbruck, Innsbruck, Austria; Krzysztof Tomaszewski, PhD, Ignatianum Academy, Krakow, Poland; Annika Waldmann, PhD, University of Luebeck, Luebeck, Germany; Teresa Young, PhD, Mount Vernon Hospital, Middlesex, United Kingdom; Matthias Rose, PhD, MD, Charité—Universitätsmedizin Berlin, Berlin, United Kingdom; Submitted on behalf of EORTC Quality of Life Group

**Aims:** Norm data obtained from the general population as well as reference values obtained from patients facilitate the interpretation of patient-reported outcomes (PRO) data. It allows to benchmark observed outcomes from either an individual or from a group of subjects to identify relevant results. The aim of this study was to systematically collect European norm data for the widely used EORTC QLQ-C30 health-related quality of life questionnaire. **Methods:** QLQ-C30 data were collected across 11 European Union (EU) countries and also Russia and Turkey. In addition, general population norm data were collected in Canada and USA for comparative purposes. The anticipated sample size was at least 1000 subjects for each of the 15 included countries. Representative samples were stratified by gender (male, female) and age groups (18–39, 40–49, 50–59, 60–69, ≥ 70 years) with at least 100 subjects in each of the substrata. After applying weights based on the United Nations' population distribution statistics, QLQ-C30 domain scores were calculated to generate a “European QLQ-C30 Norm” based on the EU countries. **Results:** A total of 15,386 respondents were recruited via online panels and completed the survey. For the EU sample, most QLQ-C30 domains showed differences by sex/age, with men scoring somewhat higher functioning and lower symptoms compared with women, while age effects varied between QLQ-C30 domains.

Substantially larger differences were seen in inter-country comparisons, with Austrian and Dutch respondents consistently showing higher functioning and lower symptoms compared with British and Polish respondents. Conclusion: This study is the first to systematically collect QLQ-C30 general population norm data across Europe and North America using a consistent data collection method. These general population norm data along with reference values greatly facilitate score interpretation of PRO data collected in various contexts such as clinical research but also clinical practice where PRO data are collected as part of routine care. Adjusted norm data for a specific setting or population can be obtained by applying appropriate weights for sex, age and country to account for selection effects. The dataset is available upon request, providing an invaluable resource to PRO researchers across the globe.

### Symposium 3: Using an implementation science approach to implement and evaluate Patient-Reported Outcome Measures (PROM/PREM) initiatives in routine care settings

**Moderator:** Angela Stover, PhD, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States; Submitted on behalf of the ISOQOL PROMs/PREMs in clinical practice implementation science work group

**Discussant:** Maria Santana, PhD, University of Calgary, Alberta, Canada

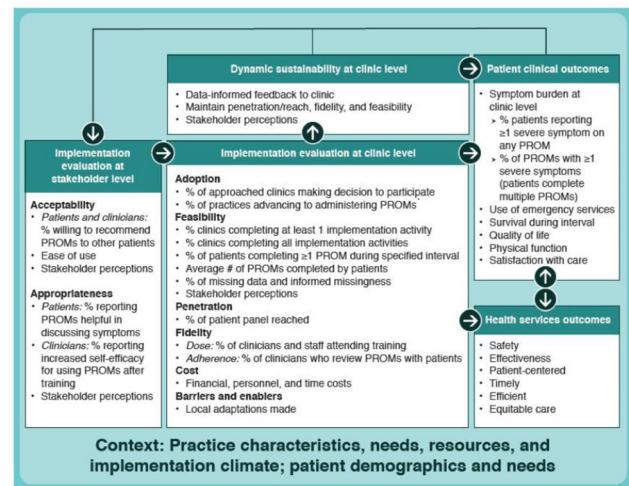
Patient-reported outcome and experience measures (PROMs/PREMs) are well established in research for many health conditions, but barriers persist for implementing them in routine care. Implementation science (IS) offers a potential way forward, but its application has been limited for PROMs/PREMs. IS is the systematic study of methods to integrate evidence-based practices into care settings. Part of IS's appeal are the theories and frameworks guiding the translation process from research to practice.

In this symposium, we will compare similarities and differences for five widely used IS frameworks and their applicability for implementing PROMs/PREMs in routine care through four case studies. Three case studies use theory to implement PROMs: (1) pain clinics in Canada; (2) oncology clinics in Australia; and (3) pediatric/adult clinics for chronic conditions in the Netherlands. One case study is using theory to plan PREMs implementation in primary care clinics in Canada. We compare case studies on theories, barriers, enablers, and implementation support strategies.

IS approaches are largely harmonious with PROMs/PREMs implementation in routine care, although no single framework or theory appears to fully capture the nuances for PROMs/PREMs across clinical contexts. Across case studies in different countries and health conditions, barriers were remarkably consistent, including technology limitations, uncertainty about benefits and concerns about negative impacts, and competing demands within established clinical workflows. A unique aspect of our study is case studies showed more variation for PROM/PREM enablers in clinics than barriers, indicating the potential for tailored solutions. A common enabling factor was designing a technology system with automated features (e.g., patient reminders) and rapid access to results. More unique enablers capitalized on local resources, such as peer champions, advertising campaigns, and providing clinics with implementation funding.

Common implementation support strategies included engaging stakeholders, changing infrastructure, providing interactive assistance, and training clinicians and staff. Evaluation in case studies was inconsistent, and thus we will present IS metrics specific to evaluating PROM/PREM implementation initiatives. Increasing the use of IS in PROM/PREM implementation studies will help advance our collective understanding of causal mechanisms to better understand how,

why, and in what circumstances IS frameworks and implementation strategies produce successful PROM/PREM implementation.



### Individual presentations

#### One vehicle, too many wheels: towards consistent use of theory for guiding the implementation of PROMs/PREMs in routine clinical practice

**Caroline Potter, PhD,** University of Oxford, Oxford, United Kingdom; **Angela Stover, PhD,** University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States; **Joanne Greenhalgh, PhD,** University of Leeds, Leeds, United Kingdom; Submitted on behalf of the ISOQOL PROMs/PREMs in Clinical Practice Implementation Science Work Group

**Aims:** Implementation science (IS) theories and frameworks can be used to guide the translation process from research evidence to clinical practice, but their use has been limited for implementing patient-reported outcomes and experience measures (PROMs/PREMs) in routine care settings. The high volume and diverse origins of IS theory, and limited numbers of theoretically informed studies, have led to disconnected and repeated implementation efforts. By identifying commonalities in IS theoretical approaches, we aim to develop a set of generalized guidelines for implementing PROMs/PREMs in routine care settings. **Methods:** We compare four widely used IS theoretical approaches through four case studies on implementing PROMs/PREMs in clinics treating pain, cancer, and chronic health conditions. **Results:** Case studies used descriptive IS frameworks categorizing barriers, enablers, and implementation support strategies. Three case studies used the Consolidated Framework for Implementation Research (CFIR), one combined CFIR with the Theoretical Domains Framework (TDF), and one used the integrated Promoting Action Research in Health Services (i-PARIHS) framework. We linked these descriptive frameworks to Normalization Process Theory (NPT), which describes four causal mechanisms of the implementation process whereby change becomes routine behavior. Table 1 shows how NPT's causal mechanisms highlight general strategies for implementing PROMs/PREMs in routine care settings. **Conclusion:** Implementation Science theoretical approaches can be used to understand why implementation succeeded, to systematize barriers, and to develop context-appropriate implementation support strategies. Consistent use of theory could yield a more systematic approach to implementing PROMs/PREMs in routine care, facilitating shared learning and coordinated implementation efforts.



Implementation mechanisms	Questions raised	Actions for implementation	Strategies used in case studies
Coherence (sense-making)	What is the work? Why should we do it?	Define the work and its value	<ul style="list-style-type: none"> <li>Stakeholder engagement</li> <li>Provide evidence</li> </ul>
Cognitive participation	Who will do what to get the work done?	Identify and mobilize communities of practice	<ul style="list-style-type: none"> <li>Training workshops</li> <li>Workflow re-design</li> <li>Implementation support team</li> </ul>
Collective action	What helps or hinders our ability to get the work done together?	Document the details of context-specific barriers and enablers	<ul style="list-style-type: none"> <li>Context assessments</li> <li>Technology support</li> <li>Practice facilitator</li> </ul>
Reflexive monitoring	What were the impacts of the work?	Refresh strategy in light of experience	<ul style="list-style-type: none"> <li>Annual evaluation meetings</li> <li>Audit and feedback</li> </ul>

**Table 1** Linking theory on causal mechanisms for implementation to strategies observed in case studies

### Implementation of PROMs in routine clinical care within an integrated pain network

**Sara Ahmed, PhD**, McGill University, Montreal, Quebec, Canada; **Diana Zidarov, PhD**, McGill University, Montreal, Quebec, Canada; **Owis Eilayyan, PhD**, McGill University, Montreal, Quebec, Canada; **Regina Visca, BEng, MBA**, McGill University Health Center, Montreal, Quebec, Canada

**Aims:** The objective of this study was to develop an implementation and evaluation plan, guided by implementation science frameworks, for integration of electronic patient-reported outcome measures (ePROMs) across an integrated chronic pain network that includes primary, rehabilitation, and hospital-based care. A secondary objective was to present preliminary results on the acceptability, adoption, usability, and feasibility of the ePROM system after 6 months of implementation. **Methods:** The Theoretical Domains Framework (TDF) was used to identify potential barriers and enablers to the use of ePROMs by primary care clinicians. In rehabilitation and tertiary care, the Consolidated Framework for Implementation Research (CFIR) was used to guide the identification of determinants of implementations, through observation of workflow, patient and clinician surveys, and clinician interviews. A mixed-method concurrent design comprising a quantitative and a qualitative analysis was used. The results were reviewed by a steering committee to iteratively inform the ePROM implementation plan. The Proctor framework of evaluation was used to guide the development of an evaluation plan for the implementation of ePROMs in the integrated chronic pain network. **Results:** Both frameworks provided similar results with respect to healthcare provider knowledge, behavior and experience interpreting PROM scores. The TDF and CFIR frameworks differed in identifying organizational-level determinants. The resultant implementation plan was structured around the adoption of PROMs to inform individual treatment planning and quality improvement. The evaluation plan focused on implementation and impact outcomes to evaluate the ePROM intervention. We will present results from the Acceptability of Intervention Measure, percent patients and clinicians using the system (adoption), Feasibility of Implementation and the End-User Computing Satisfaction Questionnaire (acceptability). **Conclusion:** The TDF and CFIR guided the development of a multi-component knowledge translation and training intervention that will address multiple gaps and barriers to implementation of PROMs across the integrated network. In addition to informing individual patient care, ePROMs will be an important component of a Learning Healthcare System to contribute outcomes that matter to patients when comparing the effectiveness of interventions and to inform health service provision.

### Using the Integrated framework Promoting Action Research in Health Services (iPARIHS) Framework to study implementation of PROMs into oncology care

**Natasha Roberts, PhD**, Queensland University of Technology, Kelvin Grove, Australia; **Natasha Roberts**, Queensland University of Technology, Brisbane, Australia; **Angela Stover**, University of North Carolina, Chapel Hill, North Carolina, United States; **Kimberly Alexander**, Queensland University of Technology, Brisbane, Australia; **David Wyld**, Royal Brisbane and Women's Hospital, Brisbane, Australia; **Monika Janda**, University of Queensland, Brisbane, Australia

**Aims:** Randomized controlled trials demonstrate improved clinical and health service outcomes in oncology care when clinicians review PROMs and address patients' expressed concerns. These results have driven a growing interest in using PROMs routinely in day to day oncology care. However, implementation is challenging, with calls for a better understanding of how to incorporate research findings into the clinical context. For this reason, we aimed to integrate symptom PROMs into an oncology outpatient setting using the "Integrated framework for Promoting Action on Research Implementation in Health Services" (iPARIHS). **Methods:** The minimum components of the intervention were PROMs completion by patients in the waiting room and clinicians acknowledging and reviewing PROMs with patients during the visit. The three research phases included pre-implementation, implementation, and evaluation. Each phase informed the next to describe, measure, and evaluate a pilot implementation strategy informed by iPARIHS. The active ingredient in iPARIHS is facilitation (implementation support strategies), with three main factors influencing implementation: (1) context: characteristics of the setting where implementation took place; (2) recipients: characteristics of anyone who interacted with the implementation process; and (3) the innovation itself: characteristics of PROMs and intervention design). Barriers to PROM completion and/or clinician acknowledgement rates were identified and addressed using the iPARIHS framework. **Results:** By measuring and evaluating implementation in short iterative cycles, the design of the intervention was refined into workflows to ensure optimal patient PROM completion rates and staff acknowledgement rates. Staff perceptions of acceptability and appropriateness during pre-implementation and post-implementation were that symptom PROMs were useful for clinical care. Clinical outcomes showed a statistically significant increase in symptom detection ( $p < 0.01$ ) and an increase in the use of current clinical pathways for managing symptoms ( $p < 0.05$ ). **Conclusion:** The iPARIHS framework was useful in the design, implementation, and evaluation of a PROM implementation initiative in routine oncology care. However, to achieve sustainability of the intervention, continued emphasis on facilitation was necessary, and implementation took much longer than anticipated. This pilot study identified key elements of success to be considered in a future large scale implementation.

### Using Implementation Science to Inform Integration of Electronic Patient-Reported Experience Measures (ePREMs) into healthcare quality improvement

**Kimberly Manalili, PhD Candidate**, University of Calgary, Calgary, Alberta, Canada; **Maria Jose Santana, MPharms, PhD**, University of Calgary, Calgary, Alberta, Canada

**Aims:** Healthcare systems are increasingly prioritizing the patient perspective in quality improvement. Collecting and monitoring the information from patients through patient-reported experience measures (PREMs) about the quality of care they receive is necessary to track changes in healthcare quality, stimulate innovation, and enhance

person-centered care. The objective of this presentation is to discuss the use of implementation science theories, models, and frameworks to assess the integration of the electronic collection of PREMs (ePREMs) in healthcare quality. Methods: To assess potential knowledge-to-practice gaps in implementing ePREMs in primary care in Alberta, the overarching implementation model that will be used is the Knowledge to Action Cycle. An integrated knowledge translation approach will ensure ongoing engagement of key stakeholders (primary care providers, quality improvement leads, and patients) throughout the study. The ePREM implementation will be informed by the identification of barriers and enablers to implementation through interviews with key stakeholders, using the theory-based Consolidated Framework for Implementation Research (CFIR). The CFIR brings an organizational perspective providing an opportunity to explore the intervention characteristics, the inner and outer context of implementation. Identified barriers and facilitators to ePREM implementation will be mapped to evidence-based implementation strategies and prioritized by stakeholders. The RE-AIM framework will be used to guide the evaluation of ePREM implementation outcomes after 6 months of implementation by assessing: Reach, Effectiveness, Adoption, Implementation, and Maintenance (sustainability). Results: This ongoing research has successfully engaged patient engagement stakeholders across Canada, through the provincial Strategy for Patient-Oriented Research networks and primary care stakeholders in Alberta. Consultations with stakeholders affirm the importance of evaluating the integrated knowledge translation approaches, as well as the implementation outcomes. Conclusion: This presentation describes how theoretical and practical considerations based on implementation science approaches could help address important ePREM implementation challenges and promote the successful uptake and use of ePREMs for quality improvement in healthcare.

#### **Implementation of the KLIK PROM portal using the Consolidated Framework for Implementation Research (CFIR) retrospectively**

**Hedy van Oers, PhD**, Emma Children's Hospital Amsterdam UMC, Amsterdam, Netherlands; Lorynn Teela, MSc, Emma Children's Hospital Amsterdam UMC, Amsterdam, Netherlands; Sasja Schepers, PhD, Princess Máxima Center for Pediatric Oncology, Utrecht, Netherlands; Martha Grootenhuis, PhD, Princess Máxima Center for Pediatric Oncology, Utrecht, Netherlands; Lotte Haverman, PhD, Emma Children's Hospital Amsterdam UMC, Amsterdam, Netherlands; Submitted on behalf of the ISOQOL PROMs and PREMs in Clinical Practice Implementation Science Group

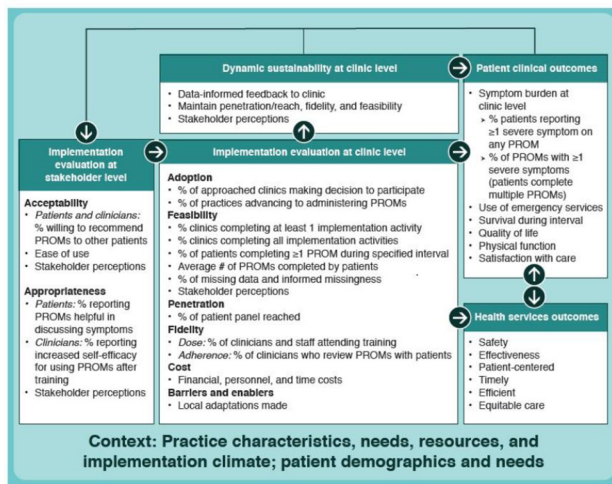
**Aims:** The KLIK Patient-Reported Outcome Measure (PROM) portal is an evidence-based intervention implemented in clinical practice in > 25 Dutch hospitals for patients (children and adults) who regularly visit the outpatient clinic. Implementation science frameworks can be used to understand why implementation succeeded or failed, to structure barriers and enablers, and to develop implementation strategies to overcome barriers. This symposium aims to (A) retrospectively describe determinants of successful KLIK PROM implementation using the Consolidated Framework for Implementation Research (CFIR), and (B) identify current barriers and match implementation strategies. **Methods:** (A) The KLIK implementation process was described retrospectively based on literature and experience, using the 39 CFIR constructs organized in five general domains: intervention characteristics, outer setting, inner setting, characteristics of individuals and implementation process. (B) The CFIR-ERIC (Expert Recommendations for Implementing Change) Implementation Strategy Matching tool identified current barriers in the KLIK implementation and matched implementation strategies that

addressed the identified barriers. **Results:** (A) The most prominent determinants of successful KLIK PROM implementation lie in the following CFIR domains: intervention characteristics (e.g., easy to use), characteristics of individuals (e.g., motivation) and process of implementation (e.g., support). (B) 13 CFIR constructs were identified as current barriers for implementing the KLIK PROM portal. The highest overall advised ERIC strategy for the specific KLIK barriers was to identify and prepare champions. **Conclusion:** Using an implementation science framework, e.g., CFIR, is recommended for groups starting to use PROMs in clinical care as it offers a structured approach and provides insight into possible enablers and barriers.

#### **Implementation science metrics to evaluate patient-reported outcome measure (PROM) implementation initiatives in routine care settings**

**Angela Stover, PhD**, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States; Lotte Haverman, PhD, University of Amsterdam, Amsterdam, Netherlands; Hedy van Oers, PhD, University of Amsterdam, Amsterdam, Netherlands; Joanne Greenhalgh, PhD, University of Leeds, Leeds, United Kingdom; Caroline Potter, DPhil, University of Oxford, Oxford, United Kingdom; Submitted on behalf of the ISOQOL PROMs/PREMs in clinical practice implementation science work group

**Aims:** Patient-reported outcome measures (PROMs) are increasingly being implemented in routine care settings, but how to optimally evaluate PROM implementation initiatives is unclear. Implementation science (IS) offers a potential way forward, but its application has been limited for PROMs. **Methods:** Three case studies were reviewed that used IS to evaluate PROM implementation: (1) pain clinics in Canada; (2) oncology clinics in Australia; and (3) pediatric/adult clinics for chronic conditions in the Netherlands. Case studies used Proctor's IS outcomes framework for evaluation, which includes eight outcomes assessing acceptability, appropriateness, adoption, feasibility, fidelity, reach/penetration, cost, and sustainability. We then mapped constructs from Proctor's evaluation framework to common PROM implementation activities. This mapping yielded PROM-specific evaluation metrics for use in routine care settings (Fig. 1). **Results:** Independent case studies used the same IS framework to evaluate their PROM implementation initiative in different care settings, but the degree of application and operationalization were inconsistent. Case studies used a range of 3–6 evaluation variables. Acceptability and appropriateness of PROMs for a specific clinic were the most common evaluation outcomes. Case studies assessed acceptability and appropriateness at the individual level (e.g., clinic providers and staff) using interviews or questionnaire responses. In the mapping exercise, a metric for acceptability was the proportion of clinicians and patients who would recommend using PROMs to similar stakeholders. The remaining six Proctor outcomes were less commonly used in case studies, in part due to their applicability in later stages of implementation and measurement at the clinic level. Mapping exercise results are shown in Fig. 1. For example, Proctor's construct of reach/service penetration can be assessed as the proportion of a clinic's patient panel completing PROMs. In addition to the metrics in Fig. 1, case studies found it useful to measure barriers, enablers, and implementation support strategies to provide context for evaluation. **Conclusion:** The use of Proctor's implementation science evaluation framework across case studies indicates its promise to standardize PROM evaluation in routine care settings. Increasing the use of implementation science frameworks in evaluating PROM implementation initiatives will help advance our collective understanding of how, why, and in what circumstances PROM implementation is successful.



#### Symposium 4: Optimizing power and precision in rare disease PRO endpoints

**Moderator:** Pronabesh Dasmahapatra, MD, MPH, Head, Heath Economics and Value Assessment Strategy, Rare Disease and Rare Blood Disorder, Cambridge, Massachusetts, United States

Within rare disease, adequately powering trials for meaningful change and treatment efficacy detection remains a challenge. Low population prevalence and heterogenous symptom presentation are two of the greatest contributors to insufficient endpoint power. While no statistical procedure can increase prevalence or decrease heterogeneity, alternative methods can be used to increase the robustness of endpoints and analyses. In this symposium, we highlight three limitations that occur in rare disease studies and three corresponding solutions. We illustrate how current practices may attenuate power in patient-reported outcome (PRO) endpoints and demonstrate alternative procedures that lead to improved estimates.

1. Identifying PRO endpoints: Identifying appropriate PRO endpoints is complicated by the heterogeneity of symptoms. This, in part, has resulted in endpoints such as Most Bothersome Symptom (MBS) gaining in popularity. However, MBS is riddled with statistical and theoretical issues that result in the attenuation of statistical power and threats to validity. These issues and alternative approaches will be presented.
2. Hypothesis testing: Issues such as skew, outliers, and unequal sample sizes in treatment arms can undermine power and precision if classic hypothesis testing is used. Furthermore, adjustments for multiplicity can overcorrect for Type I error, attenuating power to detect differences across treatment arms. Alternative procedures robust to these effects in small samples are presented and evaluated.
3. Estimating meaningful change: skew, outliers, and unequal sample sizes in treatment arms also affect estimates of meaningful change. Current practice is to use descriptive statistics and visualizations (eCDFs and ePDFs) to evaluate meaningful change sans hypothesis testing. An inferential framework for testing the separation in eCDFs is presented and extended to inferential procedures robust to small population liabilities.

The concepts are illustrated using a series of accessible examples. These examples rely upon ‘real’ and simulated PRO data. Shortcomings in current practice are highlighted and compared to more optimal approaches. Emphasis will be made throughout on how the more optimal methods fit into current regulatory guidance and positions. The goal of this symposium is to engage rare disease

stakeholders in a discussion of alternative methods useful in the construction of maximally robust endpoints and analysis strategies.

#### Individual presentations

##### Most bothersome symptom in rare disease: a look at endpoint precision and validity

**R. J. Wirth, PhD**, Vector Psychometric Group, LLC, Chapel Hill, North Carolina, United States; **James McGinley, PhD**, Vector Psychometric Group, LLC, Chapel Hill, North Carolina, United States

**Aims:** There is increased interest in ‘personalized’ endpoints for the study of rare disease. In this developing area of clinical research, change in Most Bothersome Symptom (MBS) has emerged as a leading solution for patient-centered endpoints. However, MBS relies on statistical and theoretical assumptions that, when not met, can lead to reduced statistical power and increased threats to validity. The current presentation has three aims. The first aim is to examine the measurement and statistical assumptions underlying the MBS approach. The second aim is to demonstrate how statistical power and validity are impacted by violations of these underlying assumptions. The third aim is to provide researchers with alternatives to the standard MBS approach. **Methods:** MBS assumptions are examined and the impact of violating these assumptions is demonstrated using theoretical examples, simulated data, and clinical trial data. **Examples and scenarios** are used to frame the MBS approach within psychometric theory. Clinical trial and simulated data are used to highlight the degree to which assumptions may be violated and how these violations impact the inferences drawn from statistical models. **Results:** The three-pronged approach (theoretical, clinical data, and simulated data) show that relying on an MBS approach can impact one’s ability to detect ‘true’ changes in a disease state over time. Furthermore, findings suggest that, even when change is detected, there may be little validity to basing high-stakes decisions (e.g., labeling) on MBS endpoints. **Conclusion:** The use of MBS has increased in recent years. While this method appears to have positive qualities, MBS has limited theoretical, psychometric, and statistical support. These shortcomings directly impact tests of treatment efficacy. Fortunately, alternative approaches (e.g., returning to domain scores) are readily available, statistically and psychometrically defensible, and can be implemented in practice.

##### Hypothesis tests to evaluate treatment efficacy that optimize power and precision in rare disease settings

**Charles Iaconangelo, PhD**, Pharmerit International, Brooklyn, New York, United States

**Aims:** The power of PRO endpoints in rare disease studies is often reduced by small sample sizes and subject heterogeneity. This creates challenges for statistical analysis as well as data collection and identifying appropriate endpoints. Several characteristics of the data frequently encountered in rare disease studies attenuate power: Skew Outliers Unequal sample sizes in treatment arms Adjusting for multiplicity. The attenuation in power can be substantial. This presentation will demonstrate how an alternative hypothesis testing procedure that is well established in the statistical literature—permutation tests—can be used to improve power while maintaining Type I error rates. This is illustrated via simulated data based on observed examples. **Methods:** Classic hypothesis testing entails computing a  $p$  value based on a theoretical null distribution (e.g., the  $t$ -distribution). The permutation test, in contrast, empirically estimates the null distribution via resampling. This approach is robust to skew,

outliers, unequal sample sizes, and adjusting for multiplicity. Replacing classic hypothesis tests with permutation tests can optimize power and precision. A series of simulation studies compared the power and Type I error of classic hypothesis tests and permutation tests. The impact of skew, outliers, unequal sample sizes, and multiplicity adjustments were all evaluated. Results: The simulation study evidence demonstrated that conditions commonly encountered in rare disease trials all led to problematic reductions in power when using classic hypothesis tests. Furthermore, the results show that permutation tests were robust to these issues and led to substantial improvements in power while controlling Type I error. Under one condition, both the permutation test and the adjusted t-test controlled Type I error at 0.05, however the permutation test demonstrated a power of 0.76 to identify a true treatment effect, whereas the multiplicity-adjusted t-test had a power of 0.59. Conclusion: Permutation tests were well established in the statistical literature at a time when lack of computing power prevented their widespread adoption. Advances in computing power and software availability now make permutation tests an attractive option, particularly in rare disease studies where it is difficult, if not impossible, to increase power by enrolling more subjects.

### **A new method for testing significance of eCDF separation for meaningful change in rare diseases**

**Daniel Serrano, PhD**, Pharmerit International, Bethesda, Maryland, United States

**Aims:** Meaningful change estimation remains reliant on descriptive procedures. Such procedures are known to be excessively subject to bias from missing data, skew in data, and outlying influential observations. This symposium is an extension of our line of research designed to improve the estimation of meaningful change for COA endpoints. At the 2019 ISOQOL symposium we presented a procedure for unbiasing empirical cumulative distribution function (eCDF) estimation in the presence of data missing at random (MAR). This symposium presents an easily implemented inferential framework for eCDF-based meaningful change estimation. This framework is robust to threats common in rare disease studies: missing data, skew, and outlying influential observations. The procedure relies on existing software. The aim of this symposium is to disseminate this procedure so that researchers may incorporate the technique into their ongoing rare disease research. **Methods:** eCDFs can be modeled via event/trial binomial models estimated via maximum likelihood. This framework enables the testing of separation in eCDFs between anchor or treatment groups. The corresponding odds ratio quantifies the difference in cumulative proportions achieving the estimated meaningful change criterion (e.g., a 1-point improvement). Simulation results demonstrate the unbiased and efficient properties of this framework in general. The framework is extended to the rare disease space by employing the permutation test to estimate the empirical null distribution in the presence of missing data, skew, and outlying influential observations. Superiority of this procedure will be demonstrated by comparing performance to standard inference under the same conditions. **Results:** The bias to detect the generating odds ratio of 6.75 was 0.4% across 1000 replications. Bias for corresponding group proportions achieving a 1-point improvement was less than 1%. Simulations extending this work to the rare disease space will illustrate how alternative inferential frameworks are robust to small samples, missing data, skew, and outlying influential observations. **Conclusion:** Meaningful change can be embedded within a modern maximum likelihood-based estimation framework. In addition, the statistical significance of estimated meaningful change can be tested within robust inferential frameworks. Coupling these developments circumvents the main limitations hampering precise meaningful change estimation within rare disease studies.

### **Symposium 5: Integrating adverse event data and patient-reported outcomes to better understand cancer treatment tolerability: the US National Cancer Institute-funded tolerability consortium**

**Moderator:** Gita Thanarajasingam, MD, Division of Hematology, Mayo Clinic, Rochester, Minnesota, United States; Submitted on behalf of The NCI U01 Tolerability Consortium is funded by grants from the US National Institutes of Health: U01CA233046, U01CA232859, U01CA233167, U01CA233169.

**Discussants:** Beverly Canin, Cancer and Aging Research Group, Rochester, New York, United States; Lori Minasian, MD, Department of Health & Human Services, Bethesda, Maryland, United States.

Chronically administered novel treatments including molecularly targeted agents and immune therapies are increasingly incorporated in the treatment of a broad spectrum of cancers. The toxicity profile of these agents is different than that of shorter duration conventional cytotoxic chemotherapies. In addition, toxicity profiles among patients with increased vulnerability (i.e., older adults, patients with comorbidities and/or functional limitations) have been relatively unexplored even for cytotoxic chemotherapy. Tables of high grade adverse events (AEs) according to the National Cancer Institute's (NCI's) Common Terminology Criteria for Adverse Events (CTCAE) report the most severe grades and are important for safety assessment. However, this standard approach to toxicity evaluation does not adequately represent treatment tolerability from the patient perspective. Tolerability is a patient-centered, multidimensional construct that is distinct from safety, and not intended to replace it. By definition, understanding tolerability requires the patient's perspective. However, the best patient-reported outcomes (PROs) to evaluate tolerability and the optimal analytic approaches of those metrics have not been defined. Evaluation of tolerability with PROs presents an important challenge in oncology clinical trials, in the regulatory approval of new drugs, and in real-world physician-patient decision-making.

This symposium will focus on exploring cancer treatment tolerability, including the optimal PROs, metrics, analytic approaches, and displays for evaluation of tolerability in cancer clinical trials and clinical decision-making. The NCI U01 Tolerability Consortium is a multi-stakeholder group whose goal is defining consensus metrics of tolerability and standardizing analytic approaches. The symposium moderator will introduce the Tolerability Consortium and a consensus definition of tolerability. The first presentation will describe the role of the Patient-Reported Outcomes version of the CTCAE (PRO-CTCAE) in tolerability assessment and proposed interpretation and analytic strategies. The second presentation will explore the impact of host factors on cancer treatment tolerability. The third presentation will discuss the relationships between aging-related conditions and treatment tolerability in older adults with advanced cancer. The last presentation will explore the FACT GP5 item as a single-item measure of cancer treatment tolerability.

### **Individual presentations**

#### **Analytic and interpretation approaches for PRO-CTCAE, CTCAE, and other clinical data to characterize drug tolerability**

**Amylou Dueck, PhD**, Mayo Clinic, Scottsdale, Arizona, United States; Gita Thanarajasingam, MD, Mayo Clinic, Rochester, Minnesota, United States; Lauren J. Rogak, MA, Memorial Sloan Kettering Cancer Center, New York, New York, United States; Gina L. Mazza, PhD, Mayo Clinic, Scottsdale, Illinois, United States; Ethan Basch, MD, MSc, Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, North Carolina, United States

**Aims:** This presentation will describe the role of the Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) in tolerability assessment and proposed analytic and interpretation strategies. In cancer clinical trials, it is essential to understand and communicate treatment side effects to enable patients and other stakeholders to balance risks and benefits. Historically, side effect information has been reported by clinicians on behalf of their patients, a process that has been shown to miss up to half of patients' symptoms. Therefore, the US National Cancer Institute-supported development of the PRO-CTCAE, a library of questions to enable patient self-reporting of side effects. It is not yet established how best to combine PRO-CTCAE information with traditional measures of treatment impact. **Methods:** The standard analytic approaches to clinician-based CTCAE and other clinical data (e.g., time on treatment, drug dose levels, and hospitalizations) will be reviewed. Analytic approaches employed by early PRO-CTCAE adopters will also be presented, including a review of missing data considerations in these initial trials. Finally, proposed standardized tabular and graphical PRO-CTCAE representations will be displayed. **Results:** The earliest published analytic approach for PRO-CTCAE data mirrored the most common approach to the analysis of CTCAE data and reported the maximum individual item score post-baseline per patient. A more recent published approach accounts for a patient's pre-existing symptom level (i.e., baseline item score) by reporting the maximum of individual item scores post-baseline that are strictly worse than each patient's baseline item score. Missing data have predominantly been addressed through prospective minimization in the design and data collection phases; during analysis, approaches have relied on summary measure approaches. A composite scoring algorithm for creating a single grade per symptomatic AE will be reviewed. Published tabular and graphical approaches that accompany these analytic methods will be compared. Finally, future refinements to analytic and graphical approaches will be discussed. **Conclusion:** This work will produce standardized techniques for longitudinal PRO-CTCAE data that adjust for missingness and baseline symptomatology and can inform best practices for tabulating and reporting PRO-CTCAE data for more systematic and patient-centered evaluations of tolerability in cancer trials.

#### **Measuring the tolerability of cancer therapy: incorporating host factors as moderators of toxicity and outcomes**

**Patricia Ganz, MD**, Department of Health Policy and Management, UCLA Fielding School of Public Health, Los Angeles, California, United States; **Sungjin Kim, MS**, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California, United States; **Marcio Diniz, PhD**, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California, United States; **Michael Luu, PhD**, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California, United States; **Gillian Gresham, PhD**, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California, United States; **Norah Lynn Henry, PhD**, Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, Michigan, United States; **Ron Hays, PhD**, Division of General Internal Medicine & Health Services Research, David Geffen School of Medicine at University of California, Los Angeles, Los Angeles, California, United States; **Greg Yothers, PhD**, Graduate School of Public Health, Department of Biostatistics, University of Pittsburgh, Pittsburgh, Pennsylvania, United States; **Reena Cecchini, PhD**, Graduate School of Public Health, Department of Biostatistics, University of Pittsburgh, Pittsburgh, Pennsylvania, United States; **Steven Piantadosi, MD, PhD**, Surgery at Brigham and Women's Hospital, Harvard University,

Boston, Massachusetts, United States; **Mourad Tighiouart, PhD**, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California, United States; **Andre Rogatko, PhD**, Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, California, United States

**Aims:** This presentation will examine the need to consider host factors (demographics, functional status, BMI, comorbidity, health behaviors) as important covariates in assessing the toxicity and outcomes of cancer therapy. Although randomization in clinical trials may eliminate bias in assignment of treatment, there is no guarantee that important host factors will be evenly distributed across treatment arms. **Methods:** Using existing data from several completed phase III clinical trials that include standard toxicity data reporting (CTCAE), along with patient-reported outcomes (PROs), our research group has been examining the additional value of including available host factors and PRO data collected at baseline prior to treatment in models that predict subsequent toxicity. Toxicity Index (TI), a summary measure to better discriminate patients based on their overall toxicity experiences, accounting for all observed toxicity grades rather than just the most severe one, was also employed. Probabilistic index model (PIM), logistic regression model and Cox proportional hazards regression model were employed to examine associations between host factors and toxicity. Model diagnostics were performed to ensure that the models are appropriate. Multivariable analyses are performed with variables selected a priori or using a stepwise variable selection procedure based on Akaike Information Criterion (AIC). **Results:** We will present results from several trials of breast and colorectal cancer that identify differences in tolerability based on host factors to demonstrate the added value to their inclusion in the interpretation of clinical trial outcomes. **Conclusion:** By combining existing analytical methods and applying them to standard data collected in a cancer clinical trial, we can provide a more accurate description of a patient's treatment experience that could lead to individualized dosing for better toxicity control.

#### **Relationships between aging-related conditions and treatment tolerability in older adults with advanced cancer**

**Eva Culakova, PhD**, University of Rochester, Rochester, New York, United States; **Luke Peppone, PhD**, University of Rochester, Rochester, New York, United States; **Marie Flannery, PhD, RN**, University of Rochester, Rochester, New York, United States; **Huiwen Xu, PhD**, University of Rochester, Rochester, New York, United States; **Erika Ramsdale, MD**, University of Rochester, Rochester, New York, United States; **Amita Patil, MS**, University of Rochester, Rochester, New York, United States; **Beverly Canin**, University of Rochester, Rochester, New York, United States; **Supriya Mohile, MD, MS**, University of Rochester, Rochester, New York, United States

**Aims:** Geriatric assessment (GA) utilizes validated tools to capture aging-related domains known to be associated with adverse outcomes in older patients with cancer. Few studies have comprehensively evaluated relationships between GA and cancer treatment tolerability. Our aims were to determine if GA impairments are related to symptomatic toxicities (measured by Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events [PRO-CTCAE] and clinician-rated CTCAE). **Methods:** Data are from the first 301 patients enrolled in the GAP 70 + trial. Eligible patients were age 70 +, had advanced solid tumors or lymphoma, had  $\geq 1$  GA domain impairment, and were initiating a new regimen with a known high prevalence of toxicity. Eight GA domains were scored as impaired or not based on established cut points of validated instruments. Symptom severity was assessed using PRO-CTCAEs (v1.0).

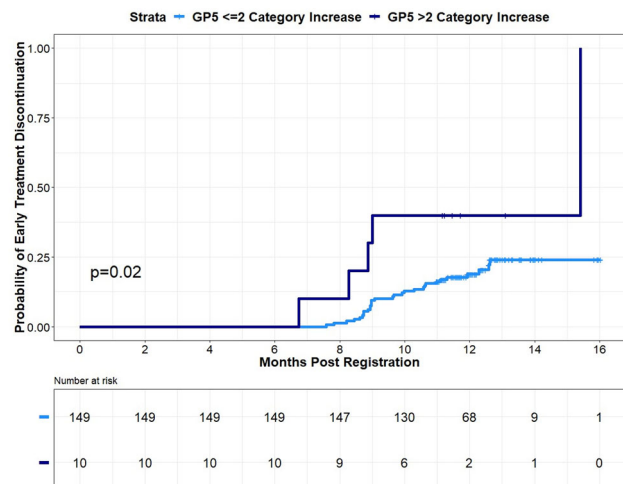
Total severity score was calculated as sum of 24 symptoms (range 0–96). CTCAE (NCI v4) was captured prospectively by staff over 3 months; medical records were reviewed centrally by blinded oncology clinicians. Association of GA domain impairments with toxicities were assessed by longitudinal linear model (PRO-CTCAE) or Cox regression (CTCAE). Results: Patients were 70–93 years (mean = 77), 45% female, 89% white; 22% had previous chemotherapy, 64% had gastrointestinal or lung cancer. An average of 4.6 GA domains was impaired, ranging from 92% in physical performance to 26% in social support. Fatigue (80%), pain (63%) and decreased appetite (57%) were the most common and severe PRO-CTCAEs. Over 3 months, total PRO-CTCAE severity score increased from 14 to 24 points. Overall, 62% patients developed any clinician-rated grade 3–5 toxicity; 92% had grade 2–5. GA impairments significantly associated with increased symptomatic toxicities included psychological status, function, nutrition, and comorbidities (mean difference = 7.55, 3.81, 3.21 and 2.97, respectively; all  $p < 0.05$ ). GA impairments associated with grade 2–5 CTCAE toxicity included psychologic status ( $p = 0.01$ ) and comorbidities ( $p = 0.05$ ). Impaired cognition was associated with increased grade 3–5 CTCAE (HR 1.35;  $p = 0.05$ ). Conclusion: Future work in evaluation of treatment tolerability should include baseline evaluation of aging-related conditions, since they are associated with PRO-CTCAE and CTCAE.

#### Analysis of ECOG-ACRIN clinical trials to advance longitudinal assessment of cancer treatment tolerability

**John Devin Peipert, PhD**, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Fengmin Zhao, PhD, MS, MHS, ECOG-ACRIN Biostatistics Center, Boston, Massachusetts, United States; Ju-Whei Lee, PhD, ECOG-ACRIN Biostatistics Center, Boston, Massachusetts, United States; Fangxin Hong, PhD, ECOG-ACRIN Biostatistics Center, Boston, Massachusetts, United States; Edward Ip, PhD, Wake Forest School of Medicine, Winston-Salem, North Carolina, United States; Ilana Gareen, PhD, MPH, Brown University School of Public Health, Providence, Rhode Island, United States; Ruth Carlos, MD, MS, FACR, University of Michigan, Ann Arbor, Michigan, United States; Ingrid Mayer, MD, MSCI, Vanderbilt University, Nashville, Tennessee, United States; Kathy Miller, MD, Indiana University Simon Cancer Center, Indianapolis, Indiana, United States; Ann Partridge, MD, MPH, Dana-Farber/Partners CancerCare, Boston, United States; Tait Shanafelt, MD, Stanford University School of Medicine, Stanford, California, United States; Keith Stewart, MB, CHB, Mayo Clinic, Phoenix, Arizona, United States; Ahmad Tarhini, MD, PhD, Winship Comprehensive Cancer Center of Emory University, Atlanta, Georgia, United States; Mary Thomas, MS, VA Palo Alto Health Care System, Palo Alto, California, United States; Matthias Weiss, MD, ThedaCare Cancer Center, Appleton, Wisconsin, United States; David Cella, PhD, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Robert Gray, PhD, Dana Farber Cancer Institute, Boston, Massachusetts, United States; Lynne I. Wagner, PhD, Wake Forest School of Medicine, Winston-Salem, North Carolina, United States

**Aims:** New methods are needed to advance patient-oriented tolerability assessment of cancer treatments. The US National Cancer Institute-supported EVOLV project brings together a multi-disciplinary team to evaluate longitudinal methods for interpreting, reporting, and visualizing patient-centered tolerability assessments, and determine these assessments' predictive value for early treatment discontinuation (ETD). Drawing from ECOG-ACRIN trials, the primary assessments include the FACT-G item GP5 ("I am bothered by side effects of treatment") and the PRO-CTCAE. **Methods:** To date, study activity has focused on GP5 as a predictor of ETD. First, we examined whether

GP5 prior to treatment (baseline) was associated with ETD using 5 phase III ECOG-ACRIN clinical trials with chronic leukemia, multiple myeloma, melanoma, and breast cancer patients; hazard ratios (HR) from Cox proportional hazard models were used to estimate this relationship. For 4 of the 5 trials, induction and maintenance phases were analyzed separately; the 5th trial was adjuvant. Next, in a phase III trial with multiple myeloma patients, we tested whether increase in GP5-rated side effect bother from baseline to cycle 7 was associated with ETD by: (1) stratifying Kaplan–Meier curves by patients who meaningfully increased in GP5 (increase of  $> 2$  response categories); (2) fitting a joint model of the longitudinal GP5 change trajectory on ETD. Results: GP5 prior to treatment was significantly associated with ETD in 4 separate analyses (2 within the same trial) across 3 of the 5 trials: 3 during maintenance and one during induction phase (HR range 1.5–5.3). In the myeloma trial, patients reporting increased bother on GP5 had higher hazard of ETD (HR 3.08 (95% CI 1.18–8.02)). (Figure) In the joint model, the estimated effect of GP5 on hazard of ETD was large: HR 9.56 (95% CI 2.41–37.82). Conclusion: To date, EVOLV has found evidence that side effect bother on the GP5 prior to treatment, and increase in GP5 while on treatment, are associated with higher likelihood of ETD. Additional analyses are ongoing to explore other single-item predictors of ETD, examining patient-reported predictors of dose modification, refinement of visualization techniques, and longitudinal, latent-variable modeling of tolerability trends. These new directions will also be presented.



#### Symposium 6: From single site to scale: what does it take to implement PROs across health systems?

**Moderator:** Danielle C. Lavalley, PharmD, PhD, University of Washington, Seattle, Washington, United States. This project was supported by grant number R01HS023785 from the Agency for Healthcare Research and Quality.

The use of patient-reported outcomes (PROs) to support screening, diagnosis, and evaluation of patient outcomes is rapidly expanding beyond research into traditional care delivery settings. Driven by a combination of policy, payer, and stakeholder incentives, health systems are recognizing the critical need for PROs to augment patient-centered care, care quality, and population health initiatives. Additionally, a growing number of health systems are prioritizing the electronic capture and presentation of PROs via health information technology to maximize efficiency and advance the patient-centeredness of medical records. While the use of PROs in clinical practice holds great promise for improving care experience and quality, PROs

bring unique considerations around measurement and reporting that health systems may not be poised to navigate. Without a thoughtful, coordinated strategy for PROs use, health systems run the risk of adding unnecessary burden for both patients and care teams, expending unsustainable resource loads, and inappropriately applying PROs data to direct patient care. In order to ensure PROs realize their potential to inform and improve care, the use of PROs must adapt to the needs and constraints of large healthcare organizations to remain sustainable, and facilitate seamless PRO data collection and integration across the care continuum. In this symposium, speakers from multiple health systems will share experiences, learnings, and best practice recommendations for scaling PRO use across large health system. Two speakers from the University of Washington will present design guidelines for PRO governance and integration across diverse health system environments, based on a 5-year AHRQ-funded project. Following, speakers from the University of Pittsburgh Medical Center and University of Utah will present complementary experiences with PRO governance and implementation in practice.

### Individual presentations

#### Recommendations for governing PRO use across health systems

**Danielle C. Lavalley, PharmD, PhD**, University of Washington, Seattle, Washington, United States

**Aims:** At the point of care, PROs provide an invaluable opportunity to understand and track patient outcomes and to inform clinical decision-making. However, at the enterprise level, the collection of PROs invites the potential for duplication, inefficiencies, fragmentation, and inappropriate data use. In this presentation, we will share recommendations for health system governance of PROs. **Methods:** This presentation draws on learnings from a five-year AHRQ-funded project to develop design guidelines for health system use of PROs in practice. The team utilized action research methodology, involving iterative cycles of planning (inquiry and identification of evidence gaps), acting (gathering data from real-world practice), observing (health system activities and initiatives related to PROs use), and reflecting (analyzing data and communicating with multiple stakeholder groups). Core activities that informed the development of governance recommendations include participation and leadership in systemwide PROs Governance committees, cataloging PROs use cases via survey and semi-structured interviews, and stakeholder analysis. **Results:** Recommended functions of PROs governance include developing infrastructure to manage the intake and prioritization of PRO measure domains, identifying repeatable and scalable models for the build and modification of electronic PROs tools, and applying PRO data within the learning health system context. Continuous stakeholder engagement is needed to sustain PROs governance efforts and align with the broader health system environment. **Conclusion:** Balancing the breadth of PRO data requests with the burden to build, maintain, and utilize PRO tools in practice requires clear and consistent governance at the systems level. As each health system's approach to PROs governance will be informed by their institutional culture and values, it is important to build communities of practice that can guide health systems in navigating PROs governance effectively.

#### Governing PROs across health systems: case study at University of Utah

**Rachel Hess, MD, MS**, University of Utah, Salt Lake City, Utah, United States

**Aims:** As the use of PROs across large healthcare delivery organizations continues to rise, it is critical to identify models of PROs governance that can provide learnings for the broader community of practice. In this presentation, the speaker will describe experiences, barriers, and successes with PROs governance at the University of Utah health system, and share practical examples of governance functions in practice. **Methods:** The University of Utah launched the use of PROs in clinical practice within a single clinical site in 2013, and moved to an enterprise-led implementation in 2015, sponsored by the Senior Vice President and led by its physician practice group. In 2019, the Utah PRO program shifted to the Chief Medical Quality Officer's office, further centralizing it in the accountability infrastructure. Utah's approach to measurement selection involved a combination of universally used (i.e., all clinics) PROMIS measures and clinic-specific selections based on clinical utility. Core features of Utah's PROs data capture and reporting tools were established and maintained by centralized resources and governance teams; however, individual practice sites tailored key attributes of PRO measure deployment to their setting. **Results:** Utah expanded its PROs implementation to over 75 clinical practices, capturing PRO measures on over 200,000 unique patients in the Utah health system. Governing the measure selection strategy supported the system's ability to efficiently scale PROs use across clinical settings, considering needs for technical development, training, and implementation monitoring via real time metrics. System-wide evaluation measures demonstrate broad scale adoption by patients and clinical teams, and highlight continued need for ongoing facilitation and sharing of continuous learning across the organization. **Conclusion:** Utah's governance experiences highlight the importance of clearly defining and leveraging a measurement strategy to support scalability and use across diverse practice settings, as well as appropriately positioning the PRO program in the organization. Future work will continue to align PRO technology with existing health system tools.

#### Recommendations for integrating PROs into clinical care delivery

**Elizabeth Austin**, University of Washington, Seattle, Washington, United States

**Aims:** The capture of PROs can introduce nuanced challenges (i.e., adaptive logic, complex scoring and interpretation) that add complexity to clinical workflow. Successful integration of PROs into clinical practice must address and adapt to the needs of diverse clinical stakeholders (e.g., patients, providers, administrators) and environments (e.g., primary care, specialty care, remote monitoring). Bringing PROs to scale across health systems can also highlight considerations around resource allocation, security, clinical care policies and procedures, and approaches to patient engagement in the use of technology such as patient portals. In this presentation, we will share recommendations for health system integration of PROs into clinical care delivery. **Methods:** This presentation draws on learnings from a 5 year AHRQ-funded project to develop design guidelines for health system use of PROs in practice. The team utilized action research methodology, involving iterative cycles of planning (inquiry and identification of evidence gaps), acting (gathering data from real-world practice), observing (health system activities and initiatives related to PROs use), and reflecting (analyzing data and communicating with multiple stakeholder groups). Core activities that informed the development of integration recommendations include formative and summative evaluation of multiple PROs implementations across multiple practice sites that included review of implementation monitoring metrics, qualitative interviews, observation and fieldwork of practice sites, and documentation of ongoing practice facilitation efforts. **Results:** Recommendations for the integration of PROs into clinical practice include clearly defining goals

for how PROs will inform care, aligning workflows for PRO capture with existing clinical environments, identifying opportunities to support users with technology and training, and engaging in active monitoring and evaluation. In particular, learnings highlight core workflow functions that can guide repeatable models of PRO implementation across diverse clinical settings. Conclusion: Health systems will need to give thoughtful attention to the needs of PRO workflows across clinical settings in order to identify opportunities to support standardization, infrastructure for training and ongoing monitoring, and ensure continued alignment between PROs data capture and goals for clinical care.

### **Integrating PROs into care delivery: case study at University of Pittsburgh Medical Center**

**Janel Hanmer, MD, PhD**, University of Pittsburgh, Pittsburgh, Pennsylvania, United States

**Aims:** The integration of PROs into clinical care requires planning across clinical, IT, operational, legal, and reporting services to ensure PRO data collection and reporting processes align with diverse contexts of clinical care. This speaker will share experiences, best practices, and recommendations for PRO integration across the University of Pittsburgh Medical Center (UPMC), and share practical examples of PRO workflow design, training, and strategies to integrate PROs into clinic culture and practice. **Methods:** UPMC began collecting PROs in Epic in 2012 and the UPMC health system launched a Patient-Reported Outcomes (PRO) Center in 2017, with the goal of improving how PRO data informed individual patient care, clinical services, and population health. Through this, the UPMC PRO Center supports clinical teams throughout the stages of planning (e.g., PRO measure selection, goals for data use at point of care and system level), design (e.g., workflow model, data capture and reporting tools via the EHR and patient portal), deployment (e.g., training), and evaluation (e.g., process and outcome metrics) of PRO tools used in practice. **Results:** The UPMC PRO Center has supported the development and management of ePRO workflows at over 273 clinical sites in the UPMC health system. UPMC's implementation model is anchored in standardized project planning and implementation tools, the use of clinical and operational champions, and routine implementation monitoring metrics. Two use cases, PRO as a process measure (depression screening in primary care) and PRO as an outcome measure (quality reporting in Physical Medicine and Rehabilitation) will be used to describe UPMC's model and distinguish key attributes of PRO workflow design and implementation support across settings. **Conclusion:** The UPMC PRO Center's experience has provided a wealth of real-world learnings around effective workflow design and approaches to supporting the scale of PRO implementation across a large health system. In particular, UPMC's experience demonstrates the critical role of defining how PRO scores will inform decision-making, and ensuring all roles are engaged in training to support the application of PRO data to clinical care.

### **Symposium 7: Application and value of qualitative in-trial or 'exit' interviews to generate supplementary evidence regarding patients' trial, treatment and disease experiences: examples from multiple disease areas**

**Moderator:** Sophi Tatlock, Adelphi Values Ltd, Bollington, United Kingdom

In clinical trials, the value of collecting additional patient experience data from trial participants beyond that provided by clinical outcome assessment (COA) endpoints is increasingly recognized. Qualitative

'embedded' or 'exit' interviews are increasingly conducted as an additional means of capturing patient experiences. Collecting qualitative data from trial participants provide an opportunity to obtain in-depth feedback regarding their experience of disease symptoms, their evaluation of treatment (both positive and negative), and perspectives regarding clinical trial participation. Such findings can be used to inform the design, refinement and/or interpretation of COAs in future trials and administration or application of the treatment in the real-world, post-approval. Regardless of clinical trial design or phase, qualitative exit interviews can be valuable to gain in-depth insights from patients that are not possible using traditional COAs.

This symposium will provide an overview of the application of qualitative exit interviews using examples in a variety of contexts for a multitude of different objectives. The methods employed and resulting findings will be used to frame discussions of the value of such data to a range of stakeholders including patients, clinicians, sponsors and regulators. The first presentation outlines the use of qualitative exit interviews with trial participants to explore participants' disease and treatment experiences. Implications of interview findings for informing future trial design, measurement strategy and for inclusion in regulatory submissions will be discussed. The second presentation aims to provide an overview of the application of qualitative exit interviews with trial investigators to inform patient education support programs and potential future use of treatment in the real-world. The third presentation outlines the value of exit interviews in a rare disease to generate evidence on treatment preferences, to inform patient-clinician discussions and promote tailored treatment decisions in the context of a novel, newly approved therapy. The fourth presentation will describe the results of qualitative exit interviews with study partners of clinical trial participants to explore observed individual experiences of treatment and meaningful changes in disease experience, with the aim of supporting treatment messaging and informing meaningful change thresholds on COAs.

### **Individual presentations**

#### **Using qualitative exit interviews to explore the disease experience and evaluate Performance-Reported Outcomes (PerFOs) with patients who experience muscle weakness**

Nicola Williamson, Adelphi Values Ltd, Bollington, United Kingdom; David Neil, GlaxoSmithKline, Upper Providence, Pennsylvania, United States; Aoife Mahon, Adelphi Values Ltd, Bollington, United Kingdom; Sophi Tatlock, Adelphi Values Ltd, Bollington, United Kingdom; Adam Gater, Adelphi Values Ltd, Bollington, United Kingdom; **Maggie Tabberer, MSc**, GlaxoSmithKline, London, United Kingdom

**Aims:** Loss of muscle mass and strength are serious consequences of many diseases. In particular, muscle weakness is associated with significant deficits in physical function and mobility, negatively impacting quality of life (QoL), limiting ability to perform daily activities and contributing to significant morbidity and early mortality. However, at present, there is a lack of approved medications for the treatment of muscle wasting disorders of any etiology and limited research concerning the patient experience of such disorders. This creates challenges for the design of clinical trials investigating novel therapeutic interventions, especially the selection and evaluation of appropriate clinical outcome assessments (COAs) to capture treatment benefit. **Methods:** This session will present findings, implications and learnings from qualitative exit interviews with patients experiencing muscle weakness associated with Chronic Obstructive Pulmonary Disease (COPD;  $n = 67$ ) in a Phase IIa randomized, double-blind, placebo-controlled study (GSK200182). During the interviews, participants were asked to reflect on their



disease experience (including symptoms and functional impairments) at the commencement of the trial and ways in which this may have changed throughout the course of the clinical study. Feedback regarding participant perspectives towards the treatment (including features, benefits, side effects and overall satisfaction) and their experience of participating in the clinical trial (including feedback on trial procedures) was also sought. Results: Interview findings provided vital information regarding the patient disease experience, generating supplementary insights to that collected via qualitative research studies conducted independent of a clinical study. The utility of data for evaluating the content validity of a diverse range of COAs (including patient-reported outcomes, clinician-reported outcomes and performance outcomes) within this specific context of use and for defining meaningful change will be discussed. During the interviews, patients described qualitative improvements in their symptoms and functional limitations beyond those captured by COAs. Conclusion: The presenters will reflect on the value of such data for characterizing treatment benefit and generating preliminary insights regarding individualized benefit risk; specifically, for communication and decision-making among different stakeholders as the emphasis of clinical development moves from pharmacokinetics and pharmacodynamics to later stage clinical studies where the focus is on comparative evidence regarding safety, efficacy and effectiveness. Funding: GSK (study NCT03359473/200182)

#### **Qualitative exit interviews with study clinical investigators to explore trial and treatment experience and feasibility of use in general clinical practice**

**Nicola Williamson**, Adelphi Values Ltd, Bollington, United Kingdom; **Jo Atkinson**, Pfizer Ltd, Surrey, United Kingdom; **Nicola Bonner**, Adelphi Values Ltd, Bollington, United Kingdom; **Rob Arbuckle**, Adelphi Values Ltd, Bollington, United Kingdom; **Katie Tinsley**, Adelphi Values Ltd, Bollington, United Kingdom; **Carl Cooper**, Adelphi Values Ltd, Bollington, United Kingdom; **Lucy Abraham**, Pfizer Ltd, Surrey, United Kingdom

**Aims:** Qualitative interviews conducted during or upon completion of clinical trials are increasingly being conducted with trial participants (patients) to obtain feedback on their disease, treatment or trial experience. However, valuable insights can also be obtained from interviewing trial investigators on study procedures, treatment experience and feasibility of treatment administration. Evidence generated can be used to inform patient support programs and understand barriers to use in the real world. The objective of this study was to conduct qualitative exit interviews with clinical investigators involved in global, Phase III, randomized, double-blind, active-controlled clinical trials in chronic pain, to obtain feedback on their experience of the trial and administering the treatment. Clinical investigators across sites in the US, UK, Spain, and Japan ( $n = 31$ ) participated in a 30-min telephone interview once all randomized patients had completed the efficacy phase of the trial, while still blinded. Questioning explored investigator perspectives on trial study procedures and training materials, potential logistical challenges related to treatment administration in clinical practice and perspectives on suitability of the treatment for different patient subgroups. This session will outline the value of qualitative exit interviews with clinical investigators relating to four key topics. Trial study procedures—Discussions of study procedures highlighted the value of patient support materials regarding study procedures to facilitate patient-clinician discussions and improve patient education. Mode and frequency of administration—Insights were gained regarding

investigator views and perceptions of patient treatment preferences in consideration of factors such as patient convenience and site burden, highlighting benefits and pitfalls of different options and aiding appropriate selection of methods for incorporation into general clinical practice. Feasibility of treatment administration—The treatment was considered feasible for administration in general clinical practice but potential barriers regarding logistical and practical considerations were highlighted, some differing by country, which can inform identification of appropriate settings for treatment administration in the real world. Use of treatment in general clinical practice—Insights into what patient subgroups are likely to be prioritized for treatment in general clinical practice, aiding identification of key product attributes and benefits. Additional considerations based on patient perceptions of the treatment will also be discussed.

#### **Patient perspectives on the benefit of a novel therapy for a rare disease: Using qualitative exit interviews to inform post-launch value messaging and to support shared treatment decision-making**

**Jane Wells**, Adelphi Values Ltd, Bollington, United Kingdom; **Parth Vashi, PharmD**, Bayer Pharmaceuticals, Whippany, New Jersey, United States; **Adam Gater**, Adelphi Values Ltd, Bollington, United Kingdom

**Aims:** Patients are increasingly active participants in the management of their health, working together with healthcare professionals to make treatment choices based on clinical evidence and their own preferences. This can present challenges for novel therapies entering the market, as data regarding patient experience of therapies (particularly in a real-world setting) may be limited. This is especially the case in rare diseases where, even if appropriate patient-reported outcome measures are available, small sample sizes in clinical studies may limit sensitivity and opportunities to capture meaningful changes. Trial exit interviews provide a means to elicit supplementary patient experience data. This session will provide insights from an exit interview study among patients participating in an extension study for a novel Hemophilia A therapy. Hemophilia A is a rare, hereditary disorder characterized by repeated and prolonged bleeds into muscles and joints resulting in pain, limitations to physical functioning, and impacts on health-related quality of life. Prophylactic treatment for Hemophilia A (requiring intravenous injections 3–4 times per week) is burdensome to patients and adherence is suboptimal. Extended half-life (EHL) factor VIII replacement therapies offer longer intervals between infusions while maintaining efficacy and safety outcomes. To explore the importance of infusion frequency and the potential benefits of reduced infusion frequency among patients receiving prophylactic treatment with an EHL product, exit interviews were conducted with patients ( $n = 16$ ) who participated in the extension phase of a Phase II/III partially randomized, open-label trial. Qualitative feedback highlighted that longer duration of factor coverage and less frequent administration (compared with conventional FVIII replacement therapies) was associated with numerous benefits, including greater ability to participate in physical activities; better vein health; less time scheduling and administering FVIII; reduced impact on work; and improved emotional well-being. In this session, challenges and solutions to conducting multinational exit interviews independent of clinical trial protocols will be discussed. In addition, the value of such evidence related to post-launch activities for Jivi® (approved in US, EU and Japan in 2018), in terms of communicating patient experiences and perspectives to a broad array of internal and external stakeholders using a variety of communication channels, will also be discussed.

### Use of qualitative exit interviews to explore individual experiences of treatment and meaningful change in two clinical trials for autism spectrum disorder

**Elizabeth Gibbons, MSc**, Clinical Outcomes Solutions, Folkestone, United Kingdom; **Tom Willgoss, PhD**, F. Hoffmann-La Roche, Welwyn Garden City, United Kingdom; **Susanne Clinch, PhD**, F. Hoffmann-La Roche Ltd, Welwyn Garden City, United Kingdom; **Michael Cladek, PhD**, Clinical Outcomes Solutions, Chicago, Illinois, United States; **Claire Burbridge, MSc**, Clinical Outcomes Solutions, Folkestone, United Kingdom

**Aims:** Exit interviews following a clinical trial provide an opportunity for obtaining a wealth of information about the individual's experience living with a condition, participating in a clinical trial, and taking a treatment. They can also be used to explore the meaning of changes on Clinical Outcome Assessments (COAs) used as endpoints in clinical trials. This approach is being adopted in two clinical trials with individuals with Autism Spectrum Disorder (ASD); one in children and another in adults. The aim of the exit interviews was to explore individual experiences of treatment and meaningful change to inform the interpretation of key clinical trial efficacy data/outcomes and contextualize the perceived benefit of treatment. **Methods:** Semi-structured interviews are being conducted with study partners of clinical trial participants within 4 weeks of completing the trial (approximately 80 study partners in each study). Study partners and interviewers are both blinded to treatment allocation. Thematic analysis of the qualitative data is being conducted to identify what changes clinical trial participants have experienced, the impact these changes have on daily life and the importance of any changes experienced. Interview data are also being used to derive anchors that will inform the estimation of meaningful change thresholds on key COAs. **Results:** The interviews are providing evidence to support treatment value messages and allow meaningful interpretation of trial data. Initial blinded results show key changes experienced and highlight the value and impact these changes are having on the daily lives of the clinical trial participants and their families. For example, study partners reported that changes in socialization led to improvements in everyday life beyond individual social interactions. Such changes increased willingness to participate in activities, improved family interactions, and increased emotional well-being of participants and their families. **Conclusion:** These initial findings demonstrate the unique value of conducting exit interviews with clinical trial participants as they provide rich descriptions of changes as well as any treatment benefit, and a more holistic understanding of the individuals' experience. Such qualitative data provide important contextual information when deriving meaningful change thresholds and how this level of change may impact daily activities and quality of life.

### Symposium 8: Making sense of sensors: selecting, incorporating, and analyzing fit-for-purpose technologies for continuous data capture in clinical trials

**Moderator:** Philip Griffiths, PhD, Adelphi Values, Bollington, UK

**Discussant:** Michelle Campbell, US Food and Drug Administration, Silver Spring, Maryland, United States

Digital technologies (e.g., wearable, in-home, and ingestible sensors) enable passive collection of patient-level data, resulting in “new” types of data: either in quality (capturing data we've never previously been able to collect) or quantity (continuous flows of data). When assessing how a patient feels or functions, especially when evaluating the potential treatment benefit of a medical product, it is important to

ensure these new technologies are not just monitoring meaningful aspects of health, but are truly “patient-centered.”

In this symposium, experts from across different stakeholder groups will draw on their applied experience implementing digital technologies to highlight key considerations and:

- Discuss approaches to efficiently facilitating collaboration of key stakeholders—sponsors, regulators, technology vendors, data scientists, statisticians, and others—to meet the needs of all involved and to ensure a successful study.
- Describe the need for a standard lexicon across stakeholders in digital health to meet the needs of the clinical trial or study.
- Review the implications of using digital technologies for endpoint assessment in clinical trials, how these can diverge from traditional studies, and what should inform the decision to use digital measures in the first place.
- Discuss potential study designs for optimizing digital sensors, practical implementation of the sensor and data collection, subsequent management of the data and pitfalls to anticipate, assessment of the measurement properties of the sensor-derived endpoint, and analysis plan construction.
- Describe core considerations for the usability of remote sensor technologies and the data they collect.

This series of brief presentations and a lively panel discussion will spur meaningful audience interaction with perspectives from: Industry sponsor(s), “digital health” specialist groups, statisticians, psychometricians, and technology providers specializing in passive data collection within clinical trials.

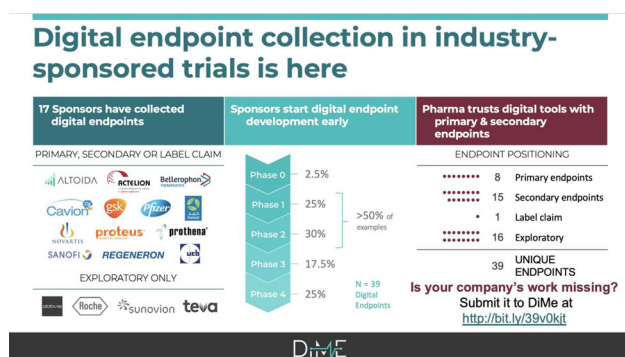
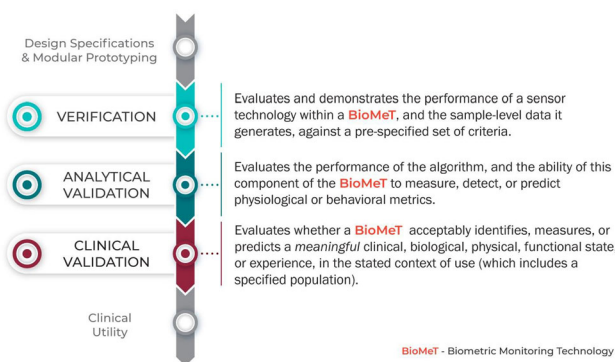
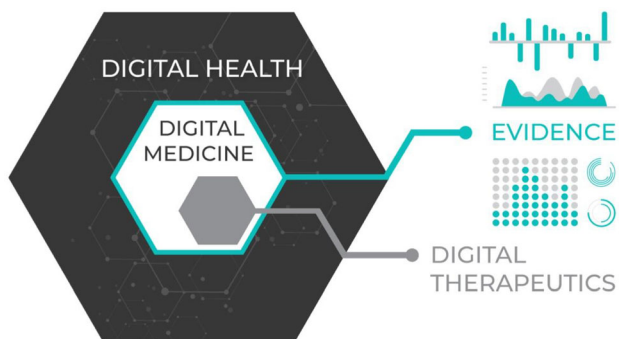
### Individual presentations

#### How can transdisciplinary collaboration help us ensure that the digital medicine tools we are being asked to place our trust in are indeed trustworthy?

**Jennifer Goldsack, MChem, MA, MBA**, Digital Medicine Society (DiMe), Boston, Massachusetts, United States; Submitted on behalf of The Digital Medicine Society (DiMe)

**Aims:** This presentation has two aims. First, to identify the transdisciplinary experts who are critical to the advancement of digital technologies to optimize health. Second, to describe how these experts can collaborate across traditional disciplinary silos, including the need for a unifying technical language. **Methods:** Founded in 2019, the Digital Medicine Society (DiMe) is the first professional organization for experts from all disciplines comprising the diverse field of digital medicine. Together, we drive scientific progress and broad acceptance of digital medicine to enhance public health. DiMe is a 501(c)(3) non-profit organization dedicated to advancing digital medicine to optimize human health. We do this by serving professionals at the intersection of the global healthcare and technology communities, supporting them in developing digital medicine through interdisciplinary collaboration, research, teaching, and the promotion of best practices. Projects completed to date include the publication of a primer on measurement in digital medicine, the development and maintenance of a library of digital endpoints being used in industry sponsored trials of new medical products, and establishing a framework for evaluating whether biometric monitoring technologies (BioMeTs) are fit for purpose. **Results:** In the 8 months since launch, DiMe has established a thriving community of over 900 individual experts from 39 different countries representing all the fields comprising digital medicine. Efforts to establish a shared lexicon continue and include defining digital health, digital medicine, and digital therapeutics; proposing a verification, analytical validation, and clinical validation framework for evaluating fit-

for-purpose BioMeTs; and engaging all disciplines—from cybersecurity experts to citizen scientists, engineers to ethicists, and regulators to researchers—in advancing the field. Conclusion: With 39 digital endpoints currently included in clinical trials of new medical products, digital is not the promise of the future of clinical trials, it is already here. Transdisciplinary collaboration is essential to ensuring that the digital tools we are being asked to place our trust in are indeed trustworthy.



#### Industry considerations for the use of technologies for continuous data capture

**Jiat Ling Poon, PhD**, Eli Lilly and Company, Indianapolis, Indiana, United States; Elizabeth (Nicki) Bush, Eli Lilly and Company, Indianapolis, Indiana, United States

Aims: The emergence and development of continuous data collection technology have allowed developers of pharmaceuticals and medical

devices, and researchers to assess clinically relevant disease-related outcomes that were previously infeasible more easily, reliably, and with potentially less burden to patients/participants. This has opened opportunities in clinical development for the inclusion of study endpoints that could previously not be efficiently assessed either outside of a clinical setting, or continuously throughout the duration of a study. Such technologies have also allowed for the assessment of existing established outcomes using new modalities that may complement or improve upon existing modalities. An overview of the role of continuous data collection technology in clinical development will be provided along with a discussion of the various considerations for endpoint selection. Methods: Drawing from experiences in clinical development programs, the discussion will focus on the considerations around appropriate endpoint selection, specifically endpoints collected through continuous data collection technology, for inclusion in clinical and/or observational research. The practicalities of implementing such data collection modalities will also be discussed. Results: The presentation will include a discussion of methods for ensuring and demonstrating that the outcomes included and assessed in clinical and/or observational studies are clinically relevant and meaningful to patients. The considerations around appropriate selection of technology best suited to assess the outcome of interest and any added value over and above traditional assessment modalities will also be discussed. Additionally, the practical implications of incorporating such technologies in studies will be covered, including the need to involve internal and external stakeholders, patient retention, and data considerations such as monitoring of data collection and data flow. Conclusion: The data that are collected from the use of cutting-edge technologies allow pharmaceutical and medical device developers to assess clinical and functional outcomes of treatment that were previously not possible, or assess existing outcomes in potentially complementary or better ways. However, endpoint selection and mode of assessment still needs to be driven by carefully considering the research question while balancing the practical aspects of study execution to produce results that are clinically meaningful and relevant to patients and other stakeholders.

#### Technologies for continuous data capture in clinical trials—the logistics of making it happen

**Paul O'Donohoe**, Medidata, London, United Kingdom

Aims: The ongoing development of increasingly accurate, increasingly powerful, increasingly user-friendly and ever-cheaper technologies has driven a growing interest in utilizing wearable and other sensor types (ingestible; in-home etc.) as a way of gaining a refined and unique insight into the patient experience in clinical trials. While there is ongoing and robust discussion around how wearables and sensors will actually support a better understanding of how patients are feeling and functioning, there has been less focus on the equally important topic of the logistics of using these technologies—namely how these devices, the data they produce, and the human interactions they require, are successfully supported within the context of a clinical trial. This presentation will give a technology providers perspective on the challenges of managing the hardware, software and support systems needed to best unlock the potential these technologies hold. Methods: Using case study examples of wearable and sensor technologies used in clinical trial settings, the hurdles, considerations and key principles for successfully managing these technologies will be reviewed and discussed. Results: Key considerations for delivery of these technologies in a clinical trial setting include:— properties of the hardware and software of the target device; integration of the target device into the broader clinical trial data ecosystem; getting devices to sites and patients; and, training and technical support of devices in the field. Conclusion: The availability of wearable and sensor technologies for continuous data capture in clinical

trials, and the number of trials actually using these technologies, is only going to increase in the coming years. This presents a huge challenge for study teams and the providers tasked with seamlessly integrating these technologies into the already extremely complex clinical trial ecosystem, while ensuring patient burden is kept to the absolute minimum. Following some key principles around the assessment of devices, integration into the broader data structure, and logistical, training and helpdesk support for patients and sites, can ensure a study is best set-up for taking advantage of these novel data sources.

### **Technologies for continuous data capture: discussing dilemmas with data and considering potential analysis approaches**

**Carrie Houts**, Vector Psychometric Group, Tempe, Arizona, United States; **James McGinley**, PhD, Vector Psychometric Group, LLC, Chapel Hill, North Carolina, United States; **Philip Griffiths**, PhD, Adelphi Values, Bollington, United Kingdom

**Aims:** Continuous data collection technology, such as ingestible sensors, heart monitors, or actigraph devices, allow researchers unprecedented access to objective patient data. While such technologies produce an enormous amount of information, optimal methods for efficiently understanding the data and testing clinically relevant hypotheses is an area of opportunity. An overview of methodological and statistical considerations related to the analysis of data from such collection modes is discussed. **Methods:** Using exemplar data from such technologies, the properties of variables obtained from these data collection methods will be reviewed. The presentation will focus on how such variables may deviate from more typical clinical trial variables, both at an individual variable level and as possible sets of variables measuring a common “concept.” Additionally, opportunities for the use of atypical analysis frameworks, made feasible by the time-intensive, but short-term longitudinal nature of such data, will be discussed. This will focus on how results from such methods can be combined with traditional outcomes to provide answers to clinically relevant questions, such as patient experience or treatment efficacy. **Results:** Properties of variables (missingness rates, distributional concerns, intercorrelation, etc.) stemming from continuous data capture technologies are discussed. The use of well-established mixed-modeling approaches is proposed as a starting point for analyzing the available data. These incorporate strong, flexible statistical methods while still providing results able to address typical clinical trials aims. Other novel methods that may be useful in understanding continuous data, like machine learning or longitudinal latent variable modeling techniques, are also briefly introduced. **Conclusion:** The information-rich digital data that are collected from cutting-edge technologies can provide researchers with the opportunity to ask and answer nuanced questions which were previously unavailable. However, initial analysis of such data has revealed potential issues with current measurement and analysis techniques often employed in clinical trials. More methodological research is needed to better understand the best way to efficiently leverage the wealth of available information digital data has to offer; analyses must be both statistically rigorous and produce clinically meaningful results that are interpretable to the numerous stakeholders in clinical research programs (e.g., patients, trial sponsors, regulators).

### **Symposium 9: Using the estimand framework to align study design and analysis with patient-reported outcome objectives: the times they are a-changin’**

**Moderators:** **Bellinda King-Kallimanis**, PhD, US FDA, Silver Spring, Maryland, United States; **Madeline Pe**, PhD, EORTC Quality of Life Department, Brussels, Belgium

Patient-reported outcomes (PROs), assessing patients’ self-reported functioning, symptoms, and general health status are critical in the evaluation of benefit/risk and relative effectiveness of new treatments. However, current PRO research objectives are frequently not clearly stated in-trial protocols and analysis plans (e.g., compare Treatment A vs Treatment B on PROs). This has the potential to lead to an unclear interpretation that may adversely impact the analysis, and robustness of PRO findings. To address this shortcoming, research objectives must be well defined to inform the analysis and interpretation of PRO results.

The 2019 International Council for Harmonization guideline for the estimand framework is a promising approach to support the development of well-defined research objectives. An estimand is defined as the target of estimation based on a scientific question of interest and composed of five attributes:

1. Treatment: interventions or combination of intervention administered concurrently
2. Target study population: which patients are the focus of the question
3. Variable of interest (endpoint): what will be measured and how
4. Intercurrent events: events that preclude observation of the variable or distort its interpretation
5. Population level summary: what is the basis for comparison

For each estimand attribute, multiple options are available. For example, the target study population could refer to all randomized patients or a specific subgroup (e.g., patients with at least a baseline PRO assessment). Making appropriate decisions for these attributes are complex and often involve different perspectives. Early multi-disciplinary, multi-stakeholder discussions are needed to ensure that decisions reflect the primary objective of assessing PROs in a specific clinical trial.

This symposium is not intended as endorsement of a study design or outcome, but rather as an illustration of the thought process that goes behind aligning the study design and analysis with research objectives using the estimand framework. The majority practice of vaguely stated PRO objectives in clinical trial protocols is hurting the community’s ability to make sense of this rich source of data. With the introduction of patient focused legislation and guidelines, this symposium will highlight how the times are a changin’ in how PRO objectives are stated in clinical trial protocols.

### **Individual presentations**

#### **An HTA perspective on the treatment policy estimand if the endpoint is repeatedly measured over time**

**Christoph Schürmann, Dr.**, IQWiG—Institute for Quality and Efficiency in Health Care, Cologne, Germany

**Aims:** Benefit assessments by IQWiG evaluate the comparative effectiveness of a new drug versus the standard of care in the approved patient population. By law, relevant endpoints include mortality, morbidity and health-related quality of life. The estimand applied is the „treatment policy“, because interest is in a treatment’s effect on a specific population eligible for treatment (target population), irrespective of intercurrent events (IE). To describe the patient experience, observing the complete period from baseline to end of study and using an overall effect measure is considered relevant. **Methods:** We consider pain as an example endpoint of a patient-reported outcome that was measured repeatedly over time in the intention to treat (ITT) population until the end of the study on some continuous scale. To assess the burden of pain over the complete study period, the effect measure of interest is the difference in means, taking into account all data from the repeated measurements. A suitable class of statistical models is linear mixed models, for which a

standard for repeated measurements is already established. Generally, IQWiG is not provided individual patient data, only aggregated outcome and IE data. A treatment policy analysis is not always provided and cannot be performed by IQWiG using the aggregated data, e.g., if due to the study design data collection stops after IEs such as progressive disease or discontinuation of treatment. Results: Employing the treatment policy estimand requires data collection for relevant outcomes to be continued after IEs. If data collected after IEs are not available, it may be possible to approximate a treatment policy estimand by statistical analyses with appropriate handling of missing data. However, this approach requires a thorough evaluation of the risk of bias with respect to patients affected by IEs. Conclusion: The treatment policy estimand is closest to the ITT principle and critical when assessing a therapeutic effect for a target population. A suitable effect measure should capture the overall patient experience, therefore, a suitable endpoint should be repeatedly observed and included in the analysis irrespective of IEs.

### Why & how to use time-to-event endpoints for COAs & how can the estimand framework help?

**Rachael Lawrance**, Adelphi Values Ltd, Bollington, United Kingdom

Aims: Describing the benefit of a new treatment regimen based on delaying time to disease progression compared to current treatment is a well-established approach in oncology clinical trials. In this breast cancer case study, the primary clinical efficacy endpoint is “disease progression.” How do we best incorporate the patient’s perspective about their treatment in a way that also aligns with the primary endpoint of the trial? In chronic disease conditions, it is generally expected that over time patients’ symptoms will worsen, functioning will decrease and general HRQoL will decline. A successful treatment should delay these declines. Therefore, the time-to-event approach is very relevant—how do we consider use of time-to-event endpoints for clinical outcome assessments & how can the new estimand framework help? Methods: We consider a naïve PRO objective, and detail how to use the five components of the estimand framework to help us construct a more specific objective. The naïve objective presented here is based on evaluating the idea that a new treatment “delays the decline in physical function.” An example estimand will be presented, as well as discussing the many potential issues when considering the use of a time-to-event endpoint for PRO data such as protocolled data collection schedules, intermittent missing data, dropout, the role of censoring, disease progression, cross-over therapy and deaths. Results: Firstly, the variable of interest must be clearly defined, and whether death should be included as a deterioration event or not must be considered. The example estimand presented focuses on the on-treatment period; patients with events would be those with decline in physical function as measured by change in baseline in PRO assessment score while on treatment; patients who died would be censored in the analysis. Other issues for consideration in construction of a time-to-event endpoint for a PRO will also be presented. Conclusion: This talk will highlight considerations needed when building a precise objective when considering a patient-reported time to event endpoint, which is a highly relevant type of endpoint in oncology clinical studies.

### Applying the estimand framework to describe patient experience while on treatment: a case study

**Mallorie Fiero, PhD**, U.S. Food and Drug Administration, Silver Spring, Maryland, United States; **Madeline Pe, PhD**, European

Organisation for Research and Treatment of Cancer, Brussels, Belgium; **Chana Weinstock, MD**, U.S. Food and Drug Administration, Silver Spring, Maryland, United States; **Laura Lee Johnson, PhD**, U.S. Food and Drug Administration, Silver Spring, Maryland, United States; **Bellinda King-Kallimanis, PhD**, U.S. Food and Drug Administration, Silver Spring, Maryland, United States; **Paul Kluetz, MD**, U.S. Food and Drug Administration, Silver Spring, Maryland, United States; **Scott Komo, PhD**, U.S. Food and Drug Administration, Silver Spring, Maryland, United States; **Rajeshwari Sridhara, PhD**, U.S. Food and Drug Administration, Silver Spring, Maryland, United States

Aims: Unclear clinical outcomes assessment (COA) research objectives are common and can lead to potentially misleading conclusions about patient experience. The estimand framework in the ICH E9(R1) addendum provides standardization of principles to improve dialogue between all disciplines involved in the development of objectives, design, conduct, analysis, and interpretation of a trial. We apply the estimand framework to a case study with a research objective assessing physical function while on treatment in an advanced cancer setting. Methods: The case study is a randomized trial of patients who have advanced breast cancer; progression-free survival is the primary endpoint. Physical function score is collected at every treatment cycle. A multi-disciplinary team formulated an estimand based on the question: “At every assessment, what is the proportion of patients on treatment who at least maintained their physical functioning for each treatment arm?” This research objective is not geared towards an efficacy claim and is descriptive. The goal is not to make direct comparisons between treatment arms. Results: The target study population includes on-treatment patients who received at least one dose of the drug and completed baseline physical function assessment. We defined the endpoint as patients who maintained or improved in their physical function based on pre-specified criteria at every assessment point until end of treatment. Intercurrent events are events that occur after randomization that may impact interpretation of patient experience. Intercurrent events of interest include death, disease progression, treatment discontinuation and initiation of subsequent therapy. If any of these intercurrent events occur, the patient will be removed from the analysis at subsequent time points because we are interested in evaluating patients on treatment. We defined the population-level summary as the proportion of on-treatment patients who maintained or improved physical function for each treatment arm at each assessment until the end of treatment. Conclusion: The estimand framework provides transparency in the questions being answered and decisions made in the analysis of COA data, which improves interpretation of patient experience in regulatory decision-making. This is not an endorsement of a study design or estimand; rather it is meant to illustrate principles in conceptualizing a COA research question and design.

### Symposium 10: The future of electronic patient-reported outcome measurement for children and adolescents: the use of computer-adapted testing

**Moderator:** **John Chaplin, PhD**, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden

This symposium will highlight the emerging research in the field of pediatric Computer-Adapted Testing (CAT). It will explore how CATs can be used to increase precision of measurement while simultaneously reducing the measurement burden. The symposium will describe the full trajectory of CAT development and use from the collection of reference data in a general population, through the implementation of a CAT in different settings and different patient groups and finally to the presentation and interpretation of the results. The symposium will examine the benefits and challenges of CAT

PROM measurement. Two CAT tools, available to assess pediatric populations, will be discussed: the pediatric PROMIS® and the Kids-CAT®.

Dr. Christiane Otto will describe the development, validation, feasibility, acceptance, and application of the Kids-CAT for self-reported health-related quality of life (HRQoL) in chronically ill and healthy children and adolescents. Doctoral student Michiel Luijten will then look at the challenges of translation and implementation of a CAT system using the example of PROMIS implemented in an App which systematically monitors quality of life of chronically ill children and their parents (KLIK). Dr. Kaveh Ardalan will assess measurement qualities of CATs compared to fixed short-form PROM measures in terms of precision and burden across the range of symptom experience. Professor Jin-Shei Lai will examine the measurement properties of a CAT in an example of a difficult-to-access child population with brain tumors. Finally, Dr. Kathrin Fischer will look at the application of the Kids-CAT and its measurement precision in longitudinal assessment over 6 months of HRQoL in children and adolescents with type 1 diabetes mellitus; she will report on the association between the blood glucose level (HbA1c) and Kids-CAT domains.

At the end of the symposium, participants will have learnt about how CATs are being used in different clinical populations and their benefits over other measurement tools in terms of precision and brevity. Participants will also have learnt how results can be presented and clinically interpreted.

### Individual presentations

#### **The Kids-CAT: a computer-adaptive tool to measure quality of life in children and adolescents. Development, validation, and implementation in clinical settings and population health reporting**

**Christiane Otto, PhD**, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; **Dana Barthel, PhD**, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; **Janine Devine, PhD**, Charité - Universitätsmedizin Berlin, Berlin, Germany; **Felix Fischer, PhD**, Charité - Universitätsmedizin Berlin, Berlin, Germany; **Otto Walter, PhD**, Charité - Universitätsmedizin Berlin, Berlin, Germany; **Sandra Nolte, PhD**, Charité - Universitätsmedizin Berlin, Berlin, Germany; **Holger Mühlen, PhD**, Ernst-Moritz-Arndt Universität Greifswald, Greifswald, Germany; **Annett Mierke, Charité - Universitätsmedizin Berlin**, Berlin, Germany; **Kathrin Irmgard Fischer, Charité - Universitätsmedizin Berlin**, Berlin, Germany; **Ann-Katrin Meyrose, PhD**, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; **Ute Thyen, PhD, Prof.**, University Medical Center Schleswig-Holstein, Lübeck, Germany; **Marcus Klein, MD**, University Medical Center Schleswig-Holstein, Kiel, Germany; **Silke Schmidt, PhD, Prof.**, Ernst-Moritz-Arndt Universität Greifswald Greifswald, Germany; **Matthias Rose, PhD, Prof.**, Charité—Universitätsmedizin Berlin, Berlin, Germany; **Ulrike Ravens-Sieberer, PhD, Prof.**, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

**Aims:** Item banks measuring child patient-reported-outcomes (PRO) have recently been developed; however, only few CAT tools are available to assess pediatric HRQOL efficiently and precisely. We aim at describing the development, validation, feasibility, acceptance and application of the “Kids-CAT,” the first computer-adaptive test measuring generic self-reported health-related quality of life (HRQoL) in chronically ill and healthy children and adolescents in Germany. The Kids-CAT was further administered in a large German

health survey in 7- to 17-year-olds from the general population, results are presented. **Methods:** In line with the US PROMIS initiative, methods of classical test and item response theories were used for item bank development of the Kids-CAT, including items of well-established measures. Kids-CAT dimensions were developed based on the structure of the European KIDSCREEN questionnaire. A child-friendly design and an immediate feedback report for physicians (the Kids-CAT Report) were created. The Kids-CAT was administered in a longitudinal prospective study at University Medical Centers in  $n = 312$  chronically ill children. Feasibility, acceptability (following a multimethod research design), and psychometric properties were investigated. The Kids-CAT was then implemented in the national German child health survey ( $n = 1483$ ). **Results:** The five Kids-CAT dimensions Physical Well-being, Psychological Well-being, Parent Relations, Social Support & Peers and School Well-being include item banks of 26 to 46 items each and show high content validity, unidimensionality, local independence, low DIF, and model conform IRCs. On average 4 to 6 items were administered with a reliability of 0.9 (SEM < 0.32). Median item response time varied with age and reading abilities (2–3 min per item bank). Kids found the tool easy to complete, pediatricians emphasize the benefit of its report for patient-doctor interaction. The Kids-CAT measures reliably, particularly in lower areas of HRQoL. Support for its convergent and discriminant validity was found in correlations to well-established measures. Further results from the child health population survey on association to illnesses, mental health, and health utilization will be presented. **Conclusion:** The Kids-CAT advances HRQoL assessment in routine pediatric care and health monitoring by allowing a precise and valid measurement, making it less burdensome for respondents, and enhancing the patient-doctor communication via instant score reports.

#### **From translation to implementation: the challenge of PROMIS® CATs in daily pediatric practice**

**Michiel Luijten**, Emma Children’s Hospital, Amsterdam UMC, Amsterdam, Netherlands; **Caroline Terwee, Dr.**, Amsterdam UMC, Amsterdam, Netherlands; **Raphaële van Litsenburg, Dr.**, Princess Máxima Center for Pediatric Oncology, Utrecht, Netherlands; **Maud van Muilekom, MSc.**, Emma Children’s Hospital, Amsterdam UMC, Amsterdam, Netherlands; **Hedy van Oers, Dr.**, Emma Children’s Hospital, Amsterdam UMC, Amsterdam, Netherlands; **Martha Grootenhuus, Professor**, Princess Máxima Center for Pediatric Oncology, Utrecht, Netherlands; **Lotte Haverman, Dr.**, Emma Children’s Hospital, Amsterdam UMC, Amsterdam, Netherlands

**Aims:** The Patient-Reported Outcomes Measurement Information System (PROMIS) is a collection of multiple item banks measuring generic domains of physical/mental/social health, using item response theory (IRT) modeling. These item banks can be administered with Computerized Adaptive Testing (CAT) to decrease the burden and repetitiveness of questionnaires for patients. In 2009 the Dutch-Flemish PROMIS National Center was founded. The pediatric section of this center aims to implement the pediatric PROMIS item banks within clinical care and research to standardize the use of pediatric PROMIS and minimize barriers. **Methods:** To implement Pediatric PROMIS in the Netherlands, we (1) translated the PROMIS pediatric item banks; (2) collected normative data and validated the item banks in the general population; (3) are working on validating the item banks in several clinical populations and assessed the efficiency of CATs compared to short forms, (4) assessed the optimal feedback option for CATs and (5) implemented the items banks and feedback option into the KLIK PROM portal. **Results:** (1) There are currently twelve PROMIS pediatric item banks translated: Anger/Anxiety/

Depression/Fatigue/Mobility/Pain Interference/Peer Relationships/Upper Extremity/Sleep Disturbance/Sleep-Related Impairment/Global Health. (2) Normative values are available as the item banks have been validated in general population samples ( $n = 1082$ ) and (3) in a clinical sample (Juvenile Idiopathic Arthritis,  $n = 155$ ) where the PROMIS item banks displayed good discriminative/concurrent validity, reliability and post hoc CAT simulations outperformed short forms. We are currently validating all item banks in pediatric sickle cell disease/hemophilia patients. (4) During focus groups health care providers reported to prefer literal representation of item responses in traffic light colors. Graphs of T-scores should have a colored background that displays severity and a consistent display of directionality (upward trend = improvement). (5) To implement the PROMIS item banks and administer CATs within KLIK a connection was built between the portal and the Assessment Center of the Dutch-Flemish PROMIS National Center. Conclusion: The PROMIS pediatric item banks are valid and reliable in the Dutch population and are available as CATs within the KLIK PROM portal ([www.hetklikt.nu](http://www.hetklikt.nu)) for use in clinical practice as well as for research.

**Comparison of PROMIS computerized adaptive testing-administered item banks versus fixed short forms in Juvenile Myositis**

Ruchi Patel, BS, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Valeria Esparza, BS, University of Illinois-Chicago College of Medicine, Chicago, Illinois, United States; Jin-Shei Lai, PhD, OTR/L, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Elizabeth Gray, Northwestern University Feinberg School of Medicine, MS, United States, Illinois, Chicago, Rowland Chang, MD, MPH, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; David Cella, PhD, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; **Kaveh Ardalan, MD MS**, Duke University School of Medicine, Durham, North Carolina, United States

**Aims:** Juvenile myositis (JM) negatively impacts health-related quality of life (HRQoL). Legacy measures underestimate JM’s impact due to floor/ceiling effects. Patient-Reported Outcomes Measurement Information System® (PROMIS) measures have undergone initial validation in pediatric rheumatic diseases, including JM, but benefits of computerized adaptive testing (CAT) vs fixed short forms (FSF) are unknown. This study compares PROMIS CAT and FSF in JM. **Methods:** JM patients (5-17 yo) and their parents were recruited. Demographic and clinical assessments were collected, including: Physician Global Assessment of Disease Activity; Disease Activity Score; muscle enzymes; and Childhood Myositis Assessment Scale. Patients (8–17 yo) self-reported and parents (of patients 5–17 yo) proxy-reported both PROMIS CAT and FSF versions of Fatigue, Pain Interference, Upper Extremity Function, Mobility, Anxiety and Depressive Symptoms. To study extreme versus middle scorers, participants were split into three groups by PROMIS CAT T-scores (< 45, 45–55, > 55). Pearson correlations, paired t-tests, and Cohen’s d were used to compare PROMIS CAT and FSF for the entire cohort and for each T-score grouping. **Results:** Data from 67 patient–parent dyads were analyzed. Most patients were 8–17 yo ( $n = 49$ ; 73%), juvenile dermatomyositis ( $n = 61$ ; 91%), female ( $n = 56$ ; 84%), and white ( $n = 51$ ; 76%). Median [IQR] age of onset was 5.2 [3.9, 7.1] and age at initial study visit was 11.8 [7.4, 15.2]. Clinical measures showed low disease activity in this prevalence sample, e.g., median muscle enzyme values in normal range. PROMIS CAT and FSF highly correlated (Pearson’s 0.79–0.92) (Table 1). Mean CAT and FSF scores were not significantly different, except parent-proxy anxiety and fatigue, with modest effect sizes (0.508 and 0.317,

respectively). Correlations between CAT and FSF varied across domains, patient/parent report, and T-score groupings (Table 2). Scatterplots show floor/ceiling effect at the less symptomatic extreme in all FSF domains (Fig. 1). Conclusion: PROMIS CAT is feasible and comparable to FSF. CAT had less pronounced floor/ceiling effects than FSF, detecting individual differences in low symptom/disability scorers. CAT is recommended for long-term follow-up of JM patients since deconditioning often persists in remission. Future studies should focus on multicenter replication, benefits of CAT vs FSF in patients with more severe symptoms, and clinical interpretability.

Table 1: Comparison of PROMIS Computerized Adaptive Testing (CAT)-Administered Item Banks and Fixed Short Forms (FSF)

Domain	CAT			FSF			CAT vs FSF			
	n	Median # Questions Administered [IQR]	Median T-Score* [IQR]	n	# Questions Administered	Median T-Score* [IQR]	Pearson's Correlations	t-Test (p-value)	Cohen's D (effect size)	
CHILD	Anxiety	49	12 [6, 12]	38.7 [31.9, 47.1]	48	8	33.5 [33.5, 44.6]	0.91	0.7864873	0.0393214
	Depression	49	6 [5, 12]	44.6 [31.8, 55.0]	48	8	45.1 [35.2, 52.4]	0.89	0.9493802	0.092124
	Fatigue	49	11 [6, 12]	36.1 [25.6, 51.1]	49	10	37.3 [30.4, 46.4]	0.92	0.6970846	0.1678414
	Mobility	49	12 [5, 12]	56.4 [46.4, 61.7]	49	8	58.5 [48.2, 58.5]	0.92	0.9422712	-0.0766145
	Pain Interference	49	12 [5, 12]	32.0 [32.0, 44.8]	49	8	34.0 [34.0, 43.4]	0.87	0.4046734	0.1201629
	Upper Extremity	49	12 [12, 12]	57.3 [46.0, 57.3]	49	8	56.7 [56.7, 56.7]	0.92	0.0781128	-0.2571733
PARENT-PROXY	Anxiety	67	5 [5, 12]	44.4 [33.7, 52.2]	66	8	34.6 [34.6, 47.6]	0.90	0.0001666	0.5080356
	Depression	67	5 [5, 12]	44.7 [33.2, 53.6]	66	6	36.2 [36.2, 48.9]	0.79	0.0556075	0.2399114
	Fatigue	67	5 [5, 12]	43.9 [32.9, 52.4]	66	10	38.8 [34.1, 49.0]	0.83	0.0123982	0.3167088
	Mobility	67	12 [5, 12]	54.0 [43.6, 60.2]	66	8	56.5 [44.2, 56.5]	0.84	0.5333983	0.0770758
	Pain Interference	67	12 [5, 12]	40.6 [36.8, 50.5]	66	8	37.8 [37.8, 47.6]	0.82	0.1239677	0.1918397
	Upper Extremity	67	12 [5, 12]	47.1 [38.6, 55.7]	66	8	54.7 [39.8, 54.7]	0.86	0.2806785	-0.1339051

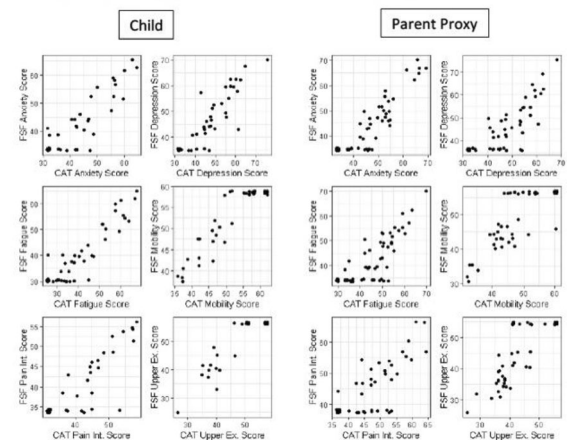
\*higher score for anxiety, depression, fatigue and pain interference indicates worse symptoms while higher score for mobility and upper extremity indicates better function

Table 2: Comparison of CAT and FSF by specific T-score ranges

Group (by T-Score)	Anxiety			Depression			Fatigue			Mobility			Pain Interference			Upper Extremity								
	n	Corr*	p†	ES‡	n	Corr*	p†	ES‡	n	Corr*	p†	ES‡	n	Corr*	p†	ES‡	n	Corr*	p†	ES‡				
CHILD	28	0.86	0.38	-0.20	35	0.66	0.05	-0.42	34	0.61	0.11	-0.27	9	0.71	0.51	-0.11	18	0.71	0.51	-0.11	11	0.76	0.21	-0.19
CHILD	7	0.41	0.19	0.34	12	0.59	0.18	0.41	6	0.86	0.05	1.05	12	0.71	0.05	-0.16	7	-0.02	0.26	0.47	7	0.63	0.05	-1.05
CHILD	9	0.64	0.08	0.33	11	0.58	0.33	0.31	9	0.41	0.03	0.92	28	N/A*	0.00	0.67	4	0.44	0.03	2.03	31	N/A*	N/A*	1.04**
PARENT-PROXY	34	0.81	0.34	0.31	33	0.70	0.13	-0.06	33	0.63	0.04	-0.16	19	0.83	0.02	-0.56	41	0.64	0.05	-0.32	38	0.70	0.08	-0.14
PARENT-PROXY	25	0.72	0.00	1.02	30	0.73	0.00	0.90	31	0.52	0.06	1.07	30	0.81	0.07	-0.43	17	0.19	0.01	0.14	7	0.39	0.44	-0.31
PARENT-PROXY	9	0.80	0.99	0.00	11	0.63	0.04	-0.02	10	0.86	0.03	0.89	27	N/A*	0.00	2.06	8	0.44	0.16	0.34	31	N/A*	N/A*	1.04**

\*Pearson Correlations between FSF and CAT  
 †Paired t-test p-value  
 ‡Effect size = mean difference/standard deviation of difference  
 \*N/A = not applicable to CAT and FSF values were the same  
 \*\*tst = t-value

Figure 1: Child self-report and parent-proxy report PROMIS CAT vs FSF score scatterplots by domain



**Using CATs to measure symptom burden reported by children with brain tumors**

Jin-Shei Lai, Northwestern University, Chicago, Illinois, United States; John Devin Peipert, Northwestern University, Chicago,

Illinois, United States; Stewart Goldman, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, Illinois, United States

**Aims:** Children with brain tumors (BT) could experience symptom burden throughout their disease continuum, from on-therapy to long-term survivorship. Monitoring of health-related quality of life (HRQL) and symptoms of patients with BT is needed yet not always feasible using current approaches. This is partially due to lack of brief-yet-precise assessments with minimal administration burden that are easily incorporated into clinics. Dynamic computerized adaptive testing (CAT) or static fixed-length short forms, derived from psychometrically sound item banks, were designed to fill this void. This study evaluated symptom burden experienced by children with BT using pediatric PROMIS (Patient-Reported Outcomes Measurement Information System) CATs and the potentially influential factors. **Methods:** Data from 230 children with BT aged 7–22 (mean age = 14 year; 52% boys; 76% white) were analyzed. Average years since last treatment was 2.6 (87% ≤ 1 year). Symptom burden was assessed Pediatric PROMIS CATs—Anxiety, Depression, Fatigue, Mobility, Upper Extremity Function (UE), Peer Relationship (PR), and Cognition. Patients and parents completed Symptom Distress Scales (SDS). Test-statistics and ANOVA were used to evaluate relationships between PROMIS measures and potentially influential variables. **Results:** Participants completed each CAT within 2 min. Significant results ( $p < 0.01$ ) showing impact of symptom burden included: (1) all PROMIS measures were correlated with SDS reported by patients and parents; (2) Fatigue, Mobility and UE were associated with Karnofsky functional performance status, number of treatment modalities (0–3), and time since last treatment (≤ 1 year, > 1 year); (3) Fatigue and Cognition were associated with educational program (regular classroom without an Individualized Education Plan (IEP) versus those that had an IEP); (4) Mobility and UE were associated with time since last radiation; and 5) Mobility, UE, and Anxiety were associated with time since last chemotherapy. **Conclusion:** Treatment type and time since treatment impacted burden. Significant planned associations were found between PROMIS measures and other variables, including SDS, functional performance, and educational programs. Given the brevity of administering CATs in the clinical setting, results from this study support using PROMIS CATs to comprehensively evaluate BT's symptom burden their follow-up care.

#### **Application of the Kids-CAT in the clinical setting: predictors of health-related quality of life in children and adolescents with chronic diseases**

**Kathrin Fischer**, Charité Universitätsmedizin Berlin, Berlin, Germany; Dana Barthel, PhD, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Felix Fischer, PhD, Charité Universitätsmedizin Berlin, Berlin, Germany; Christiane Otto, PhD, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Ute Thyen, Prof., Universität zu Lübeck, Lübeck, Germany; Marcus Klein, MD, Christian-Albrechts-Universität, Kiel, Germany; Otto Walter, PhD, Charité Universitätsmedizin Berlin, Berlin, Germany; Matthias Rose, Prof., Charité Universitätsmedizin Berlin, Berlin, Germany; Ulrike Ravens-Sieberer, Prof., University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Sandra Nolte, PhD, Charité Universitätsmedizin Berlin, Berlin, Germany

**Aims:** The Kids-CAT, a computer-adaptive test measuring generic health-related quality of life (HRQL) has been implemented in two specialized outpatient clinics in chronically ill children and adolescents. This study investigates HRQL and predictors of HRQL in young patients over the course of 6 months. Moreover, we examined the association between the blood glucose level (HbA1c level) and Kids-CAT domains in a subsample of children and adolescents with diabetes

over time. **Methods:** The Kids-CAT covers the domains physical well-being, psychological well-being, parent relations, social support & peers, and school well-being. Effects of sociodemographic, disease, health-related, and psychosocial factors on HRQL according to the Kids-CAT were investigated by means of individual growth modeling in the mixed clinical sample of 7 to 17 year-olds ( $n = 248$ ) using longitudinal data. In the diabetic subsample ( $n = 203$ ), path analyses were performed to explore the association between HRQL and HbA1c level over 6 months including three measurement points. **Results:** HRQL in young patients was comparable to an age-matched German-speaking reference population. The predictor disease control was positively associated with physical and psychological well-being, whereas health complaints were negatively related to all Kids-CAT domains over time. Only a minimal relationship between HbA1c and the Kids-CAT domains was found indicating a small negative impact of HbA1c on the domains' physical well-being, psychological well-being, and parent relations. **Conclusion:** Children and adolescents with chronic conditions reported good HRQL. Factors such as disease control, health complaints or clinical parameters, such as HbA1c, can impact HRQL over time and should be considered in pediatric health care. In light of the minimal association between HbA1c and HRQL underscores the added value of evaluation PROs in addition to classical clinical outcomes.

## **Oral Sessions**

### **101: PROs in Cancer Research I**

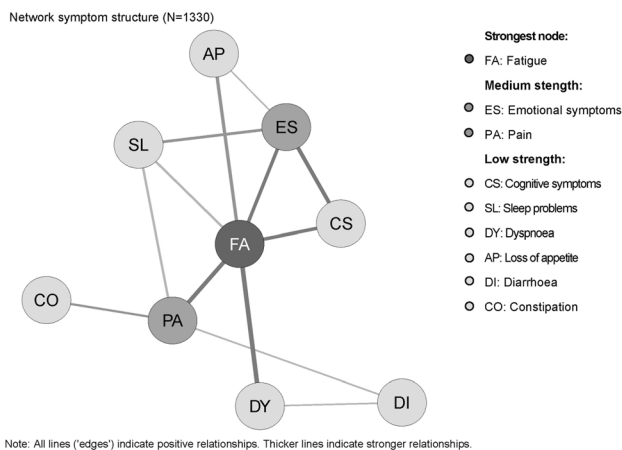
#### **(101.1) Symptom clusters in survivors of 7 cancer types from the PROFILES registry: a network analysis**

Belle de Rooij, PhD, Netherlands Comprehensive Cancer Organisation (IKNL), Eindhoven, Netherlands; Simone Oerlemans, PhD, Netherlands Comprehensive Cancer Organisation (IKNL), Eindhoven, Netherlands; Floortje Mols, PhD, Department of Medical and Clinical Psychology, Tilburg University, Tilburg, Netherlands; Olga Husson, PhD, Division of Psychosocial Research and Epidemiology, The Netherlands Cancer Institute, Amsterdam, Netherlands; Nicole P.M. Ezendam, PhD, Netherlands Comprehensive Cancer Organisation (IKNL), Eindhoven, Netherlands; Kelly M. de Ligt, PhD, Netherlands Comprehensive Cancer Organisation (IKNL), Eindhoven, Netherlands; Meeke Hoedjes, PhD, Department of Medical and Clinical Psychology, Tilburg University, Tilburg, Netherlands; Katrijn Van Deun, PhD, Department of Methodology and Statistics, Tilburg University, Tilburg, Netherlands; Lonneke V. van de Poll-Franse, PhD, Netherlands Comprehensive Cancer Organisation (IKNL), Eindhoven, Netherlands; Dounya Schoormans, PhD, Department of Medical and Clinical Psychology, Tilburg University, Tilburg, United Kingdom

**Aims:** Frequently reported symptoms among cancer survivors, including fatigue, pain and cognitive problems, often co-occur in the same individual. However, research has mostly focused on distinct symptoms - once at a time, as cause or effect - ignoring the complex interplay between symptoms. Research into the clustering of symptoms may improve our understanding of the underlying mechanisms that impact survivors' symptom burden. We applied network analysis in a balanced sample of cancer survivors to: 1) explore the clustering of symptoms; and 2) assess differences in symptom clustering between cancer types. **Methods:** We used cross-sectional survey data from the population-based PROFILES registry, collected between 2008 and 2018, including survivors of 7 cancer types (colorectal, breast, ovarian, thyroid, chronic lymphocytic leukemia [CLL], Hodgkin lymphoma and non-Hodgkin lymphoma). Using network analysis, we explored and visualized the associations between self-



reported symptoms (EORTC QLQ-C30) and the centrality of these symptoms in the network (i.e., how strong a symptom is connected to other symptoms), for the total sample and each cancer type separately. Results: Within our sample ( $n = 1330$ ), fatigue was the most central symptom in the network with moderate to strong direct relationships with dyspnea ( $r = 0.34$ ), pain ( $r = 0.30$ ), cognitive symptoms ( $r = 0.24$ ), emotional symptoms ( $r = 0.24$ ), lack of appetite ( $r = 0.23$ ) and sleep problems ( $r = 0.15$ ). In addition, a strong relationship was found between emotional and cognitive symptoms ( $r = 0.28$ , Fig. 1). These relationships persisted after adjustment of sociodemographic and clinical characteristics. Connections between fatigue and dyspnea, pain, emotional symptoms and lack of appetite were consistently found across all cancer types ( $n = 190$  each). In CLL patients, fatigue and cognitive symptoms were not directly connected, but indirectly through an additional connection between pain and cognitive symptoms. Survivors receiving chemotherapy showed a similar network compared to those who did not. Survivors receiving radiotherapy ( $n = 493$ ) showed an additional direct connection between lack of appetite and cognitive symptoms ( $r = 0.14$ ). Conclusion: In a heterogenous sample of cancer survivors, fatigue consistently clustered with dyspnea, pain, emotional symptoms, and lack of appetite. Although longitudinal data are needed to build a case for the causal nature of these symptoms, fatigue could be a starting point for interventions to reduce the overall symptom burden of cancer survivors.

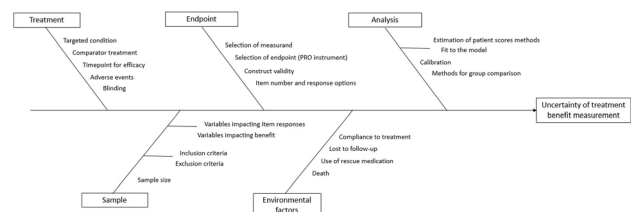


### (101.2) The MD Anderson Symptom Inventory (MDASI) disturbed sleep item as a possible rapid and efficient screener of sleep quality among early-phase clinical trials clinic patients with advanced cancer

Goldy George, PhD, The University of Texas MD Anderson Cancer Center, Houston, Texas, United States; Tito Mendoza, PhD, The University of Texas MD Anderson Cancer Center, Houston, Texas, United States; Eucharia Iwuanyanwu, The University of Texas MD Anderson Cancer Center, Houston, Texas, United States; Sarina Piha-Paul, MD, The University of Texas MD Anderson Cancer Center, Houston, Texas, United States; Aung Naing, MD, The University of Texas MD Anderson Cancer Center, Houston, Texas, United States; Charles Cleeland, PhD, The University of Texas MD Anderson Cancer Center, Houston, Texas, United States; David Hong, MD, The University of Texas MD Anderson Cancer Center, Houston, Texas, United States

Aims: Poor sleep quality is prevalent in patients with advanced cancer in early-phase clinical trials. However, extant validated sleep measures are often time-intensive and difficult for patients with advanced

cancer to complete, and complicated for clinicians to administer and compute in busy clinical trial settings. We investigated whether the MDASI single “disturbed sleep” item might provide initial screening for poor sleep quality in patients with advanced cancer in early-phase clinical trials clinics. Methods: Patients completed the validated core MDASI that included the “disturbed sleep” single-item, and 12 other symptoms (each item was rated on a 0-10 scale; higher scores indicated worse severity). Patients also completed the 19-item Pittsburgh Sleep Quality Index (PSQI), a validated and global measure of sleep quality. Statistical computations included Spearman’s rho, and receiver operating characteristic (AUC-ROC) curves to calculate area under the curve (AUC), sensitivity, specificity, positive (PPV) and negative (NPV) predictive values. Results: Early-phase clinical trial clinic patients ( $n = 246$ , 52% female, 79% White, 87% age  $\geq 45$  years) reported a median MDASI disturbed sleep item score of 2 (IQR = 5). The MDASI disturbed sleep item was effective in distinguishing “poor sleepers” with global PSQI score  $> 5$  (AUC = 0.78). At a cut point of 3 or greater, the MDASI “disturbed sleep” item exhibited sensitivity = 82.0%, specificity = 60%, NPV = 85.0%, and PPV = 55%. The PSQI component subscales that the MDASI disturbed sleep item correlated most with were the subjective sleep quality ( $\rho = 0.63$ ,  $p < 0.001$ ), sleep latency ( $\rho = 0.46$ ,  $p < 0.001$ ), and sleep disturbance ( $\rho = 0.41$ ,  $p < 0.001$ ) domains. Worse MDASI disturbed sleep was linked to worse MDASI distress ( $\rho = 0.53$ ,  $p < 0.001$ ), drowsiness ( $\rho = 0.52$ ,  $p < 0.001$ ), fatigue ( $\rho = 0.48$ ,  $p < 0.001$ ), sadness, nausea, and pain ( $\rho = 0.42$ ,  $p < 0.001$  for each). Conclusion: The MDASI “disturbed sleep” item has sufficient sensitivity and negative predictive value and can serve as an efficient preliminary screen for poor sleep quality in early-phase clinical trial clinics. This research reiterates that patient-reported single-item outcome measures can provide rapid and effective screening of selected symptoms in patients with advanced cancer. This is relevant in situations such as the current COVID-19 pandemic where health resources are overextended, clinician-patient interactions more limited, and stress-induced sleep disturbance in patients may be more prevalent.



### (101.3) Development and psychometric validation of BREAST-Q scales measuring cancer worry, fatigue and work life impact

Manraj Kaur, PhD, Brigham and Women’s Hospital, Boston, MassachusettsOntario, United States; Anne Klassen, McMaster University, Hamilton, Ontario, Canada; Elena Tsangaris, Brigham and Women’s Hospital, Boston, Massachusetts, United States; Louise Bordeleau, McMaster University, Hamilton, Ontario, Canada; Toni Zhong, University of Toronto, Toronto, Ontario, Canada; Stefan Cano, Modus Outcomes, Toronto, United Kingdom; Trisia Breitkopf, McMaster University, Hamilton, Ontario, Canada; Andrea Pusic, Brigham and Women’s Hospital, Boston, Massachusetts, United States

Aims: The BREAST-Q is a modular patient-reported outcome instrument for breast cancer surgery used in research and clinical practice. The aim of this study was to develop and validate three new BREAST-Q scales that assess the impact of breast cancer diagnosis or treatments on cancer worry, fatigue, and work life. Methods:

Interpretive descriptive approach was used to conduct in-depth, semi-structured, and qualitative interviews with a heterogeneous sample of adult women (18 years and older) diagnosed with breast cancer (any stage, any treatment). The participants were recruited from three tertiary cancer centers (two in Canada, one in the United States). The interviews were audio-recorded, transcribed and coded using line-by-line approach. Constant comparison was used to refine the codes. The item pool led to the development of the three new breast cancer scales that are currently not a part of the BREAST-Q—cancer worry, fatigue, and work life impact. The new scales were refined through three rounds of one-on-one cognitive interviews with patient participants and one round of expert feedback obtained through REDCap. The scales were field-tested in sample recruited from the Army of Women, an online community of women, and Rasch measurement theory (RMT) analysis was used to refine the scales and examine their psychometric properties. Results: Interviews with 57 women with breast cancer (mean age,  $55 \pm 10$  years) were used to develop the three new BREAST-Q scales. Feedback from 9 patients and 23 experts from eight countries was used to refine the scale's instructions, items, and response options. An RMT analysis of BREAST-Q new scales data from 1680 women (one-week test–retest,  $n = 1006$ ) showed that all items had ordered thresholds, mapped out a targeted clinical hierarchy, and had good test–retest reliability (intraclass correlation co-efficient,  $> 0.84$ ). The reliability statistics for the person separation index was  $> 0.81$  and for Cronbach's alpha was  $> 0.89$ . The items in the BREAST-Q scales were highly correlated to similar items in the Euro-Qol-5D and EORTC-8D questionnaires. Conclusion: The new BREAST-Q scales can be used in clinical care and research to assess cancer worry, fatigue, and work life impact in women with breast cancer, irrespective of the stage of breast cancer or type of treatments.

#### **(101.4) Identifying difficulties and most-interested findings for analyzing patient-reported outcome data from perioperative settings**

Qiuling Shi, Chongqing Medical University, Chongqing, China; Larissa A Meyer, MD, University of Texas MD Anderson Cancer Center, Houston, Texas, United States; Wei Dai, MD, Sichuan Cancer Hospital, Sichuan, China; Wei Xu, MS, Chongqing Medical University, Chongqing, China; Yuxian Nie, MS, Chongqing Medical University, Chongqing, China; Charles S Cleeland, PhD, University of Texas MD Anderson Cancer Center, Chongqing, Texas, United States; Xin S Wang, MD, University of Texas MD Anderson Cancer Center, Houston, Texas, United States

Aims: Longitudinal assessments of patient-reported outcomes (PROs) are becoming more integrated into patient care and research in the perioperative setting. This increased application of PROs is resulting in enormous demand for statistical support. As part of an organized effort to establish programmatic guidance on PRO data procedures, we conducted a professional survey to identify both difficulties and areas of clinical and research interest to guide perioperative PRO analyses. Methods: We surveyed researchers and clinicians from a department of thoracic surgery at a tertiary hospital involved in a series of clinical research projects that utilized PROs as major outcomes. Expert-generated questions with a 0–10 scale were used for defining two domains: difficulties (23 items, 0 = no difficulty, 10 = as difficult as you can imagine) and interest of findings (10 items, 0 = not interested; 10 = as interested as you can imagine) in PRO data analysis. We identified the top 5 items with the highest mean scores for each domain. Results: A total 25 of 28 approached clinicians and researchers (89.3%) responded to the survey, including 15 (60%) surgical professionals and 4 (16%) data analysts. The 5 most difficult tasks were modeling longitudinal data (mean  $\pm$  SD  $7.57 \pm 2.25$ ), planning further analysis for

negative findings ( $7.09 \pm 1.86$ ), dealing with missing data ( $6.57 \pm 1.73$ ), searching guidelines for data analysis ( $6.43 \pm 2.01$ ), and choosing figures and tables for publication ( $6.43 \pm 2.01$ ). The 5 areas of greatest interest included relationship between PROs and clinical outcomes ( $8.74 \pm 1.25$ ), trajectories of PROs over the course of recovery ( $8.56 \pm 1.31$ ), relationships between symptoms and functioning/quality of life ( $8.48 \pm 1.38$ ), preoperative baseline levels of PROs ( $8.35 \pm 1.74$ ), and time to alleviation of symptoms ( $8.26 \pm 1.28$ ). Conclusion: This study suggests that professionals are experiencing difficulties in handling PROs data analysis in perioperative research and clinical care and highlights the need for more formalized guidance and statistical support. Results from this survey will inform the establishment of standardized recommendations and detailed guidance for PRO data analysis to encourage and support well-designed research and to promote the implementation of PROs into surgical practice.

#### **102: Advancing theory in the analysis of PRO data**

##### **(102.1) Contemplating clinical trials as measurement systems: A new perspective for measurement uncertainty associated with treatment benefit demonstration**

Angély Loubert, PharmD, MSc, Modus Outcomes, Lyon, France; Antoine Regnault, PhD, Modus Outcomes, Lyon, France; Véronique Sébille, PhD, UMR INSERM 1246—SPHERE, Nantes, France; Jean-Benoit Hardouin, PhD, UMR INSERM 1246—SPHERE, Nantes, France; Jeanette Melin, PhD (Medical sciences), RISE Metrology, Gothenburg, Sweden; Stefan Cano, PhD, CPsychol, AFBPsS, Modus Outcomes, Gothenburg, United Kingdom; William P. Fisher, Jr., PhD, Living Capital Metrics LLC, Sausalito, California, United States

Aims: The objective of a clinical trial is to generate supportive evidence on the benefit of a new treatment. For this purpose, the treatment effect is defined, modeled, estimated, evaluated, and interpreted as a unit quantity. From the perspective of measurement science and following the International Vocabulary of Metrology (VIM), results from a clinical trial can be seen as a measured quantity value from a measurement system, which comes with uncertainty of measurement. Our objective was to develop a conceptual model describing clinical trials as a measurement system and list the uncertainty sources associated with it, in the specific case of a patient-reported outcome (PRO) endpoint. Methods: The metrological principles described in the Guide to the expression of Uncertainty in Measurement (GUM) were applied to the case of clinical trials. Specifically, we searched for the various sources of uncertainty pertaining to the measurement result. This was achieved by exploring systematically the various aspects of clinical study design and the available literature. The sources of uncertainty identified were classified in a typology. This exercise was applied to the case of a trial with a primary PRO endpoint, and within the framework of Rasch measurement theory (RMT). Results: A graphical conceptual model of clinical trials as a measurement system was created. The links with the notion of estimand recently introduced for the analysis of clinical trials were examined. The uncertainty sources identified were classified in the following categories: sampling; treatment; endpoint; environmental factors; and statistical analysis method. An example, summarized Ishikawa diagram derived from the model is displayed in the attached figure. Conclusion: Our theoretical model describing a clinical trial as a measurement system provides a new perspective for clinical trial design, with an holistic approach to the various decisions to be made, underpinned by the common reference to measurement uncertainty. In practice, a typology of the sources of uncertainty relevant to a clinical trial was also created and will be used as a basis for future research that will attempt to quantify these sources in an uncertainty budget.

### **(102.2) Imputation strategies within the estimand framework to evaluate the overall likelihood of patient improvement in longitudinal trails**

Lysbeth Floden, PhD, MPH, Clinical Outcomes Solutions, Tucson, Arizona, United States; Stacie Hudgens, MA, Clinical Outcomes Solutions, Tucson, Arizona, United States; Hailin Yu, MPH, Clinical Outcomes Solutions, Chicago, Illinois, United States; Melanie Bell, PhD, University of Arizona College of Public Health, Tucson, Arizona, United States

**Aims:** In clinical trials, patient-reported endpoints can be dichotomized by a meaningful within-patient change threshold to compare proportions of patients who achieve meaningful improvement. Advances in methodology of principled approaches to analyze incomplete data highlight the need to better define trial estimands. Estimands encompass four components: the research objective, the target population, the analytical approach, and the handling of post-randomization events including non-compliance and dropout. In this study, we defined estimands for modeling the likelihood of meaningful response of patient-reported outcomes (PROs) and demonstrated the use of missing at random (MAR) and missing not at random (MNAR) imputation methods within an estimand framework. **Methods:** We simulated data including PRO adherence indicators and missingness due to dropout from a two-arm trial measuring a PRO score which was dichotomized then modeled using a generalized estimating equation (GEE). We defined and evaluated a de jure and de facto estimand, considering scenarios with missingness and protocol non-compliance. As a component of the estimand we specified a multiple imputation (MI) approach to evaluate improvement while on treatment (MI using the fully conditional specification), and to evaluate improvement considering real-world compliance (control-based MI using the fully conditional specification). We evaluated bias relative to the true value and linked the imputation method to the estimand. **Results:** When patient missingness was not related to adherence, the estimates from MI were similar to the true estimate from the population with fully observed values. When PRO missingness included non-adherence to the study protocol, the estimates using control-based MI were similar to the true estimates in scenarios with non-adherence. These of estimates differed from each other and suggest that the mechanism of missingness may be less important than defining the estimand and using an appropriate imputation approach. **Conclusion:** Likelihood of meaningful improvement using standard MI estimated the difference in PRO responders due to treatment taken as directed making it the best imputation choice for the de jure estimand. Likewise, likelihood of meaningful improvement using MNAR MI best characterized the difference in PRO responder proportions due to the treatment regimens (assuming those who drop out are nonadherent), making it a good choice for the de facto estimand.

### **(102.3) Response shift in Patient-Reported Outcomes Measures: a formal definition and a revised model**

Antoine Vanier, MD, PhD, Inserm U1246 SPHERE—University of Nantes, Nantes, France; Frans Oort, University of Amsterdam, Research Institute of Child Development and Education, Amsterdam, Netherlands; Leah McClimans, University of South Carolina, Department of Philosophy, Columbia, South Carolina, United States; Bernice G Gulek, University of Washington, Harborview Medical Center, Seattle, Washington, United States; Jan R Böhnke, University of Dundee, School of Health Sciences, Dundee, United Kingdom; Nikki Ow, McGill University, Center for Outcomes Research and Evaluation, Dundee, Quebec, Canada; Nancy Mayo, McGill University Health Centre Research Institute, Department of Medicine,

Division of Clinical Epidemiology, Montreal, Quebec, Canada; submitted on behalf of Response Shift—in Sync Working Group

**Aims:** Response shift has been defined as a change in the meaning of one's self-evaluation of a target construct. Up to 2009, theoretical models to explain response shift have been published, introducing new questions and dilemmas, and theoretical debates continue. To stimulate empirical research and to address these dilemmas, we propose a formal definition of response shift and a model depicting the components engaged in explaining both changes in the target construct (e.g., HRQoL) and in its measure (e.g., PROM score) at two points in time. **Methods:** This work is an international collaborative effort involving both experienced and new researchers on response shift. It involved a critical assessment of the literature, a two day face-to-face working group meeting and writing activities to draft a revised model and explanatory paper. **Results:** Three main dilemmas were identified. First, response shift definition can be confusing as the phenomenon is both described as a discrepancy between observed and target change and as an effect on the target construct itself. Second, previous models have explained change in the construct, but not the variability of the construct at each time of measurement which renders the chain of causality unclear. Third, extant models do not explicitly discriminate the measure from the construct. The formal definition and revised model aim to address these dilemmas. Here, we define response shift as an effect occurring whenever observed change (e.g., change in PROM scores) is not fully explained by target change (e.g., change in the construct intended to be measured). This discrepancy is illustrated in a revised model centered on a catalyst, personal and biopsychosocial factors, and mechanisms of adaptation, learning and growth, which are causally related in explaining the appraisal and variability of both the measure (e.g., PROM) and the target construct (e.g., HRQoL) at two times of measurement. **Conclusion:** This new model specifically differentiates between the multiple pathways leading to both direct (e.g., impact of the catalyst) and mediated effects (e.g., adaptation, response shift) on the target construct and its measure. This revised model will help clarifying the whole chain of causality explaining changes both in the target construct and in PROMs.

### **(102.4) Psychometric analyses of repeated measures using longitudinal extensions of the Rasch model and partial credit model with invariant frame of reference over time**

Sarah Marquis, Ph.D, University of California, Santa Barbara, Santa Barbara, California, United States; Antoine Regnault, Ph.D, Modus Outcomes, Lyon, France

**Aims:** Repeated administration of clinical outcome assessment (COA) over time is common in clinical research offering a wealth of potentially useful data for psychometric analysis, especially in contexts where any data are precious such as rare diseases. However, running analysis on longitudinal data raises some questions as possible time effects or within-individual correlations are a concern in this setting. Our objective was to demonstrate how specific longitudinal extensions of the Rasch model and partial credit model (PCM) could be useful for psychometric analyses on repeated measures. **Methods:** We developed longitudinal extensions amending the dichotomous Rasch model and PCM by adding a subject-dependent time parameter. This specification reflects the assumption that the item parameters should be fixed to provide a stable frame of reference over time. We formally examined the mathematical properties of these extensions, focusing on the key desirable properties of the family of the Rasch model. The extension of the PCM was also applied to the Unified Parkinson's Disease Rating Scale (UPDRS) data from the Parkinson's Progression Markers Initiative (PPMI)

study for illustration purposes. Results: Our extensions of the Rasch model and PCM were demonstrated to maintain the key properties of the Rasch models, namely parameter separation and statistical sufficiency for all three types of parameters (items, persons, and time). Sufficient statistics were formally derived for all parameters in the model. The independence of the item estimates from the time estimates warranted the analysis of repeated measures with psychometric purposes, such as calibration, with our model. The application of the extension of the PCM to the PPMI data showed the consistency of the UPDRS items estimates with the “standard” PCM applied to pooled data from all visits. Conclusion: The longitudinal extensions of the Rasch model and PCM will be useful when repeated measurements are used to perform psychometric analyses of the COA. Additionally, these longitudinal extensions will be appropriate to characterize how a targeted concept, defined by an invariant frame of reference, changes in a group of respondents over time.

### 103: Cancer Clinical Care

#### (103.1) Can you handle the truth- does sharing quality of life scores with oncologists affect patients’ responses?

Claire Snyder, PhD, Johns Hopkins, Baltimore, Maryland, United States; Amanda Blackford, Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, Maryland, United States; Kate Absolom, University of Leeds, Leeds, United Kingdom; Ethan Basch, University of North Carolina-Chapel Hill, Chapel Hill, North Carolina, United States; Galina Velikova, Leeds Institute of Cancer and Pathology and St James’s Institute of Oncology, Leeds, United Kingdom

Aims: Quality of life (QOL) measures can be used in clinical practice, with individual patient’s QOL scores reported to oncologists to inform care and management. However, questions have been raised about whether cancer patients report their QOL differently when they know their provider will see the scores, either under-reporting problems out of fear the provider might stop their therapy or over-reporting problems to get attention. Methods: We conducted a secondary analysis of data from a randomized controlled trial of patients commencing cytotoxic or biologic treatment who were expected to have at least 3 visits. Patients were randomized 2:1 to Feedback (scores shared with oncologists) and No Feedback (scores not shared with oncologists). Patients in both arms completed the EORTC QLQ-C30 (15 domains) and Hospital Anxiety & Depression Scale (2 domains) in the waiting room via touch-screen prior to each visit. Our primary analysis tested for differences in the QOL scores of Feedback vs. No Feedback patients at the first assessment using t-tests and linear regression models adjusting for performance status, with the primary interpretation focused on whether the 95% confidence intervals (CI) included an effect size of 0.5, a medium difference. Secondary comparisons examined longitudinal scores using the interaction p value to detect differences over time between arms. Results: In the Feedback Arm, 104 patients had 3 assessments, 5 only 2; in the No Feedback Arm, 47 patients had 3 assessments, 1 patient only 2. Across the 17 domain scores at the first assessment, the effect size 95% CI overlapped 0.5 for 6 domains, with the No Feedback Arm reporting better scores on 4 and the Feedback Arm better on 2. The only difference to reach statistical significance was less severe dyspnea reported in the No Feedback Arm ( $p = .01$ ). In the longitudinal analyses, no statistically significant differences between arms were found on the 17 domains. Conclusion: We found no evidence of systematic differences in the reporting of QOL based on whether the scores were shared with oncologists, suggesting that QOL scores can be fed back to providers for use in patient care and management without concern of biased reporting.

#### (103.2) Development and field testing of a patient-reported symptom index for use with non-muscle invasive bladder cancer patients using mixed methods

Claudia Rutherford, The University of Sydney, Sydney, Australia; Margaret-Ann Tait, The University of Sydney, School of Psychology, Sydney, Australia; Daniel Costa, Pain Management Research Institute, Royal North Shore Hospital, St Leonards, Australia; Madeleine King, The University of Sydney, School of Psychology, Sydney, Australia; David Smith, Cancer Council New South Wales, Sydney, Australia; Shomik Sengupta, Monash University, Eastern Health Clinical School, Sydney, Australia; Manish Patel, The University of Sydney, Sydney Medical School, Sydney, Australia; submitted on behalf of Response Shift—in Sync Working Group

Aims: Non-muscle invasive bladder cancer (NMIBC) is a chronic condition requiring treatment and lifelong monitoring with regular endoscopic examinations. In this clinical context, patient-reported outcomes (PROs) have enormous potential to inform treatment assessment and recommendations for NMIBC; however, current PRO measures are inadequate for NMIBC because they lack key NMIBC-specific symptoms and side-effects associated with contemporary treatments. We aimed to develop and evaluate a patient-reported NMIBC Symptom Index (NMIBC-SI) that was acceptable, reliable, valid, and responsive to treatment effects. Methods: We conducted a systematic review and interviewed 26 patients and 20 clinicians to develop a conceptual framework of PROs important to NMIBC. The 125 issues in the conceptual framework were phrased as questions and pre-tested in 12 cognitive interviews to develop a draft 104-item NMIBC-SI. In Field Test 1 (FT1), we administered the NMIBC-SI to patients on active treatment from nine Australian sites. NMIBC-SI item responses were considered for exclusion if they had low prevalence, were conceptually similar, or highly correlated ( $\geq 0.50$ ). Nine Urologists reviewed the results and final items for inclusion. In Field Test 2 (FT2), patients from 16 sites across four countries completed the final version NMIBC-SI at baseline and four follow-up times. Results: For FT1, we recruited  $n = 220$  (178 male, mean age 69) representing: Low 27.7%; Intermediate 13.2%; High 50.9% risk groups. 80% patients did not experience 21 items, 7 items were highly correlated, and 4 excluded as  $> 50\%$  of urologists rated them not related to NMIBC treatment. The final 56-item NMIBC-SI used in FT2 included a 23-item symptom burden scale, 2 treatment-specific modules, and 3 function scales. NMIBC-SI has currently been administered to  $n = 248$  newly diagnosed patients (186 male, mean age 67),  $n = 206$  before treatment,  $n = 170$  1-week post-surgery,  $n = 134$  end of induction therapy, and  $n = 34$  1-year post-treatment. Conclusion: The NMIBC-SI allows comprehensive assessment of patients’ self-reported symptom burden and functioning impairment. This prospective longitudinal study evaluates the validity and reliability of the NMIBC-SI, assessing key PROs across treatments, disease trajectory (acute to 1-year survivorship), and risk categories. The NMIBC-SI will be suitable for use in clinical practice and future clinical trials of treatments for NMIBC.

#### (103.3) Age and rural/urban disparities in access to and completion of electronic patient-reported outcomes (ePROs) in an EHR-facilitated Cancer Symptom Control (E2C2) trial

Joan Griffin, PhD, Mayo Clinic, Rochester, Minnesota, United States; Jennifer Ridgeway, PhD, Mayo Clinic, Rochester, Minnesota, United States; Kristin Fischer, MPH, Mayo Clinic, Rochester, Minnesota, United States; Lila Rutten, PhD, Mayo Clinic, Rochester, Minnesota, United States; Aaron Leppin, MD, MSc, Mayo Clinic, Rochester, Minnesota, United States; Nathan Tesch, MS, PMP, Mayo Clinic,

Rochester, Minnesota, United States; Amanda Nelson, BSW, Mayo Clinic, Rochester, Minnesota, United States; Sarah Redmond, PhD, Mayo Clinic, Rochester, Minnesota, United States; Casey Fazer, MPAS, PA-C, Mayo Clinic, Rochester, Minnesota, United States; Amy Johnson, APRN, CNP, Mayo Clinic, Rochester, Minnesota, United States; Margaret Wagnerowski, APRN, CNP, Mayo Clinic, Rochester, Minnesota, United States; Margaret Wagnerowski, MSN, RN, Mayo Clinic, Rochester, Minnesota, United States; Diedre Pachman, MD, Mayo Clinic, Rochester, Minnesota, United States; Kathryn Ruddy, MD, MPH, Mayo Clinic, Rochester, Minnesota, United States; Andrea Cheville, MD, MsCE, Mayo Clinic, Rochester, Minnesota, United States

**Aims:** Rural cancer survivors experience greater depression and anxiety, attend fewer healthcare visits, and receive less guideline-concordant care than urban cancer survivors. Elderly cancer patients face similar challenges, with greater chronic disease burden, loss of physical function and disability, and greater risk for drug interactions and toxic treatment side effects. Monitoring symptoms using electronic patient-reported outcomes (ePROs) may reduce these disparities by addressing manageable symptoms. However, ePRO systems may be differentially adopted by elderly and rural patients. **Methods:** The Enhanced, EHR-facilitated Cancer Symptom Control (E2C2) care model is a remote symptom monitoring and management system. Patients receive ePROs for Sleep, Pain, Anxiety, Depression, Energy deficit (SPADE) symptoms and limitations in physical function in their patient portal, and can respond electronically or in-clinic via tablet. Using discrete EHR-embedded algorithms, ePRO responses trigger evidence-based symptom-management approaches. Patients reporting moderate symptoms/limitations receive low-touch, automated self-management resources. Patients reporting intense (severe) symptoms and/or functional limitations receive nurse-managed collaborative care. As part of a Hybrid II stepped-wedge (five blocks) cluster-randomized pragmatic trial to evaluate E2C2, disparities in “intervention reach” (i.e., ePRO response rates, response mode, access to patient portal, and portal use) are being assessed among elderly and rural-dwelling patients with cancer. **Results:** From the first block of randomized clinical sites, nearly two-thirds of patients responded to ePROs, but only 52% of those did so via the patient portal. We found no differences in intervention reach variables between rural and urban patients. Differences were found by age, with younger patients more likely to have portal accounts (< 65, 88%; 65–74, 83%; > 75, 72%); use the portal (< 65, avg. 18 times in last 90 days; 65–74, 15 times; > 75, 13 times) and respond to ePROs using the portal rather than in clinic (< 65, 60%; 65–74, 51%; > 75, 41%). **Conclusion:** While ePRO response rates were relatively high in this first intervention block, priorities for further increasing completion of ePROs should focus on older patients. Failure to increase activation and portal use among this population could dilute the potential of a remote intervention to reduce disparities in cancer symptom management.

#### (103.4) Financial toxicity and the impact on health-related quality of life in patients with plasma cell disorders

Rahma Warsame, MD, Mayo clinic, Rochester, Minnesota, United States; Angela Dispenzieri, MD, Mayo Clinic, Rochester, Minnesota, United States; Amylou Dueck, PhD, Mayo Clinic, Scottsdale, Arizona, United States

**Aims:** Advances in novel therapies and supportive care have contributed to improved outcomes for patients with Multiple myeloma (MM) and Light Chain Amyloidosis (AL). This progress comes at a high price; with high costs to patients increasing distress, while decreasing compliance and even survival. We aimed to screen patients for financial toxicity and understand its impact on various domains of health-related quality of life (HRQOL). **Methods:** Prospective study of adult patients with MM or AL

that receive follow-up care at Mayo Clinic, Rochester, MN. Financial toxicity was measured using the COmprehensive Score for financial Toxicity (COST) questionnaire. HRQOL was measured using the Patient-Reported Outcome Measurement Information System (PROMIS)-29. Baseline demographic information and clinical characteristics were abstracted from the medical record. Statistical analysis included descriptive statistics, Spearman correlations, and comparison of COST scores between groups by Jonckheere-Terpstra, Kruskal–Wallis, and Wilcoxon rank-sum tests for ordered, unordered, and binary categorical variables, respectively. **Results:** To date 77 patients have been enrolled, 90% MM, 55% male, 52% age 65 or older, 91% white, 92% non-Hispanic, 56% at least college graduate, 47% income at least \$75,000USD/year, and 36% employed (49% retired). The mean COST score was 26.5 (SD 10.0) with 42% reporting high financial toxicity. Financial toxicity significantly differed by age, gender, education, and income (all  $p < 0.05$ ). PROMIS-29 domain descriptive statistics and correlations with financial toxicity appear in Table 1. **Conclusion:** Financial distress was prevalent in this well-educated and high-earning MM and AL cohort that is able to receive tertiary care at Mayo Clinic. Longitudinal assessment is ongoing.

**Table 1:** PROMIS-29 Descriptive Statistics and Spearman Correlations with COST Scores

PROMIS-29 Domain	N	Mean	Standard Deviation	Range	Spearman Corr	p-val
Physical Function T-score	77	45.7	9.1	3.0 - 44.0	0.12	0.28
Anxiety T-score	76	47.6	7.9	22.5 - 57.0	-0.35	0.002
Depression T-score	77	46.1	6.6	40.3 - 65.3	-0.28	0.01
Fatigue T-score	75	53.4	7.3	41.0 - 62.2	-0.03	0.79
Sleep Disturbance T-score	76	49.4	7.0	32.0 - 73.3	-0.05	0.67
Social Roles T-score	77	49.5	10.3	50.5 - 63.8	0.19	0.09
Pain Interference T-score	77	53.2	10.1	41.6 - 75.6	-0.24	0.04
Pain Intensity	76	2.6	2.2	0 - 8	-0.30	0.009

#### (103.5) Can Patient-Reported Outcome (PRO) measures used in clinical practice predict survival at disease progression in patients with advanced lung cancer

Rasmus Blechingberg Friis, MD, PhD student, Department of Oncology, Hospital Unit West Jutland, Herning, Denmark; Niels Henrik Hjøllund, MD, Professor, AmbuFlex/WestChronic, Occupational Medicine, University Research Clinic, Aarhus University, Herning and Department of Clinical Epidemiology, Aarhus University Hospital, Herning, Denmark; Helle Pappot, MD, Professor, Department of Oncology, University Hospital of Copenhagen, Copenhagen, Denmark; Gry Assam Taarnhøj, MD, Department of Oncology, University Hospital of Copenhagen, Copenhagen, Denmark; Jesper Medom Vestergaard, M.Sc.(I.T), Department of Occupational Medicine, Danish Ramazzini Centre, University Research Clinic, Herning, Denmark; Halla Skuladottir, MD, Med.Sc.D., Department of Oncology, Hospital Unit West Jutland, Herning, Denmark

**Aims:** Patient-Reported Outcome (PRO) measures used in clinical practice can improve supportive care, patient-physician communication, patient satisfaction and even survival. However, the predictive value of PRO measures as a tool for decision-making at the time of disease progression has not previously been described. This study aimed to explore the predictive value of PROs at the time of disease progression in patients with lung cancer. **Methods:** Patients with lung cancer have since 2014 completed an electronic version of the EORTC QLQ-C30/LC-13 questionnaires prior to every outpatient visit at our Department of Oncology. The patients' responses are used in routine clinical practice. Patients receiving palliative antineoplastic treatment were eligible for the analyses if the questionnaire had been completed before initiation of first-line treatment and at the time of disease progression. Longitudinal mean changes in PRO scores were compared using paired t-test. The predictive value of the scores was

evaluated using Cox proportional hazard model. A  $p$  value  $< .01$  was considered statistically significant. Results: A total of 94 patients met the inclusion criteria. At the time of disease progression, survival could be predicted by the absolute score of the global health scale, three functional scales (physical, role, emotional) and seven symptom scales (fatigue, pain, dyspnea, hemoptysis, lung cancer dyspnea, chest pain). In addition, changes in hemoptysis, dysphagia, dyspnea, and chest pain predicted survival at the time of progression. Conclusion: PRO measures used in clinical practice may provide clinicians with relevant predictive information about patients with lung cancer at the time of disease progression. These results show the potential value of PRO measures when used in clinical decision-making.

#### 104: Cancer research in pediatric and older populations

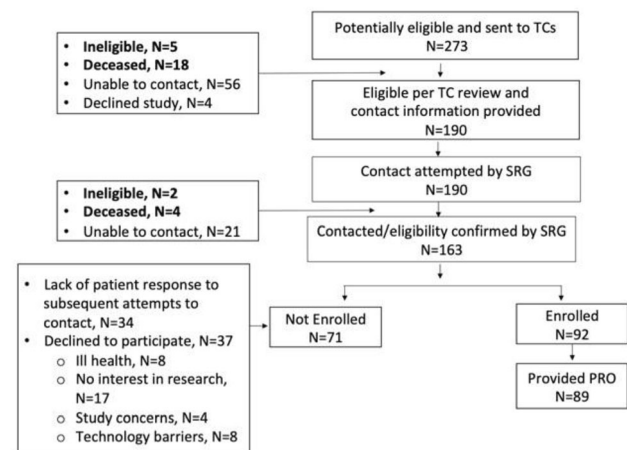
##### (104.1) Allogeneic hematopoietic cell transplantation (alloHCT) for patients over 65 years old is not associated with worse symptoms/function than younger patients - a Center for International Blood and Marrow Transplant Research (CIBMTR) study

Bronwen Shaw, CIBMTR, Milwaukee, Wisconsin, United States; Ruta Brazauskas, CIBMTR, Milwaukee, Wisconsin, United States; Linda J Burns, CIBMTR, Minneapolis, Minnesota, United States; Deborah Mattila, CIBMTR, Minneapolis, Minnesota, United States; Alisha Musseter, CIBMTR, Minneapolis, Minnesota, United States; Min Chen, CIBMTR, Minneapolis, Wisconsin, United States; Sumithra Vasu, Ohio State University, Columbus, Ohio, United States; Joseph Pidala, H. Lee Moffitt Cancer, Tampa, Florida, United States; Lori Muffly, Stanford University, Stanford, Connecticut, United States; Joseph Uberti, Karmanos Cancer Center, Detroit, Michigan, United States; Roni Tamari, Karmanos Cancer Center, Detroit, Michigan, United States; Roni Tamari, Memorial Sloan Kettering Cancer Center, New York, New York, United States; Erin Leckrone, CIBMTR, Minneapolis, Minnesota, United States; Judith Myers, CIBMTR, Milwaukee, Wisconsin, United States; Lih-Wen Mau, CIBMTR, Minneapolis, Minnesota, United States; J Douglas Rizzo, CIBMTR, Milwaukee, Wisconsin, United States; Wael Saber, CIBMTR, Milwaukee, Wisconsin, United States; Mary M Horowitz, CIBMTR, Milwaukee, Wisconsin, United States; Stephanie J Lee, Fred Hutch Cancer Research Center, Seattle, Washington, United States; Kathryn E Flynn, CIBMTR, Milwaukee, Wisconsin, United States

**Aims:** CIBMTR is an outcomes database that collects longitudinal clinical data on HCT recipients ( $> 540,000$  current participants), but patient-reported outcomes (PROs) are not routinely collected. The primary aim of this study was to determine whether symptoms/function are worse in older ( $\geq 65$  years) than younger (55–64 years) alloHCT recipients. The secondary aim was to test the feasibility of an electronic PRO (ePRO) system in CIBMTR registry patients. **Methods:** This was a cross-sectional study of patients  $\geq 55$  years old with primary/secondary myelodysplastic syndrome (MDS) undergoing an alloHCT under a Centers for Medicaid & Medicare Coverage with Evidence Development protocol. The primary endpoint of treatment related mortality was comparable in older and younger patients. Additional study inclusion criteria were:  $\geq 6$  months from alloHCT, English/Spanish, and an active email address. CIBMTR confirmed eligibility and obtained contact details for patients from six participating transplant centers. All further contact was by the CIBMTR Survey Research Group (SRG). Consent was electronic and PROMIS measures (Table 1) were delivered using computerized adaptive testing (CAT). Results: 244 of 273 patients eligibility confirmed, and 163 could be contacted; 92 patients enrolled and 89 provided PROs (Fig. 1). Participation for eligible patients did not differ by age. Median time post HCT for patients completing

PROs was 32 (range 9–94) months. The average number of questions per domain was 4.2–7.4, and the total time taken was 18.3 min (IQR). Older patients did not report worse function/symptoms in any domain, and both pain interference and sleep disturbance were significantly lower in the  $\geq 65$  vs  $< 65$  year group (Table 1). In multivariable analysis (Table 2), addressing both clinical and socio-economic predictors, there was a negative impact of needing a caregiver (both), lower income (pain), unemployment (sleep) and having secondary MDS (sleep). Conclusion: Older patients do not have worse symptoms/function than younger patients after an alloHCT for MDS, in fact younger patients reported worse pain interference and sleep disturbance. Patients of all ages reported worse physical functioning than population norms, consistent with past research. Inability to contact patients was the largest barrier to accrual and is being addressed by a number of strategies.

Figure 1: Consort diagram



#### bold, ineligible (n=29)

Table 1: Median (range) of scores of individual PROMIS domain by age.

Domain	Age <65 years (n=27)	Age $\geq 65$ years (n=62)	p-value
Social roles*	49.4 (21.99-66.39)	49.14 (27.05-67.37)	0.799
Anxiety*	48.24 (33.11-71.37)	48.12 (33.11-72.66)	0.574
Depression*	49.64 (34.28-71.94)	49.64 (34.28-61.92)	0.774
Emotional support*	52.13 (36.4-66.13)	55.51 (36.06-66.13)	0.162
Fatigue*	49.23 (32.45-78.32)	50.77 (25.33-68.49)	0.978
<b>Pain interference*</b>	<b>53.55 (38.75-68.97)</b>	<b>43.61 (38.76-66.98)</b>	<b>0.050</b>
Physical function*	45.46 (29.39-52.42)	43.42 (28.28-61.71)	0.862
<b>Sleep disturbance*</b>	<b>54.73 (38.26-68.74)</b>	<b>47.7 (26.51-72.06)</b>	<b>0.004</b>

\* higher score is better; # higher score is worse

p-value of  $\leq 0.05$  was considered significant as indicated in bold font

Table 2: Predictors of PRO outcomes for pain interference and sleep disturbance

Variable	Group	Pain Interference		Sleep Disturbance	
		Estimate (SE)	p-value	Estimate (SE)	p-value
Intercept		-56.50 (2.42)	$\leq 0.0001$	49.45 (2.73)	$\leq 0.0001$
Age*	<65 years	0.00	.	0.00	.
	$\geq 65$ years	-5.39 (2.06)	0.0106	-5.06 (2.03)	0.0149
Income*	<60,000 US dollars	0.00	0.0163 (overall)		
	$\geq 60,000$ US dollars	-6.40 (2.22)	0.0050		
	Missing	-2.84 (3.27)	0.3864		
Needing a caregiver*	No	0.00	.	0.00	.
	Yes	7.45 (3.14)	0.0198	6.88 (2.93)	0.0215
Employment*	Part/full time			0.00	0.0252 (overall)
	Unemployed/on disability			9.94 (3.66)	0.0081
	Retired			3.29 (2.76)	0.2368
Secondary MDS*	No			0.00	.
	Yes			-5.02 (2.29)	0.0314

\*at the time of PRO completion; #at transplant; SE, standard error; US, United States

#### **(104.2) Agreement between child self-report and caregiver proxy-report for symptoms and functioning of children undergoing cancer treatment**

Bryce Reeve, PhD, Duke University School of Medicine, Durham, North Carolina, United States; Jennifer Mack, MD, MPH, Dana-Farber Cancer Institute/Boston Children's Hospital, Boston, Massachusetts, United States; Molly McFatrach, MPH, Duke University School of Medicine, Durham, North Carolina, United States; Janice Withycombe, PhD, RN, MN, Clemson University School of Nursing, Clemson, South Carolina, United States; Scott Maurer, MD, University of Pittsburgh School of Medicine, UPMC Children's Hospital of Pittsburgh, Pittsburgh, Pennsylvania, United States; Shana Jacobs, MD, Children's National Hospital, Pittsburgh, District Of Columbia, United States; Li Lin, MS, Duke University School of Medicine, Durham, North Carolina, United States; Nicole Lucas, BS, Duke University School of Medicine, Durham, North Carolina, United States; Justin Baker, MD, FAAP, FAAHPM, St. Jude's Children's Research Hospital, Memphis, Tennessee, United States; Courtney Mann, MA, CCRP, Duke University School of Medicine, Durham, North Carolina, United States; Lillian Sung, MA, CCRP, Duke University School of Medicine, Durham, North Carolina, United States; Lillian Sung, MD, PhD, The Hospital for Sick Children, Toronto, Ontario, Canada; Deborah Tomlinson, RN, MN, The Hospital for Sick Children, Toronto, Ontario, Canada; Pamela Hinds, PhD, RN, Children's National Hospital, Washington, D.C., United States

**Aims:** Although patient-reported outcomes are increasingly used in adult cancer care, there has been more limited acceptance that children can self-report their own health-related quality of life (HRQOL). Caregiver report is therefore often used as a proxy for child's HRQOL. The study aim is to examine the association between child self-report and caregiver proxy-report for Patient-Reported Outcomes Measurement Information System® (PROMIS®) HRQOL domains among children with cancer, and to identify factors associated with better child and caregiver-proxy agreement. **Methods:** Children (7–18 years) with a first cancer diagnosis and their caregivers completed surveys at 2 time points, 72 h preceding treatment initiation (T1) and at follow-up (T2), when symptom burden was expected to be higher (e.g., 7–17 days later for chemotherapy). Data collection from nine pediatric oncology hospitals was from October 2016 to September 2018. PROMIS measures included mobility, pain interference, fatigue, depressive symptoms, anxiety, and psychological stress. Intraclass correlations (ICCs) evaluated agreement between child and parent-proxy, and multivariable mixed-effect models, adjusting for child and parent sociodemographic factors and the caregiver's own self-reported HRQOL, identified factors associated with better or worse agreement. **Results:** 482 child-caregiver dyads completed surveys at T1 (response rate 83.1%). 46% of children were female, 17% were black, and 15% Hispanic. ICCs between child self-report and caregiver proxy-report were moderate for mobility (ICC = 0.57) and poor for symptoms (ICCs = 0.32–0.42). In the multivariable model, caregivers reported the child's mobility score 6.00 points worse than the child's self-report at T2, exceeding the PROMIS minimally important difference of 3 points. Caregivers overestimated the child's self-reported symptom levels, ranging from 5.79 points (psychological stress) to 13.69 points (fatigue). The caregiver's own self-reported HRQOL was associated with greater discrepancy between child and caregiver scores for all domains except mobility. **Conclusion:** This study found that agreement between child self-report and caregiver proxy-report for symptoms was poor and for mobility was moderate. Caregivers consistently overestimated symptoms and underestimated mobility relative to children themselves; this discrepancy got larger when the caregivers' own HRQOL got worse. These results argue for elicitation of the child's own report whenever possible in pediatric oncology research and healthcare delivery settings.

#### **(104.3) Geriatric patient experience after thoracic surgery-how to help older patients expect the unexpected**

Gita Mody, MD MPH, University of North Carolina, Chapel Hill, North Carolina, United States; Antonia Bennett, PhD, University of North Carolina, Chapel Hill, North Carolina, United States; Mahrokh Irani, MPH, Harvard School of Public Health, Boston, Massachusetts, United States; Clara Kerwin, BS, University of Colorado School of Medicine, Denver, Colorado, United States; Michael Jaklitsch, MD, Brigham and Women's Hospital, Boston, Massachusetts, United States; Laura Frain, MD, Brigham and Women's Hospital, Boston, Massachusetts, United States

**Aims:** To evaluate the postoperative experience of elderly patients undergoing thoracic surgery in order to identify differences in expected and actual outcomes as well as explore additional themes for future research. **Methods:** A purposive sample of 10 patients over 70 years old having undergone thoracic surgery at least 1 year previously were identified from a combined clinic held by a thoracic surgeon and geriatrician in an academic hospital setting. Patients underwent comprehensive geriatric assessments preoperatively including a frailty score calculated from comorbidities, medications, and physical exam and categorized as robust, pre-frail, and frail. Semi-structured telephonic individual interviews were done using an interview guide designed to elicit perceptions on the postoperative recovery experience in the domains of physical and emotional health-related quality of life (HRQOL), postoperative symptoms, and recommendations for other patients. Interviews were recorded. Field notes were taken, and exemplar quotes were extracted during review of the recordings. Two study personnel independently performed inductive coding of the field notes and quotes. Descriptive statistics of patient demographics were calculated. **Results:** Ten interviews were obtained in 7 participants who underwent pulmonary resection for lung cancer and 3 who underwent esophageal procedures. Seven participants were women and the mean age was 77.7 years (SD ± 6.1). Frailty scores were robust (1), pre-frail (4), and frail (5). Dominant themes that emerged were the unexpected duration of physical recovery time (5), improvement in emotional HRQOL postoperatively (3), and eventual return to baseline or better physical function postoperatively (7). Three participants stated they would have preferred to have been made more aware preoperatively of the possibility of the complications they experienced. **Recommendations** for other patients were heterogeneous. **Conclusion:** Elderly thoracic surgery patients experience a gap between their expected and actual postoperative outcomes particularly the duration of physical recovery. Increased patient-surgeon communication and patient engagement may reduce this gap. Lessons from patient engagement models in other elderly patients, including using patient-reported outcomes (PROs) to monitor postoperative symptoms, physical functioning, and HRQOL may improve preoperative education as well as clinical management of these issues. Integration of PROs into perioperative care may serve as a novel area for future thoracic surgical research.

#### **(104.4) Multilevel social determinants of patient-reported outcomes among childhood cancer survivors: a report from the PEPR Consortium**

Jin-ah Sim, PhD, Department of Epidemiology & Cancer Control, St. Jude Children's Research Hospital, Memphis, Tennessee, United States; Justin N. Baker, MD, Department of Oncology, St. Jude Children's Research Hospital, Memphis, Tennessee, United States; Kevin R. Krull, PhD, Department of Epidemiology & Cancer Control, St. Jude Children's Research Hospital, Memphis, Tennessee, United States; Kirsten K. Ness, PhD, Department of Epidemiology & Cancer Control, St. Jude Children's Research Hospital, Memphis, Tennessee,

United States; Deo Kumar Srivastava, PhD, Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, Tennessee, United States; Christopher B. Forrest, MD, PhD, Department of Pediatrics, Children's Hospital of Philadelphia, Memphis, Pennsylvania, United States; Melissa M. Hudson, MD, Department of Oncology, St. Jude Children's Research Hospital, Memphis, Tennessee, United States; Leslie L. Robison, PhD, Department of Epidemiology & Cancer Control, St. Jude Children's Research Hospital, Memphis, Tennessee, United States; I-Chan Huang, PhD, Department of Epidemiology & Cancer Control, St. Jude Children's Research Hospital, Memphis, Tennessee, United States

**Aims:** The impact of demographic and treatment factors on patient-reported outcomes (PROs) in pediatric cancer populations is well established; however, the influence of contextual/social factors on PROs is understudied. We aimed to investigate the associations of contextual/social factors at parental, family, and community levels with PROs among childhood cancer survivors. **Methods:** Study participants were 293 childhood cancer survivors who took part in the PEdiatric Patient-Reported Outcomes in Chronic Diseases Consortium (PEPR). Inclusion criteria were survivors of pediatric malignancies who were 8–18.9 years of age at the time of study. Eight contextual/social factors were chosen and classified into 3 levels: parental/family (parental loneliness perceptions, household income, family conflict, etc.), census track (distance to major roads, etc.), and county environment (socioeconomic status, healthcare resources, physical infrastructure, etc.). Parental/family factors were self-reported by primary caregivers. Census track and county factors were created by geocoding the participant's home addresses and linked to national databases (US Census Bureau, CDC, etc.). Survivors' PROs were self-reported using PROMIS measures (fatigue, pain intensity, sleep disturbance, mobility, positive affect domains). Seemingly unrelated regression (SUR) was used to test associations of contextual/social factors with PROs adjusting for survivors' age, sex, cancer diagnoses. **Results:** Mean age of survivors was 14.2 years ( $SD = 2.9$ ); 50.7% were male; major diagnoses included solid tumors (48.0%), leukemia (31.8%), brain tumors (13.6%), and lymphoma (4.3%). Poorer fatigue and mobility were significantly associated with cancer diagnosis, especially brain tumors vs. leukemia survivors (fatigue  $\beta = 0.19$ ,  $p = 0.01$ ; mobility  $\beta = -0.23$ ,  $p = 0.003$ ), but not contextual/social factors ( $p$ 's  $> 0.05$ ). In contrast, contextual/social factors rather than diagnosis were significantly associated with poor PROs on other domains. At the parental/family level, lower family income was associated with higher pain intensity ( $\beta = 0.18$ ,  $p = 0.01$ ); higher parental loneliness or family conflict were associated with more sleep disturbance ( $\beta = 0.17$ ,  $p = 0.04$ ;  $\beta = 0.17$ ,  $p = 0.02$ ). At the community level, lower area socioeconomic status was associated with more pain intensity and sleep disturbance ( $\beta = 0.16$ ,  $p = 0.02$ ;  $\beta = 0.20$ ,  $p = 0.004$ ). **Conclusion:** Contextual/social (parental, family, and community factors) contributed to poor self-reported health among childhood cancer survivors, which were independent from demographic and clinical factors. Future studies are warranted to design interventions to improve PROs by addressing contextual/social challenges.

## 105: Efficiency and quality in the development and application of PROs

### (105.1) Translatability assessment and linguistic validation of the ABOUT–Dependence instrument into German, Italian, Japanese and Russian

Catherine Acquadro, ICON plc, Language Services, Lyon, France; Jennifer Lambe, ICON plc, Language Services, Lyon, France;

Adeline Verne, ICON plc, Language Services, Lyon, France; Agnes Bacso, PMI R&D, Philip Morris Products S.A, Neuchâtel, Switzerland; Esther Afolalu, PMI R&D, Philip Morris Products S.A, Neuchâtel, Switzerland; Linda Abetz-Webb, Patient-Centered Outcomes Assessments Ltd, Neuchâtel, United Kingdom; Christelle Chrea, PMI R&D, Philip Morris Products S.A., Neuchâtel, Switzerland

**Aims:** The ABOUT™–Dependence instrument was developed in US English to address the need for a fit-for-purpose instrument for assessing perceived dependence associated with the use of different tobacco and nicotine-containing products (TNP). The instrument is composed of 12 items covering three domains: extent of use (two items), signs and symptoms (five), and behavioral impact (five). The objective of this research was to assess the applicability of the source version of the ABOUT™–Dependence to other languages (German, Italian, Japanese, and Russian). **Methods:** A translatability assessment (TA) was performed on a draft version of the instrument as part of its cross-cultural development. The final version went through a linguistic validation (LV) process consisting of five steps: conceptual analysis, forward and back translation into English, testing through cognitive interviews ( $n = 6$  for German, Italian, and Japanese;  $n = 5$  for Russian), external review, and proofreading. Translation issues found by LV were categorized as cultural, idiomatic, semantic, or syntactic. **Results:** The TA identified a few potential cultural and linguistic issues and recommended solutions for optimizing the source instrument for future translations. The LV process raised a total of 34 different concerns, in which 5 items contained 2 or more issues related to semantic (7), syntactic (2), or idiomatic (5) aspects. The majority of issues were highlighted in the translation step, allowing for revision of the original. An unclear concept definition in the “use your product(s) in a situation where you weren't supposed to” English item resulted in a conceptually non-equivalent Russian translation, potentially implying TNP use in situations or places where it is banned or forbidden. The Russian translation was modified to match the original concept and the concept definition of this item was revisited to clarify the original English meaning. **Conclusion:** Combination of the TA and LV processes led to translation of the ABOUT™–Dependence instrument, adequately capturing the concepts of the original version and being reliably applicable to Germany, Italy, Japan, and Russia. These translations provide opportunities for comparing perceived dependence across multiple products and types of users globally. The findings also emphasize the importance of clear a priori conceptual definition of items.

### (105.2) Using patient-reported measures to assess quality of healthcare from the patient's perspective—The German approach

Konstanze Blatt, Dr., IQTIG, Berlin, Germany

**Aims:** There is a tradition of regular and obligatory assessments of healthcare quality in Germany. The Institute of Quality Assurance and Transparency in Healthcare (IQTIG) is commissioned by the Federal Joint Committee (G-BA) to develop disease-specific quality indicators based on patient surveys. On the example of two different disease groups - patients with schizophrenia and patients with percutaneous cardiac intervention—the general method established to develop quality indicators to be assessed by patient-reported measures in Germany will be illustrated. **Methods:** To address concrete aspects of care, quality indicators should refer to facts and situations that can be reported by patients. Thus, instead of measures of patient satisfaction, patient-reported outcome measures (PROMs) and patient-reported experience measures (PREMs) were set as general frame. The development of disease-specific instruments was divided in five



phases: 1. systematic literature search for quality-related aspects of care, 2. focus groups/interviews with patients and medical professionals to gather their criteria of high quality, 3. expert group meetings to assess the so far defined criteria, 4. development of questionnaires including two-phase-pretesting (cognitive and conventional pretesting), 5. definition of quality indicators. Results: Patient-reported measures were developed for patients with schizophrenia and patients with percutaneous cardiac intervention. For that, focus groups and interviews were conducted with 86 patients and 64 professionals. Especially process-related aspects turned out to be of high relevance, e.g., information/explanation about treatment, medication, patient-health professional-interaction, which were addressed by PREMs and had disease-specific facets. There are also outcomes crucial to the quality of the specific care reflected by PROMs. The questionnaires underwent cognitive pretesting with 114 patients. Thirty-three quality indicators were defined on the base of the conventional pretesting with 1570 patients. The pretesting hints to stable instruments addressing disease-specific quality indicators which are important to patients. Conclusion: The IQTIG applies a complex method for the development of disease-specific patient-reported measures to assess quality of care. The direct participation of patients' and their view as well as the professionals' experiences and research literature produces instruments which strengthen patient-centered care and give patients' a voice within the definition of requirements and thus determination of healthcare.

### (105.3) Validation of a positive affect item bank—a supplement to the EORTC CAT Core

Morten Petersen, Palliative Care Research Unit, Department of Geriatrics and Palliative Medicine GP, Bispebjerg og Frederiksberg Hospital, Copenhagen NV, Denmark; Eva-Maria Gamper, Department of Psychiatry, Psychotherapy and Psychosomatics, University Hospital of Psychiatry II, Innsbruck, Austria; Anna Costantini, Psychoncology Unit, Sant'Andrea Hospital, Faculty of Medicine and Psychology Sapienza University, Rome, Italy; Johannes M. Giesinger, Department of Psychiatry, Psychotherapy and Psychosomatics, University Hospital of Psychiatry II, Innsbruck, Austria; Bernhard Holzner, Department of Psychiatry, Psychotherapy and Psychosomatics, University Hospital of Psychiatry II, Innsbruck, Austria; Colin Johnson, Surgical Unit, Faculty of Medicine, University of Southampton, Innsbruck, United Kingdom; Monika Sztankay, Department of Psychiatry, Psychotherapy and Psychosomatics, University Hospital of Psychiatry II, Innsbruck, Austria; Teresa Young, East & North Hertfordshire NHS Trust incorporating Mount Vernon Cancer Centre, Northwood, Middlesex, United Kingdom; Mogens Groenvold, Palliative Care Research Unit, Department of Geriatrics and Palliative Medicine GP, Bispebjerg og Frederiksberg Hospital, Copenhagen, Denmark; submitted on behalf of the EORTC Quality of Life Group

**Aims:** The EORTC CAT Core developed by the European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Group (QLG) is a computer-adaptive test (CAT) instrument providing assessment of the 14 symptom and functional domains of the widely used EORTC QLQ-C30 questionnaire. During the development of the emotional functioning item bank, five positive affect items were developed. However, they were discarded after extensive psychometric analyses, as they did not fit together with the other items. The aim here was to evaluate whether the positive affect items may form a separate, unidimensional item bank. **Methods:** We followed the EORTC QLG's general approach for psychometric validation of CAT item banks.

Based on a large international sample of cancer patients, this includes evaluations of dimensionality, item response theory (IRT) (generalized partial credit) model fit, differential item functioning (DIF), measurement precision and known groups validity. Results: A total of 995 cancer patients from Austria, Denmark, Italy, and the UK responded to the positive affect items. Evaluations indicated acceptable fit of a unidimensional model: Cronbach's alpha = 0.87, one factor explained 72% of variation, fit indices CFI = 0.99, RMSEA = 0.10. Inspection of item residuals and infit and outfit indices indicated minor redundancy/local dependence for some items. Possible inflation of slope parameters because of dependencies was accounted for during IRT model calibration. Of 45 evaluations of DIF, three indicated significant DIF. However, evaluations indicated the possible impact of these to be trivial. High reliability (> 0.80) was observed for about 3½ standard deviations of the measurement continuum. The item set clearly distinguished between groups expected to differ (median effect size = 0.6 across known groups comparisons). Conclusion: The items constitute a unidimensional IRT calibrated item bank which can be applied across patient groups (no DIF) and may be administered as CAT or as short form. It is ready for further validation in new data. If this confirms the validity, the item bank may supplement the 14 item banks of the EORTC CAT Core, thereby expanding the domain coverage of the EORTC CAT item banks to include positive emotional affect.

### (105.4) Increasing the efficiency of computerized adaptive tests even further: The case of empirical priors

Muirne Paap, University of Groningen, Groningen, Netherlands; Niek Frans, Oslo University Hospital, Oslo, Norway; Johan Braeken, CEMO, University of Oslo, Oslo, Norway

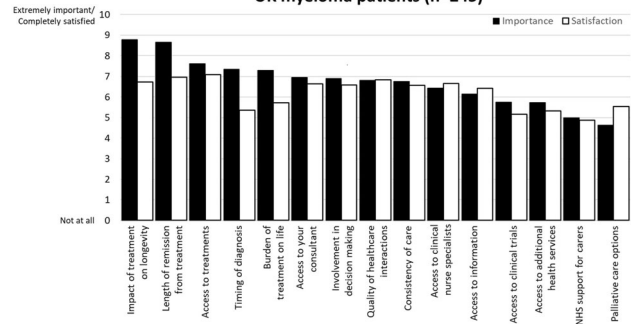
**Aims:** Computerized adaptive tests (CATs) are instruments that are adapted to the patient on the fly, and are increasingly used in health measurement. Several studies conducted in the field of educational testing have shown that empirical information about a test-taker can be used to make a CAT more efficient and more precise. However, it is currently not sufficiently clear to what degree these findings can be generalized to health measurement. The aim of this simulation study is to evaluate the risks and return associated with using prior information in a healthcare setting. **Methods:** An empirical item bank calibrated with the graded response model was used to simulate unidimensional CATs. The bank was based on an instrument measuring personality dysfunction, which consisted of 12 polytomous items. The prior was based on a global score set by a clinician prior to administering the instrument, which was scored on an ordinal scale (0–4). The correlation between the prior and the estimated trait scores (personality dysfunction) equaled .76. Five-hundred simulees were generated, and two CATs were administered to them: one with a standard normal prior (default) and one with an empirical prior. For each global score, the theta value with the highest probability was taken as the mean for the empirical prior. The lower this highest probability, the larger the variance of the empirical prior: N(-3, 0.685) if global score is 0 N(-1, 0.625) if global score is 1 N(0, 0.612) if global score is 2 N(1, 0.812) if global score is 3 N(2, 0.500) if global score is 4 **Results:** For 63.6% of the simulees, using the empirical prior resulted in a reduction in test length; for the remaining 31.7% test length was equal across the two conditions. On average, the empirical prior was associated with a reduction in test length of 20%. Bias in the final estimates was comparable across the two conditions, and was generally low (well below 1.5). More detailed results will be presented at the conference. **Conclusion:** Based on these findings, using empirical prior information can be expected to be advantageous in healthcare settings where polytomous items are used.

### (105.5) What matters most to different stakeholder groups? Using a novel anchored Best–Worst Scaling Health Pathway Index to explore alignment across UK patients, carers and clinicians in myeloma

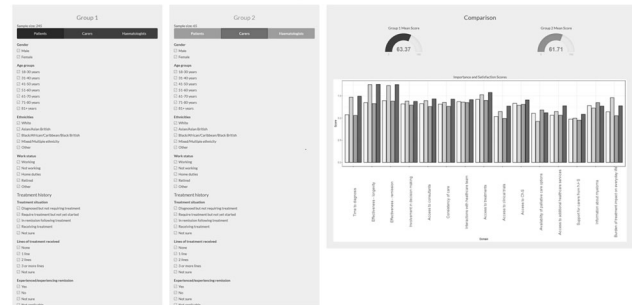
Simon Fifer, B.Ec., MAppStat, PhD (Choice Modelling), Community and Patient Preference Research (CaPPRe), Sydney, Australia; Bronwyn West, BA, Community and Patient Preference Research (CaPPRe), Sydney, Australia; Liisa Wilton, PhD, BSc (hons), Janssen-Cilag Ltd, High Wycombe, United Kingdom; Cicely Kerr, PhD, BSc (hons), Janssen-Cilag Ltd, High Wycombe, United Kingdom; Jayne Galinsky, PhD, Myeloma UK, Edinburgh, United Kingdom; Adele Goman, PhD, Myeloma UK, Edinburgh, United Kingdom; Sarah McDonald, MSc, Myeloma UK, Edinburgh, United Kingdom

**Aims:** This research explored UK myeloma stakeholder preferences and satisfaction across the Health Pathway (HP). It investigated alignment between patients, carers and clinicians on what matters most across the myeloma HP, to inform routine outcome-measurement for shifting to value-based healthcare. **Methods:** Best–Worst Scaling (BWS) is a survey technique that takes advantage of people's ability to reliably identify extremes ('best' and 'worst') in sets of items, eliciting discriminating rankings free of scale bias. This study implemented a novel anchoring process to rescale importance and satisfaction BWS scores for factors across the myeloma HP which could be compared and combined into an HP Index (HPI). This method has been used in Australia to measure personal well-being for the Government. In-depth interviews and focus groups were conducted in the UK with myeloma stakeholders (10 patients, 7 carers, 9 hematologists, 3 payers). **Results:** were discussed with a steering committee including a hematologist, cancer research nurse and patient, to identify 15 factors across the HP. 350 UK myeloma stakeholders (245 patients, 65 carers, 40 hematologists) completed the HPI BWS survey to rank the 15 factors on two dimensions: importance and satisfaction. The HPI, generated from the anchored combined scores, ranged from 0 to 100. **Results:** Two treatment factors, Impact of treatment on longevity and Length of remission from treatment, were most important for patients. The largest gaps between satisfaction and importance were found in four of the top five factors that patients' ranked most important (Fig. 1). Mean HPI scores were higher for myeloma patients than for carers and hematologists (63.4, 61.7 and 59.7 respectively) suggesting relatively higher patient satisfaction with important factors than the other stakeholder groups. HPI data implemented in an interactive dashboard facilitate further data interrogation and comparison across stakeholders (Fig. 2). **Conclusion:** Gaps between satisfaction and importance identified by this novel HPI approach highlight potential areas for impactful improvement of the myeloma HP (factors with high stakeholder importance but low satisfaction). It also puts key stakeholders' preferences and satisfaction at the center to inform funding and policy decisions. Decision-makers could use this robust tool to also capture improvement over time.

**Figure 1: Rescaled anchored satisfaction and importance scores for UK myeloma patients (n=245)**



**Figure 2: Interactive Dashboard Allowing Subgroup & Stakeholder Group Comparisons**



### (105.6) Evolution of multimorbidity patterns and their association with quality of life: findings from SHARE

Tatjana Makovski, Luxembourg Institute of Health, Luxembourg, Luxembourg; Michael Schnell, Luxembourg Institute of Health, Luxembourg, Luxembourg; Laetitia Huiart, Luxembourg Institute of Health, Luxembourg, Luxembourg; Marjan van den Akker, Institute of General Practice, Goethe University, Luxembourg, Luxembourg

**Aims:** Living with multimorbidity (here defined as 2 + diseases) is typically accompanied with lower quality of life (QoL). Understanding how diseases group, along with how they impact patients QoL would allow improvement of care. The aim of the study is to: 1) identify the existing multimorbidity patterns and their evolution over time, and 2) clarify their association with QoL. **Methods:** Longitudinal study on Survey of Health, Ageing and Retirement in Europe (SHARE), using data collected in 2006/07, 2011/12, 2013 and 2015 in 10 countries. Only individuals aged 50 + who participated in all 4 waves were considered ( $n = 10,257$ ). QoL was assessed by the Control, Autonomy, Self-Realization and Pleasure (CASP-12v1) questionnaire. Exploratory factor analysis based on tetra-choric correlations, using 15 conditions, consistent across waves was applied to identify multimorbidity patterns in the last

wave. Evolution of disease combinations within patterns and associated QoL will be observed retrospectively. Here, we estimate QoL decline for each pattern in the last wave. For this purpose having at least 2 diseases from a pattern qualified as belonging to that pattern. Mixed-effects linear regression model was applied to evaluate the association between patterns and QoL adjusting for socio-economic factors, number of symptoms, difficulties with activities and instrumental activities of daily living, and number of contacts with a doctor. Results: Over half of the population (57%) had at least 2 diseases in the last wave. Preliminary findings indicated existence of 2 patterns: cardio-metabolic and mixed. Cardio-metabolic included heart attack, hypertension, high cholesterol and diabetes. Mixed pattern combined stroke, chronic lung diseases, stomach problems, cataracts, hip and other fractures, Alzheimer disease, arthritis and depression. Kaiser–Meyer–Olkin test confirmed adequacy of the sample [0.68]. Mixed pattern showed steeper decline of a QoL compared to cardio-metabolic [-5.5 (95%CI: -6.3; -4.6)] vs [-1.0 (95%CI: -2.0; -0.1)]. Disease combinations within both patterns evolved substantially across waves. Conclusion: Further investigation will provide detailed description of the evolution of patterns and their associated change in QoL. Understanding this dynamic would assist in planning more effective preventive and curative measures which could enable better health outcomes and best QoL for patients living with multimorbidity.

## 106: Rehabilitation Research

### (106.1) PROMIS short forms for pain and function show good measurement properties in total knee arthroplasty patients

Anika Stephan, MA, Department of Teaching, Research and Development, Schulthess Clinic., Zurich, Switzerland; Vincent Stadelmann, PhD, Department of Teaching, Research and Development, Schulthess Clinic, Zürich, Switzerland, Zurich, Switzerland; Stefan Preiss, M.D., Department of Hip and Knee Surgery, Schulthess Clinic, Zurich, Switzerland; Michael Leunig, Prof., Department of Hip and Knee Surgery, Schulthess Clinic, Zurich, Switzerland; Franco Impellizzeri, Prof., Faculty of Health, University of Technology Sydney, Sydney, Australia

**Aims:** The Patient-Reported Outcomes Measurement Information System (PROMIS) aims to provide a common metric of health for many medical conditions. While PROMIS is mainly designed for computer-adaptive testing, its static short forms (SF) are used when a paper-pencil format is preferred. We examined the measurement properties of the German PROMIS-SF for pain intensity (PAIN), pain interference (PI) and physical function (PF) in patients undergoing total knee arthroplasty (TKA). **Methods:** PROMIS-SF were collected from TKA patients pre-, 6 and 12 months post-surgery. Higher scores indicate more PAIN, higher PI and better PF. Oxford Knee Score (OKS) was the main reference measure. At follow-up, patients rated their global treatment outcome (GTO) and symptom-specific well-being (SSWB) on five-point-Likert scales. A subsample completed their baseline or 6-months PROMIS-SF forms twice within 14 days, to test reliability. Measurement properties were assessed according to the COSMIN guidelines. **Results:** From 214 eligible patients, 164 (77%) could be included in the study and received questionnaires. 144 (67%, 57 males, 87 females,  $68.4 \pm 8.9$  years) returned a baseline questionnaire, and 120 (56%), a 12-month questionnaire. 51 patients provided test-retest data. Correlations ( $r$ ) with OKS were as follows: PAIN,  $-0.7$ ; PI,  $-0.8$ ; PF,  $0.8$ . Correlations with SSWB were  $|r| \geq 0.6$ . Cronbach's  $\alpha$  were: as follows PAIN,  $0.84$ ; PI,  $0.90$ ; PF,  $0.88$ . Intraclass correlation coefficients were as follows: PAIN,  $0.92$ ; PI,  $0.90$ ; PF,  $0.97$ . Standard Errors of Measurement were as follows: PAIN,  $3.1$ ; PI,  $3.3$ ; PF,  $1.7$ . These represent 19%, 21% and 16% of the

mean score change, respectively. Smallest detectable change thresholds (SDC90) were as follows: PAIN,  $7.2$ ; PI,  $7.8$ ; PF,  $4.0$ . Minimal important changes were as follows: PAIN,  $9.4$ ; PI,  $7.7$ ; PF,  $8.4$ . Follow-up at 12 months showed ceiling effects (best score) for all three scales: PAIN, 42%; PI, 53%; PF, 30%. Correlations ( $r$ ) of PROMIS change scores with OKS change scores were as follows: PAIN,  $-0.65$ ; PI,  $-0.6$ ; PF,  $0.5$  and with GTO ratings were [ $0.4 \leq |r| \leq 0.5$ ]. OKS change scores correlated moderately with GTO ( $r = -0.45$ ). GTO showed a ceiling effect (71%). Conclusion: PROMIS-SFs of pain and function could be used in TKA patients. Our results confirmed the construct validity, reliability, and responsiveness. Measurement precision is sufficient to detect minimal important changes.

### (106.2) Patient-reported experience with patient-reported outcomes in adult patients seen in rheumatology clinics

Brittany Lapin, PhD MPH, Cleveland Clinic, Cleveland, Ohio, United States; Ryan Honomichl, PhD, Cleveland Clinic, Cleveland, Ohio, United States; Nicolas Thompson, MS, Cleveland Clinic, Cleveland, Ohio, United States; Andrew Schuster, BS, Cleveland Clinic, Cleveland, Ohio, United States; Abby Abelson, MD, Cleveland Clinic, Cleveland, Ohio, United States; Chad Deal, MD, Cleveland Clinic, Cleveland, Ohio, United States; Susannah Rose, PhD, Cleveland Clinic, Cleveland, Ohio, United States; Irene Katzan, MD, Cleveland Clinic, Cleveland, Ohio, United States

**Aims:** Patient-reported outcome measures (PROMs) are increasingly utilized in the evaluation of patients with rheumatic diseases, where self-perception of the disease and its treatment is critically important. In order to maximize the value of collecting PROMs, it is important to understand the patient experience with them. The aim of our study was to assess the patient experience with completing PROMs within rheumatology clinics, and to identify patient characteristics associated with a more positive experience. **Methods:** We conducted a retrospective cross-sectional study of adult patients seen in rheumatology clinics at Cleveland Clinic between 1/1/2017 and 6/30/2017. As part of standard of care, patients completed the RAPID3, PHQ-9 depression screen, 3 PROMIS domain scales and PROMIS Global Health. Patients were included in the study if they completed at least one patient-reported experience question following completion of PROMs within the study window. Patient characteristics associated with a more positive experience were identified through multivariable proportional odds models. **Results:** 12,597 adult patients (mean age  $59 \pm 15$ ; 76% female; 84% white) completed PROMs, as well as questions on their experience with completing PROMs. The majority of patients agreed/strongly agreed that PROM questions were easy to understand (97%), useful (84%), helped their physician understand their health (78%), improved communication with their provider (78%) and improved control over their own care (70%). After adjustment for other factors, being younger, non-white, having lower household income, fewer comorbidities, and being a new patient were independent predictors of better experience with PROMs. Moderate to severe depressive symptoms and worse physical function, pain interference, fatigue, and global health also predicted a better experience with PROMs. **Conclusion:** Our study found a positive patient experience with PROMs, which is a crucial component of their successful implementation and utilization. Findings from this study suggest PROMs may be particularly beneficial in new patients, minorities, those with lower incomes, and worse self-reported quality of life. Collecting PROMs as part of standard of care could provide opportunities to improve patient-provider communication and enhance control over care for rheumatology patients who could most benefit.

### **(106.3) Elderly patients' perspectives on treatment, care and rehabilitation after hip fracture: A qualitative systematic review**

Charlotte Abrahamsen, Department of Orthopedic Surgery, Kolding Hospital a part of Lillebaelt Hospital, Kolding, Denmark; Birgitte Nørgaard, Ass. Professor, Department of Public Health, University of Southern Denmark, Odense, Denmark

**Aims:** Hip fracture surgery is a distressing and life-changing event for patients. Treatment, care and rehabilitation of hip fracture patients is governed by evidence-based recommendations; patients' preferences are sparsely represented. In order to develop a more patient-centered healthcare system involving patients' individual assessments and preferences, PRO data can be used advantageously. However, only multiple generic and domain-specific PRO tools are used for hip fracture patients. The aim of this systematic review was to identify what elderly hip fracture patients consider important in relation to their fracture. The results should contribute to developing a patient-derived and hip-fracture-specific PRO tool. **Methods:** We conducted a systematic review and searched the following electronic bibliographic databases: PubMed, CINAHL, PsycINFO and Embase for qualitative studies. We included studies of patient with hip fracture aged 65 years or older reporting on patients' perspectives. The protocol was registered with PROSPERO (ID CRD42018091981). Both authors independently screened and identified studies meeting the inclusion criteria. The quality of all included studies was evaluated using the Critical Appraisal Skills Programme (CASP) checklist. Data were extracted and analyzed by both authors using content analysis and categorized by similarity in meaning as either health-related factors or healthcare-related experiences. **Results:** Seventeen qualitative studies met the inclusion criteria. With CASP quality scores of 6.5–9.5, the quality of the studies varied. The health-related factors identified included: (1) symptoms and complications, (2) physical health, (3) mental health and (4) social relationships and (5) personal goals. Healthcare-related experiences revolved around: (1) waiting time, (2) information, (3) participation and respect and (4) discharge. A total of 162 findings important to hip fracture patients were identified. **Conclusion:** Regaining physical functioning, mobility and independence is considered most important by elderly hip-fracture patients. Above all, they want to return to their preferred activities in everyday life at home. Their social network, a surplus of mental resources and the reduction of pain and complications are vital. All of them, factors relevant to incorporate in a patient-derived and hip-fracture-specific PRO tool for the purpose of planning and delivering care based on what matters to patients.

### **(106.4) Burden of care and quality of life among informal caregivers of individuals with spinal cord Injury in Ibadan, Nigeria**

Rapheal Omosihin, University of Ibadan, Ibadan, Nigeria; Joshua Ojo, MSc Physiotherapy, University of Benin Teaching Hospital, Benin, Nigeria; Oluwaseun Erinle, University of Ibadan, Ibadan, Nigeria; Nse Odunaiya, PhD, University of Ibadan, Ibadan, Nigeria

**Aims:** Informal caregivers of individuals with spinal cord injury may experience burden of care that impact negatively on their quality of life. Understanding what constitute burden to informal caregivers and factors associated with it are pivotal in planning programs to take care of informal caregivers. **Methods:** The study utilized a mixed-method design of cross-sectional survey and qualitative component. Cross-sectional aspect of the study involved 34 consenting informal caregivers of patients with spinal cord injury from the physiotherapy outpatient clinic and neurosurgery ward of University College

hospital, Ibadan. The Zarit Burden Interview (ZBI) questionnaires and 36-short form health survey questionnaire were used to assess the level of burden of care and the quality of life of informal caregivers. Seven consenting informal caregivers participated in the qualitative component of this study. Descriptive statistics and non-parametric techniques of Cramer's V test, Mann–Whitney test and spearman correlation were used to analyze the quantitative study. The level of significance was set at 0.05. The qualitative data were analyzed using content thematic analysis. **Results:** The mean age of participants in this study was  $41.26 \pm 11.39$  years. The cross-sectional study showed that 28 (82.4%) participants had a high level of burden of care. There was a significant association between burden of care and level of income of participants, and burden of care and number of hours spent caring for relative. No significant association was found between burden of care and all the domains of quality of life except on vitality a mental health component. The qualitative study provided further insight. Specific factors that constitute a burden are hospital administration logistics, financial difficulties, and negative attitude of health workers. Participants also mentioned how burden of care affect their social, psychological and sexual functioning. **Conclusion:** Majority of informal caregivers experiences a high level of burden, associated with poor hospital administration, health workers attitude and high cost of care. Costs of care constitute a burden, as most of the people pay out of pocket. Context specific solutions such as strengthening of health system and better insurance coverage among others are important to reduce burden of care.

### **(106.5) Identifying a core set of patient-centered outcomes for spinal cord injury care**

Sara Ahmed, PhD, McGill University, Montreal, Quebec, Canada; Diana Zidarov, McGill University, Montreal, Quebec, Canada; W. Ben Mortenson, University of British Columbia, Vancouver, British Columbia, Canada; Dalton Wolfe, Parkwood Institute Research Lawson Health Research Institute, London, Ontario, Canada; Vanessa Noonan, Rick Hansen Institute, Vancouver, British Columbia, Canada; Richard Riopelle, Ontario Neurotrauma Foundation, Vancouver, Ontario, Canada; David Tulskey, University of Delaware, Newark, New Jersey, United States; Susan Bartlett, McGill university, Montreal, Quebec, Canada

**Aims:** To identify a minimal battery of core patient-reported outcomes (PROs) for community spinal cord care (SCI-CORE) that can inform shared-decision-making and patient-centered treatment using a participatory stakeholder-driven process. Our goals were to: (1) bring patients, families, and other stakeholders together to prioritize a core set of PRO domains that are valued by patients, caregivers, and SCI clinicians; (2) identify user preferences for a real-time electronic data capture, scoring, and reporting (EDCR) system; and (3) understand potential barriers and facilitators and strategies for successful implementation at each site. **Methods:** We invited patients, families, clinicians, patient advocates, and decision-makers in three Canadian provinces to serve on an advisory committee. Together we developed an online Delphi exercise, now underway with 200 SCI stakeholders throughout North America. Delphi participants rate the importance of candidate domains. Domains endorsed by > 70% of participants will be brought forward into the next round until consensus is reached and mapped back to the SCI-QOL measurement system. Open-ended questions will explore users' preferences for SCI-CORE EDCR features. Barriers and facilitators will be identified through a deductive approach based on the Consolidated Framework for Implementation Research. Data will be reviewed by the stakeholder advisory committee and compared between sites and stakeholder groups. **Results:** We will present the consensus results of prioritized SCI QOL domains to be included in the SCI CORE EDCR. Preliminary considerations

including potential barriers and facilitators, IT support, language and cultural adaptations used to inform implementation planning will also be presented. Conclusion: This study provides a stakeholder-driven system for developing and implementing a core set of e-PROs to inform SCI care in real-world settings. Our next steps are to assess acceptability, and feasibility and fidelity of the SCI-CORE EDCR into routine care, and its impact on patient activation, shared decision-making, and the care experience. We will also evaluate how SCI PROs can empower individuals to self-manage and establish goals in collaboration with their rehabilitation team to improve long-term outcomes and QOL.

## 107: Advancing quantitative methods in PRO data analysis and interpretation

### (107.1) Estimating treatment effect on patient-reported outcomes subject to dropout: comparing traditional, contemporary and causal inference approaches

Andrew Trigg, MSc, Adelphi Values, Bollington, United Kingdom; Jessica Roydhouse, PhD, University of Tasmania, Hobart, Australia

Aims: Patient-reported outcomes data are often missing due to dropout, the likelihood of which can depend on how patients feel or function during the trial. This analysis aimed to compare estimates of treatment effect through several methods. Methods: Data were from two randomized controlled trials in patients with major depressive disorder with differing extent of dropout, focusing on the difference between duloxetine and placebo in mean change from Baseline to Week 8 in depressive symptoms as measured by HAM-D score. This was estimated through a complete-case ANCOVA (CCA), mixed-model for repeated measures (MMRM), a Bayesian selection model (SM), and a multiple-imputation-based pattern-mixture model (PMM-MI). In addition, principal stratification without (PS) and with baseline adjustment (PSA) was used to estimate the effect of treatment among patients who would complete PROs regardless of treatment assignment. Results: Dropout by Week 8 was 35.0% and 8.0% in each trial. Estimates of treatment effect by each method, within the high and low dropout trials, are provided in Table 1. Estimates were closer to CCA in the low dropout trial than the high dropout trial, as expected. While the MRMM and PMM-MI estimates were similar, the SM parameter linking current PRO score and dropout seemed unstable. PSA estimates were closest to CCA, likely due to limited baseline covariates available to predict strata membership. Conclusion: There was no consistent pattern in the magnitude of treatment difference shown by each method. Therefore, the full range of methods to handle dropout should be assessed in comprehensive sensitivity analyses, especially if employing the SM. The possibility of missing data and the necessity of advanced methods should be considered when collecting baseline covariates. Comparing the methods in a larger number of trials, including intermittent missing data, is warranted.

Table 1. Point estimates by method and trial

Trial	CCA	MMRM	SM	PMM-MI	PS	PSA
High dropout (35%)	-1.72	-2.46	-3.14	-2.45	-1.29	-1.73
Low dropout (8%)	-1.74	-2.01	-1.63	-1.93	-1.90	-1.71

### (107.2) Utility weights for the EORTC QLU-C10D for Spain and investigation of a potential “Corona effect”

Georg Kemmler, PhD, Medical University of Innsbruck, Innsbruck, Austria; Jens Lehmann, MSc, Department of Psychiatry, Psychotherapy and Psychosomatics, Medical University of Innsbruck, Innsbruck, Austria; Ludwig von Butler, MSc, SurveyEngine GmbH, Berlin, Germany; Shaad Ansari, MSc, SurveyEngine GmbH, Berlin, Germany; Richard Norman, PhD, School of Public Health, Curtin University, Perth, Australia; Madeleine King, Prof., University of Sydney, Faculty of Science, School of Psychology, Perth, Australia; Bernhard Holzner, Prof., Department of Psychiatry, Psychotherapy and Psychosomatics, Medical University of Innsbruck, Innsbruck, Austria; submitted on behalf of EORTC

Aims: The EORTC QLU-C10D is a cancer-specific utility instrument based on the EORTC core quality of life (QOL) questionnaire (QLQ-C30). The QLU-C10D covers four functional (physical, role, social, emotional) and six symptom domains (pain, fatigue, sleep disturbances, nausea, etc.). Within an EORTC project, QLU-C10D valuation studies have been performed in seven European countries. Spain was the last country for which utility weights were determined before the coronavirus disease 2019 (COVID-19) pandemic. A second valuation study in Spain during the pandemic has just been completed. Aims: of the presentation are:—to compare Spanish QLU-C10D utility weights and those of other European countries,—to compare Spanish utility weights and respondents’ QOL before and during the COVID-19 Crisis. Methods: The first valuation study was run in an online sample of the Spanish general population, quota sampled for age and sex, in August 2019. A discrete choice experiment (DCE) was applied to elicit utilities. The survey also included socio-demographic and clinical information and QOL profile data (QLQ-C30). Recruitment and assessment were contracted to a company specialized in DCEs. Data were analyzed by conditional logistic and mixed logit models. A second study applied the same methods during the COVID-19 pandemic (April 2020). Results: In the first valuation study, 1010 respondents (mean age 47.1, 50.5% female) were eligible for analysis. Among QLU-C10D domains, physical functioning received the largest utility weights, followed by pain, role functioning, nausea, and social functioning, similar to rankings in other European countries. In the second study ( $n = 504$ , during COVID-19 pandemic), QLU-C10D utility weights showed a similar order of domains, but the impact of physical functioning increased substantially. Participants’ self-reported QOL indicated significantly lower role and emotional functioning, and somewhat less fatigue and pain compared to pre-COVID-19 data. Conclusion: The Spanish valuation results conformed with those of other countries, adding to the face validity of the QLU-C10D. Formal psychometric investigations are underway. The COVID-19 pandemic has severely affected daily life in most countries. In Spain, this was reflected in significant changes in QOL and in health preferences. Before new QLU-C10D valuations are performed, further monitoring of the impact of the pandemic is advised.

### (107.3) Missing patient-reported diary data—rule-based versus model-based imputation?

Pavol Kral, IQVIA, Bratislava, Slovakia; Konstantina Skaltsa, IQVIA, Barcelona, Spain; Matthew Reaney, IQVIA, Reading, United Kingdom; Michael O’Kelly, IQVIA, Dublin, Ireland

Aims: Missing data in patient-reported outcome (PRO) diary data are common, either when intermittent, i.e., missing days within a period, or monotone when patients withdraw early. Average weekly scores would typically be calculated and used as efficacy assessments. A common rule applied is that more than half of the daily scores should be

observed for a weekly average to be calculated, else this is considered missing. Regulators are increasingly questioning this heuristic and are looking for an optimal rule that would not obscure, neither exacerbate treatment effects. This research explores the possibility of imputing missing daily scores, instead of using the prescribed rules for filling them in. We explore the impact this has on the weekly average estimates through a simulation study. Methods: A 0–10 numerical rating scale (NRS), e.g., worst pain item, was considered as the score collected daily for 7 days. A simulation study was set-up with 100 samples generated by two normal distributions, one for each treatment arm. Sample size was 200 for each arm. Low variability in NRS scores was considered between consecutive days. Missing data were generated by deleting 1, 2, 3, 4, 5, 6 days from 25% of the subjects exploring three scenarios: under the missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR) assumptions. Multiple imputation assuming MAR was performed. Average weekly scores were calculated using the 4-non-missing-days rule, as well as on the datasets with missing data imputed. Estimates for each scenario were assessed with respect to their bias and its variability compared with the true average weekly score from the initial full dataset. Results: Bias is low both when applying the 4-non-missing-days rule and with multiple imputation of missing days. Higher variability was observed as the amount of missing days increased in both approaches, however variability was lower for multiple imputation for all scenarios. This could have a significant impact when further analyzing these scores in search of a treatment effect. Conclusion: Methods: that account for missingness and its causes should be applied also in the diary setting. Consideration should be carefully given to ensure true treatment effects are not obscured or exacerbated.

#### (107.4) The establishment of minimal clinically important differences for symptom recovery from lung cancer surgery

Wei Xu, Chongqing Medical University, Chongqing, China; Qiuling Shi, Chongqing Medical University, Chongqing, China; Wei Dai, Sichuan Cancer Hospital, Sichuan, China; Yang Pu, Chongqing Medical University, Chongqing, China; Hongfan Yu, Chongqing Medical University, Chongqing, China; Qingsong Yu, Chongqing Medical University, Chongqing, China

Aims: To establish the improvement minimal clinically important difference (MCID) of major symptoms for patients recovering from lung cancer surgery. Methods: We used the MD Anderson Symptom Inventory-lung cancer module (MDASI-LC) to assess symptoms of patients. We identified the top 5 symptoms based on the mean severity scores of each symptom on the first day post-surgery. Then, we calculated the symptoms change scores and interference difference respectively, between first day post-surgery and discharge day. We used linear regression and Pearson correlations to select significantly relations between symptoms difference and MDASI six-interference difference. With the change scores of six MDASI interference as anchor, multivariate analysis of variance (MANOVA) was applied to determine the MCID for symptom recovery. MCID are based on the largest F ratio from MANOVA. We calculated both the absolute difference and relative difference of symptom change. For validation, we used the MCID to categorize patients into improvement and without improvement group in terms of their symptom change scores, and assessed how the categories were concordant with single-item change score of quality of life (QoL). Results: Of 482 patients, according to our inclusion criteria, we finally choose symptoms as the pain ( $5.42 \pm 2.61$ ), fatigue ( $4.76 \pm 2.89$ ), and shortness of breath ( $3.59 \pm 2.83$ ) to establish MCIDs, which were significantly increased with MDASI six-interference (all  $p < 0.0001$ ). Based on the MANOVA, recovery MCID of absolute difference and relative difference was 2 and 30% for pain, respectively. 2 and 30%

for fatigue, respectively. 3 and 40% for shortness of breath, respectively. The MCID for symptom were validated using QoL, symptoms change scores categorize to improvement group (difference  $\geq$  MCID), and without improvement group (difference  $<$  MCID). Pain, fatigue, and shortness of breath were significantly concordant with QoL (all  $p < 0.01$ ). Conclusion: We used the anchor-based method, through absolute difference and relative difference, to establish the MCIDs for recovery of pain, fatigue, shortness of breath from lung cancer surgery. Results: from this survey will promote routine assessment and management of patient-reported symptom recovery. Meanwhile, These MCID may facilitate the conduct and interpretation of clinical evaluation, symptom epidemiology, and clinical trials.

Table 1. MCID Analysis using 3 Criteria  
(Rating symptom absolute difference as 1–5 and relative difference as 10%–50%)

MCID for recovery cutpoint	Wilks lambda	Pillai's trace	Hotelling's trace	F	P
<b>Pain</b>					
2	0.9043	0.0957	0.1058	5.78	<.0001
30%	0.9202	0.0798	0.0867	4.7	0.0001
<b>Fatigue</b>					
2	0.8716	0.1284	0.1474	6.7	<.0001
30%	0.8641	0.1359	0.1573	7.1	<.0001
<b>Shortness of breath</b>					
3	0.8773	0.1227	0.1399	4.92	<.0001
40%	0.8975	0.1025	0.1142	3.98	0.0009

Table 2. Validation with QoL [Mean (SD)]

	QoL mean (SD)	F	P
<b>Pain-2</b>			
improvement	5.98 (2.45)	5.83	0.0163
without improvement	5.23 (2.45)		
<b>Fatigue-2</b>			
improvement	5.87 (2.21)	6.46	0.0115
without improvement	5.05 (2.02)		
<b>shortbreath-3</b>			
improvement	5.94 (2.19)	6.02	0.0149
without improvement	5.22 (2.18)		
<b>Pain-30%</b>			
improvement	6.03 (2.19)	8.92	0.003
without improvement	5.15 (2.43)		
<b>Fatigue-30%</b>			
improvement	5.95 (2.21)	11.21	0.0009
without improvement	4.92 (1.99)		
<b>Shortness of breath-40%</b>			
improvement	5.91 (2.16)	5.78	0.0171
without improvement	5.17 (2.11)		

#### (107.5) Evaluating the predictive validity of PRO measures: generalizing ROC curve analysis to non-binary outcomes

Xiaochen Lin, Optum, Johnston, Rhode Island, United States; Regina Rendas-Baum, Optum, Johnston, Rhode Island, United States

Aims: Establishing the predictive validity of a patient-reported outcome (PRO) measure significantly improves its applicability, namely in predicting a patient's response to a medical intervention. Receiver operator characteristic (ROC) curve analysis has been typically used when the objective is to predict membership into a binary classification. However, when the underlying classification is not binary (e.g., non-response, partial response, or full response), dichotomizing leads to misclassification and the loss of statistical power. This study illustrates how to generalize ROC curve analysis to evaluate a PRO's performance in predicting non-binary classifications and how PRO's distributional parameters impact accuracy indices used to evaluate

predictive validity. Methods: Data from a clinical trial in cirrhosis were used to illustrate how to evaluate a PRO measure's performance to predict an ordinal outcome (4-level disease stage) and a continuous outcome (albumin). For each outcome, the PRO measure was simulated under scenarios with varying (1) correlation between the PRO and the outcome (0.4–0.8), (2) mean of the PRO measure (40–50), and (3) standard deviation of the PRO measure (7.5–12.5); (4) scale granularity: number of unique values (a proxy for number of items in the PRO measure). Accuracy indices analogous to the traditional area under the ROC curve (AUC) were computed for each scenario. Results: Accuracy improved for both types of outcomes as the correlation with the PRO measure increased, but was stable across simulated scenarios with varying distributional characteristics of the PRO measure (Figure), including its granularity. Specifically, as the correlation increased, accuracy increased from 0.62 to 0.79 for the continuous outcome, and from 0.67 to 0.91 for the ordinal outcome. Acceptable values of accuracy ( $\geq 0.7$ ) were achieved at lower levels of PRO–outcome correlation when the outcome was ordinal. Dichotomizing the original outcomes lead to widely different accuracy indices, even when the underlying correlation was constant. Conclusion: Generalized ROC curve analysis should be used for assessing predictive validity of PRO measures when the prediction involves non-binary outcomes. Traditional ROC curve analysis can be generalized to ordinal and continuous outcomes, to avoid the bias introduced in measures of accuracy when the outcome is dichotomized. These methods are easily applied though infrequently used.

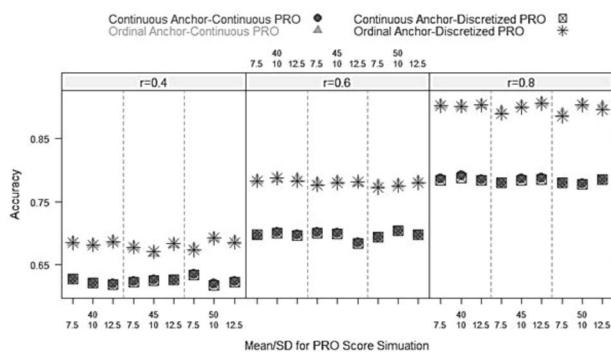


Figure. Accuracy of Predicting Ordinal and Continuous Outcomes Using Simulated PRO Measure

## 108: Innovation in a Digital World

### (108.1) Using observation and visual methods to explore quality of life and well-being with technology-dependent young people

Elizabeth Straus, University of British Columbia, Vancouver, British Columbia, Canada; Helen Brown, University of British Columbia, Vancouver, British Columbia, Canada; Fuchsia Howard, University of British Columbia, Vancouver, British Columbia, Canada; Gail Teachman, Western University, London, Ontario, Canada

Aims: The concepts of quality of life and well-being have garnered significant attention in research with persons using home mechanical ventilation (HMV) technology; however, little is known about what constitutes quality of life and well-being for young people (ages 16–40 years) living with HMV. In the absence of young people's perspectives, normative assumptions about what constitutes a good life could be taken up and contribute to unmet needs and marginalization. The aim of this presentation is to articulate how observation and visual methods were used in the context of a critical narrative inquiry study to explore quality of life and well-being for young people living with

HMV. Methods: A critical narrative inquiry methodology that views meanings as understood and lived in relation with others and social contexts was developed. Phase 1 consists of observations with participants in their everyday lives. Phase 2 involves photo-elicitation, in which participants generate a collection of photographs over 2–3 weeks that reflect their daily lives, what is important or distressing to them, and what makes their lives easier or harder. They then caption and share stories about the photographs. In each phase, specific methods are co-constructed with each participant. Results: Exemplars from the critical narrative inquiry will be presented to illustrate how observing these young people in their daily lives and engaging in photo-elicitation can generate knowledge of what matters to them, the influence of social contexts and the ways their stories reflect or challenge assumptions about what constitutes a good life. Conclusion: Critical narrative inquiry using observations and visual methods is well suited to exploring the varied ways in which quality of life and well-being are understood and lived and the ways social contexts feature in and influence their lives and meanings ascribed to these concepts. This knowledge can inform programs and policy as well as the selection of methods to assess quality of life and well-being that align with young people's perspectives and contexts. This novel methodological approach may also open up possibilities for quality of life research with other groups of persons with disabilities and/or communication impairments.

### (108.2) Bipolar Bridges: a digital health innovation targeting quality of life in bipolar disorder

Emma Morton, PhD, Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada, Vancouver, British Columbia, Canada; Steven Barnes, Department of Psychology, University of British Columbia, Vancouver, BC, Canada, Vancouver, British Columbia, Canada; Erin Michalak, PhD, Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada, Vancouver, British Columbia, Canada

Aims: Quality of life (QoL) is increasingly recognized as a key outcome of self-management interventions for bipolar disorder (BD). Mobile phone applications (apps) have enormous potential to increase access to evidence-based self-management strategies and provide real-time support. However, apps developed to facilitate mood-monitoring do not address the full spectrum of QoL domains that individuals with BD have nominated as important. Bipolar Bridges is a platform under development that aims to enable users to keep track of and optimize their QoL. Users of the app will grant permissions to collect, securely store, and integrate data from other third party health apps (e.g., sleep/activity trackers), and passively collected data (e.g., step count). The app will provide recommendations for evidence-informed self-management strategies based on the user's QoL profile. Methods: The Bipolar Bridges platform builds on the web-based adaptation of a BD-specific QoL self-assessment measure, and will integrate material from a web-based portal providing information on evidence-informed self-management strategies in BD. A combination of community-based participatory research methods, user-centered design principles, and online surveys will be used to inform the content and design of the beta app. Results: Progress to date has included validation of the psychometric properties on a web-based adaptation of the QoL.BD, a BD-specific QoL measure ( $n = 498$ ). A mixed-methods evaluation of the web-based information portal found significant impacts on QoL ( $n = 94$ ). An online survey of how patients ( $n = 141$ ) and healthcare providers ( $n = 19$ ) currently use technology to support QoL will inform which third-party apps are integrated with the beta app. A persona development exercise (currently in the pilot phase of development) will communicate user needs and preferences to the development team. Conclusion: Apps are able to facilitate access to self-management strategies in BD, however user

preferences for tools that support outcomes beyond symptoms are not being adequately addressed, nor are advances in online technologies being fully leveraged. The Bipolar Bridges project will build on a decade of research on QoL in BD along with innovative technological approaches to develop an app that will enable access to tools to optimize health and QoL.

#### **(108.4) Engaging youth and caregivers in the development of Live Yes! JA INSIGHTS—an online platform to longitudinally track juvenile arthritis-related Patient-Reported Outcome Measures (PROMs)**

Rebecca Butcher, MS, MPH, The Dartmouth Institute for Health Policy & Clinical Practice, Lebanon, New Hampshire, United States; Erin Knight-Zhang, PhD, The Dartmouth Institute for Health Policy & Clinical Practice, Lebanon, New Hampshire, United States; Andrea Ring, B.S., The Arthritis Foundation, Atlanta, Georgia, United States; Karen Schifferdecker, PhD, MPH, The Dartmouth Institute for Health Policy & Clinical Practice, Lebanon, New Hampshire, United States

**Aims:** The Arthritis Foundation's Live Yes! INSIGHTS program is an online population-health PROM platform quantifying arthritis-related experiences over time in three domains: physical health, social-emotional health, and experience of care. INSIGHTS data inform programming, online resources and research. Youth and caregiver input played an essential role in the selection of PROMs and overall design and implementation of a Juvenile Arthritis (JA) version of INSIGHTS. We describe the systematic mixed-methods approach used to collaboratively develop this program. **Methods:** We identified 23 JA-specific PROMs across the three domains through a comprehensive literature review. Three candidate instruments were selected for each domain by a workgroup based on instrument psychometric strength and practicality of administration (e.g., respondent burden, cost). We recruited teens and young adults with JA ( $n = 12$ ; ages 15–24) and adult caregivers ( $n = 39$ ) for a modified Delphi process to review the candidate PROMs and assess relevance and ease of answering. Teens reviewed self-report versions of measures; caregivers reviewed parent-proxy versions. Results: were then discussed in five in-person nominal groups (3 parent; 2 teen) to reach consensus on preferred instruments. We convened four additional focus groups ( $n = 35$  parents, 8 teens), controlled for diversity across diagnoses, disease parameters, geography, and age to gather additional input on JA INSIGHTS implementation, including recruitment, messaging and sustaining participation. Results: were triangulated to determine final tool selection and deployment strategies. Results: JA teens and caregivers were initially mixed in their measure preferences. Teens favored longer, more specific physical function measures, but shorter, high-level measures of social-emotional health. In contrast, caregivers opted for shorter instruments overall and voiced challenges in rating their child's social-emotional health. Given mixed early results, focus groups played a critical role in reaching consensus for selecting PROMIS-25 and the Patient Self-Advocacy Scale. Additional key input included if and how families could share data, options for personal tracking, and how INSIGHTS could support families and promote future JA research. **Conclusion:** Our Patient-Centered Outcomes Research approach ensured robust patient engagement in the development of JA INSIGHTS without compromising PROM psychometrics. The Live Yes! INSIGHTS program is now being implemented nationally through the Arthritis Foundation's patient networks.

#### **(108.5) Developing a core set of mobility domains among individuals with acquired brain injury (ABI): Empowering the creation of core outcome sets using natural language processing (NLP)**

Rehab Alhasani, PT, BSc., MSc., PhD (candidate), McGill University, Montreal, Quebec, Canada; Mathieu Godbout, MSc., Université Laval, Quebec, Quebec, Canada; Audrey Durand, PhD, Université Laval, Quebec, Quebec, Canada; Sara Ahmed, PhD, McGill University, Montreal, Quebec, Canada

**Aims:** To develop a core set of mobility domains among individuals with acquired brain injury (ABI) using NLP, unsupervised machine learning, guided by the International Classification of Functioning, Health, and Disability (ICF) ontology. **Methods:** An umbrella review of 47 reviews evaluating the content of mobility measures among individuals with ABI was conducted. A search was performed on 5 electronic databases between 2000 and 2020. Two independent reviewers retrieved copies of the measures and extracted mobility domains. A pre-trained BERT model (state-of-the-art model for NLP), provided vector representations (i.e., embeddings) for each sentence using ICF terms as a guide. A Principal Component Analysis (PCA)<sup>2</sup> was then applied to reduce the embeddings' dimension before applying a k-means algorithm<sup>3</sup> to retrieve clusters of similar sentences. The resulting embedding clusters were evaluated using the Silhouette score,<sup>4</sup> a clustering metric based on inter- and intra-cluster distances, high Silhouette score means that elements in a given cluster are similar and that different clusters are distinct. **Results:** The study included 474 domains extracted from 246 mobility measures. Encoding the clusters using the ICF ontology helped in clustering the domains in a way that is more closely related to mobility terminology. Our best grouping according to human evaluation obtained a 0.47 Silhouette Score and recognized the following clusters (guided by the ICF): Self-care, Environmental factor, Physical functioning, Cognition, Psychosocial and Sensory functions and pain. **Conclusion:** Improved outcome assessment by developing a core set of domains can substantially improve clinical research by allowing comparisons across studies and clinical settings. Compared to traditional manual consensus, utilizing NLP helps researchers to develop a core set of domains more efficiently and synthesize literature that manually is nearly impossible. Adding the ICF ontology to the pool of vectors before clustering forced the clusters to be centered around ICF terms related to mobility. Some limitations related to NLP: Silhouette score is only based on distances between output embeddings and therefore may not necessarily indicate the best grouping (from an expert perspective); and the process of improving clusters is inevitably related to experts' judgments.

### **109: Cardiology Clinical Research**

#### **(109.1) Group concept mapping: capturing important patient-reported outcome domains for Left Ventricular Assist Device recipients**

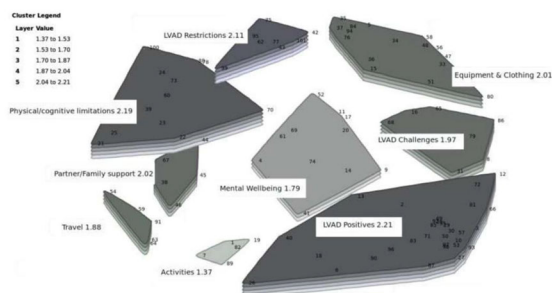
Anita Slade, PhD, University of Birmingham, Birmingham, United Kingdom; Margaret, E. O'Hara, PhD, University Hospitals Birmingham, Birmingham, United Kingdom; David Quinn, FRCS (C-Th), MD, Department of Cardiac Surgery, University Hospitals Birmingham, Birmingham, United Kingdom; Laura Marley, BSc Hons, University Hospitals Birmingham, Birmingham, United Kingdom; Stephen Griffith, Centre for Patient Reported Outcomes



Research, University of Birmingham, Birmingham, United Kingdom; Melanie Calvert, PhD, Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, United Kingdom; M. Sayeed Haque, PhD, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Sern Lim, MD, University Hospitals Birmingham, Birmingham, United Kingdom; Deirdre, A. Lane, PhD, Liverpool Centre for Cardiovascular Sciences, University of Liverpool and Liverpool Heart and Chest Hospital, Liverpool, United Kingdom

**Aims:** Advanced Heart failure carries a high mortality rate. Heart transplantation offers good outcomes for patients, but low donation rates limit this option. A left ventricular assist device (LVAD) is an alternative therapy. This mechanical pump is implanted into the recipient and assists the failing left ventricle, relieving symptoms, while improving prognosis and quality of life. However, living with an LVAD requires significant psychological and physical adaptation. Discussions with a patient and public involvement group including LVAD recipients found a range of issues that needed investigation before key patient-reported outcomes could be identified. **Aim:** To develop a conceptual map of key areas and domains that reflect LVAD recipients' experiences, and their importance in patient-reported outcomes. **Methods:** Groupwisdom™ concept mapping software was used. Group concept mapping (GCM) is a semi-quantitative mixed-methods approach that recognizes participants' expertise in their own health experiences. Participants were recruited from a regional transplant center in the UK. After consent was obtained, participants were given a unique ID, password and link to the GCM software. GCM consists of 3 steps: item generation, item sorting, and rating items for importance, relevance and frequency of impact. Each activity was completed by all participants before moving to the next step. Multidimensional scaling and hierarchical cluster analysis produces visual representations of their experiences as a cluster map, and average rating of items across the clusters. **Results:** 18 LVAD recipients consented to take part. 101 items and 9 clusters were generated. Clusters represented: Activities; Partner/family support; Travel; Mental wellbeing; LVAD challenges; Equipment and clothing; Physical and cognitive limitations; LVAD restrictions, and LVAD positives. LVAD Positives was the most homogenous group, and along with restrictions rated high for frequency, relevance and importance. Physical and cognitive limitations were rated high for importance and frequency. Equipment was rated high for relevance and frequency, and Challenges was rated high on relevance. **Conclusion:** GCM is a useful tool for mapping key areas of importance for LVAD recipients when prioritizing patient-reported outcome domains for use in clinical practice, and future research. Items identified within clusters will be used to identify potential PROMs for future use with LVAD recipients.

Figure 1 Cluster map with average rating value for importance of items within each cluster

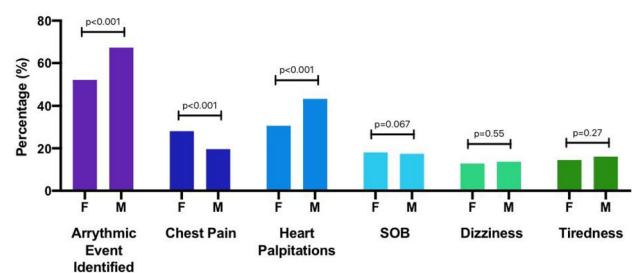


Participants were asked to rate each item for importance (1 = Not important, 2 = Important, 3 = Very important) layers represent average rating scores per item per cluster. The more layers the higher the average scores across the cluster. Label includes average score per cluster.

### (109.2) Sex-based differences in cardiopulmonary symptoms among patients with atrial fibrillation

Brian Zenger, University of Utah, Salt Lake City, Utah, United States; Jeffrey L. Turner, DO, University of Utah, Slc, Utah, United States; T. Jared Bunch, MD, University of Utah, Slc, Utah, United States; Rachel Hess, MD, University of Utah, Slc, Utah, United States; Benjamin A. Steinberg, MD, University of Utah, Slc, Utah, United States

**Aims:** Gender differences in patient-reported outcomes with atrial fibrillation (AF) and other cardiac arrhythmias have been well documented. However, the differences in symptoms have not been corroborated with real-time ambulatory rhythm monitoring. We sought to characterize the gender differences in symptom-rhythm correlation among patients with AF. **Methods:** Clinically ordered ambulatory rhythm monitoring studies among patients with a history of AF were analyzed (duration 7–30 days). Patients were provided with standard instructions to trigger and document symptoms (including shortness of breath (SOB), chest pain, dizziness, palpitations, or tiredness). Heart rhythm was simultaneously recorded and annotated. **Results:** 293 women and 381 men underwent ambulatory rhythm monitoring. Overall, automatically triggered and patient triggered totaled to 8,140 events ( $n = 4008$  female, %). Arrhythmia occurred without symptoms in 1018 events for women (35% of events with either arrhythmia or symptoms) versus 66% for men ( $n = 1815$ ;  $p < 0.001$ ). Patients reported symptoms in a total of 2885 events ( $n = 1918$  in women, 66.5%). Females' symptoms were correlated with a documented arrhythmic event ( $p < 0.001$ ). Females were more likely to note chest pain (28.3% vs. 19.8%,  $p < 0.001$ ) and less likely to report heart palpitations (30.4% vs 43.3%,  $p < 0.001$ ) compared with males. There were no differences in reported symptoms of fatigue, SOB, dizziness, or tiredness between females and males. **Conclusion:** Females' symptomatic events were less likely to correlate with cardiac arrhythmia, compared with males, and significant self-reported variation in symptoms exists between females and males. As the presence of symptoms often drives treatment of AF, these data indicate that reported symptoms and identified cardiac arrhythmias may have significant differences between sexes, and should be considered in clinical practice.



### (109.3) Differential item functioning due to comorbid chronic health conditions in the quality of life measure in the National Health and Aging Trends Study

Frances Yang, PhD, University of Kansas Medical Center, Kansas City, Missouri, United States; Makenzie Danley, University of Kansas Medical Center, Kansas City, Kansas, United States; Solon Kao, DDS, University of Missouri Kansas City, Kansas City, Missouri, United States

**Aims:** With the rapid onset of the COVID-19 pandemic, older adults are one of the highest at-risk populations with decreased immunity to the virus culminated with potentially multiple chronic health conditions. The concern of older adults is their quality of life and well-being before and after the COVID-19 pandemic in the US in the midst of having comorbid chronic health conditions. To identify potential differences in responses to quality of life (QOL) questions and the impact of chronic health conditions, we examine the potential measurement differences in quality of life before and after the consideration of the presence of one or more chronic health conditions in older Americans. **Methods:** In a national sample of 5,150 older adults, aged 65 and older, the latest round of data collection in a national longitudinal cohort known as the National Health and Aging Trends Study. Mplus version 8 was used to conduct differential item functioning (DIF) analysis in a latent variable modeling framework. There were nine items for QOL. The following nine chronic health conditions were examined simultaneously together and on both the general and specific factors in a step-wise manner using modification indices to detect DIF: arthritis, cancer, diabetes, hbp, heartattack, heartdisease, lung, osteoporosis, and stroke. **Results:** Based on standard published criteria for best fitting models (Comparative Fit Index (CFI) > 0.95, Root Mean Squared Error of Approximation (RMSEA) < 0.06) Hu and Bentler, 1999), the bifactor model showed adequate fit ( $X^2 = 718$ ,  $df = 34$ ,  $p < .001$ , Comparative Fit Index = 0.96, RMSEA = 0.06). There were two specific factors of sentimentality (often feel cheerful, often feel bored, often feel full of life, and often feel upset) and vitality (life has meaningful purpose, feel confident, gave up trying, like living situation, and finds a way); and the general QOL factor that had higher loadings than the specific factors. **Conclusion:** Differential item functioning was found on both the general and specific factors of QOL due to one or more of the chronic health conditions. Scores for QOL are reported as adjusted for DIF to show different profiles of chronic health conditions in relation to inflated as compared to deflated scores of QOL in the context of the COVID-19 pandemic.

#### (109.4) Impact of socio-cultural factors and environmental resources on the health-related quality of life of children and adolescents after congenital heart disease surgery: parental experiences from a low middle income country

Laila Akbar Ladak, PhD, MScN, BScN, RN, The Aga Khan University, Karachi, Pakistan; Robyn Gallagher, PhD, BA, MN, RN, FAHA, FESC, Susan Wakil School of Nursing and Midwifery, Sydney Nursing School, The University of Sydney, Sydney, Faculty of Medicine and Health, Sydney, New South Wales, Australia. Charles Perkins Centre, Sydney Nursing School Faculty of Medicine and Health, Sydney, New South Wales, Australia, Sydney, Australia; Babar Sultan Hasan, M.D., FAAP, The Aga Khan University, Pakistan, Karachi, Pakistan; Khadija Awais, MBBS, The Aga Khan University, Pakistan, Karachi, Pakistan; Ahmed Abdullah, MBBS, The Aga Khan University, Pakistan, Karachi, Pakistan; Janice Gullick, PhD, MArt, BFA, RN, Susan Wakil School of Nursing and Midwifery, Sydney Nursing School, The University of Sydney, Sydney, Faculty of Medicine and Health, Sydney, New South Wales, Australia, Karachi, Australia

**Aims:** Health-related quality of life (HRQOL) is an important indicator of long-term well-being, influenced by environmental factors such as family, culture, societal norms and available resources. This study aimed to explore the influence of individual, socio-cultural and environmental factors that influence parents' health-related decision-making and the perceived HRQOL for children or adolescents after Congenital Heart Disease (CHD) surgery in Pakistan. **Methods:** A descriptive, qualitative design, guided by the Social Ecological model was utilized to explore the experiences and perceptions of 18 parents of children/

adolescents who had surgery for congenital heart defects (CHD) in a low middle income country (LMIC), Pakistan. A content analysis checklist and COREQ (Consolidated criteria for REporting Qualitative research) were used to ensure study rigor. **Results:** At the intrapersonal level, unrealistic expectations of surgery, residual CHD symptoms and difficulty maintaining educational progress were of greatest concern. There were low levels of health literacy and understanding about CHD among family and friends, however, strong kinship ties were an important resource at the interpersonal level. These families lived in poverty and the mothers carried the sole burden of care for their sick children. From the institutional-level aspect, there were unclear expectations of the child's needs at school, and parents had poor access to psychological, family planning and genetic counseling, and poor access to CHD education resources. At sociocultural level, religion and trust in God were important coping factors, however, CHD was a gendered experience with particular concerns around scarring and the marriageability of girls. Parents noted the deficit of antenatal and specialist CHD services, and felt the consequence of a lack of a universal health care system at the public policy level. **Conclusion:** Diverse socio-ecological factors explain the different HRQOL outcomes in LMICs for children living with CHD after surgery. These specific contexts should inform future improvements and interventions in countries like Pakistan.

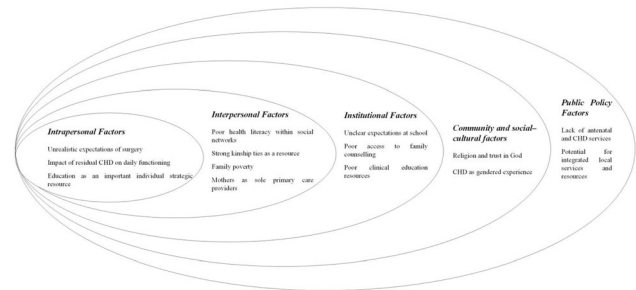


Figure 1 Factors influencing the health related quality of life in congenital heart disease surgical patients as perceived by their parents using the socioecological model

#### (109.5) Multimorbidity and health-related quality of life after myocardial infarction: A multicenter longitudinal patient-reported outcomes study

Theresa Munyombwe, PhD, University of Leeds, Leeds, United Kingdom; Suleman Aktaa, MD, University of Leeds, Leeds, United Kingdom; Chris Wilkinson, MD, University of Newcastle, Newcastle, United Kingdom; Marlous Hall, PhD, University of Leeds, Leeds, United Kingdom; Tatendashe Dondo, Phd, University of Leeds, Leeds, United Kingdom; Oras Alabas, PhD, University of Leeds, Leeds, United Kingdom; Ben Hurdus, MBBS, Leeds Teaching hospital, Leeds, United Kingdom; West Robert, Phd, University of Leeds, Leeds, United Kingdom; Hall Alistair, PhD, University of Leeds, Leeds, United Kingdom; Chris Gale, PhD, University of Leeds, Leeds, United Kingdom

**Aims:** To determine the association between multimorbidity and changes in HRQoL in survivors of MI. **Methods:** A longitudinal national cohort study of 9566 hospital survivors of MI was conducted from 77 National Health Service hospitals in England between 1st November 2011 and 24th June 2015 and followed over one year. EuroQol 5 dimension 3 levels (EQ-5D-3L score) and Visual Analogue Scale (VAS) were used to measure HRQoL at hospitalization, and at one, six and twelve months following AMI. Baseline multimorbidity subgroups were determined using latent class analysis(LCA), and the association between these subgroups and changes in HRQoL quantified using multilevel modeling. **Results:** Of 9566 survivors of MI (7154 [75%] men, mean age 64.1 years [SD

11.9)], more than half 53.5% ( $n = 5119$ ) had one or more conditions (multimorbidity) including hypertension 4,078 (42.6%), previous angina 1792 (18.7%), peripheral vascular disease (PVD) 428 (4.5%), diabetes mellitus 1,714 (17.9%), COPD 1166 (12.2%), cerebrovascular disease (CVSD) 428 (4.5%), chronic renal failure 289 (3.0%), heart failure 212 (2.2%). LCA identified three distinct multi-morbidity clusters a severe multi-morbidity class (6.5%) with predominantly hypertension, diabetes, chronic renal failure, heart failure and COPD; a moderate multi-morbidity (47.6%) with predominantly hypertension and diabetes, and mild multi-morbidity (45.9%). Compared to the moderate and mild multimorbidity classes, patients in the severe multimorbidity class were older (mean age 74.8 vs 68.8 and 57.5 years) and more commonly presented with NSTEMI (86.0% vs. 66.7% and 49.3%). The severe multimorbidity class had a lower EQ-VAS score compared to patients in the moderate class (difference 3.32, 95% CI 1.84 to 4.80) and the mild multimorbidity class (difference 3.34, 1.66 to 5.01). Conclusion: For hospital survivors of MI, multimorbidity is common and associated with poor HRQOL especially in older people, women and patients with NSTEMI. Distinct multimorbidity HRQOL clusters may be readily identified, for whom interventions could be designed and tested to improve HRQOL.

## 110: Advancing qualitative methods in the development and application of PROs

### (110.1) Quality of life in colorectal cancer survivors: A systematic synthesis of qualitative research

Chloe Lim, PhD Candidate, The University of Sydney, Sydney, Australia; Rebekah Laidsaar-Powell, PhD, The University of Sydney, Sydney, Australia; Jane Young, PhD; Professor, The University of Sydney, Sydney, Australia; Steven Kao, BHB MBChB PhD FRACP, Chris O'Brien Lifehouse, Sydney, Australia; Yuehan Zhang, PhD Candidate, Australian National University, Canberra, Australia; Phyllis Butow, PhD; Professor, The University of Sydney, Canberra, Australia

**Aims:** Colorectal cancer (CRC) is the world's third-most common cancer, with a five-year survival rate of 65-90%. For clinicians to understand unmet needs and provide optimal care for CRC survivors, qualitative research into their psychosocial experiences and quality of life (QoL) is imperative. No qualitative systematic reviews in this area have been identified. We aimed to fill this gap. **Methods:** Five databases (PsycINFO, MEDLINE, Embase, CINAHL, PubMed) were searched with terms related to "colorectal," "cancer," "survivorship," and "qualitative research." A second search was conducted in PubMed with an "advance\*" search term replacing "survivorship." Additional searching of reference lists, author names, and citations occurred. We included survivors' experiences across the survivorship trajectory, excluding the palliative phase. Titles, abstracts, and full texts were screened. Included articles underwent data extraction, bias ratings using the CASP qualitative checklist, and thematic synthesis. Approximately 10% of articles at each stage of review were cross-checked by a second rater; disagreements were discussed until agreement was reached. **Results:** De-duplication left 1871 articles. After title/abstract screening, 284 full text articles were reviewed, with a final 77 articles included. Studies primarily originated from western countries (mostly USA and Europe) and focused on curative subpopulations and short-term outcomes. Specific treatment procedures of participants were poorly reported. Thematic synthesis revealed 7 overarching themes of CRC survivorship: physical symptoms; functional limitations; psychosocial impacts; financial impacts; interactions with the healthcare system; coping strategies; and, positive outcomes of cancer. Studies showed that bowel functioning was the main cause of functional limitations and

negative QoL. Additionally, stomas posed threats to survivors' body image and confidence. Returning to work was challenging for survivors, due to physical symptoms and financial burdens. Survivors' unmet needs included lack of, or conflicting, information provided by healthcare professionals regarding symptom expectations and health management, and lack of ongoing support throughout follow-up and recovery. **Conclusion:** CRC impacts survivor's QoL in all areas, and thus a co-ordinated supportive care response is required to address survivors' unmet needs. To address research gaps identified, future qualitative studies should focus on advanced CRC subpopulations, treatment-specific impacts on QoL, and long-term (> 5 years) impacts on CRC survivors.

### (110.2) Results from expert concept elicitation interviews to support development of an international, disease-specific PRO in transthyretin amyloidosis (ATTR)

Lynne Broderick, MPH, Optum, Johnston, Rhode Island, United States; Montserrat Vera Llonch, MD, MPH, MSc, Akcea, Boston, Massachusetts, United States; Meg O'Connor, MPH, Optum, Johnston, Rhode Island, United States; Duncan Brown, PhD, Akcea, Boston, Massachusetts, United States; Isabelle Lousada, Amyloidosis Research Consortium, Newton, Massachusetts, United States; Kristen McCausland, PhD, MPH, Optum, Newton, Rhode Island, United States; Kerry Noonan, Akcea, Boston, Massachusetts, United States; Kristen Hsu, Amyloidosis Research Consortium, Newton, Massachusetts, United States; Michelle White, PhD, Optum, Johnston, Rhode Island, United States

**Aims:** Transthyretin amyloidosis (ATTR) is a rare, debilitating condition caused by misfolded protein deposits in different organ systems. There are no validated ATTR-specific patient-reported outcome measures (PROs) that adequately measure the multifaceted symptoms and impacts experienced by patients. The heterogeneous presentation of ATTR requires an exploration of the patient experience in many countries to ensure the PRO is sensitive to potential differences. This research builds on a prior US-based study and presents findings from qualitative interviews with experts in 9 countries to support the development of an ATTR-specific PRO. **Methods:** Researchers conducted qualitative concept elicitation interviews with clinicians and patient advocates in Brazil, Canada, France, Italy, Japan, Portugal, Spain, Sweden, and the UK to document the relevant and significant symptoms and impacts experienced by patients with ATTR. Interviews were 60-min in duration and conducted by telephone. Interviews followed a semi-structured interview guide and, excepting Japan, were conducted in English. Data were analyzed using thematic analysis. **Results:** Eleven clinicians and 2 patient advocates from 9 countries participated in interviews. Results: indicated the need for an ATTR-specific PRO to: (1) include a comprehensive symptom list with symptoms presented in plain language, and (2) measure disease severity in terms of impacts on daily function, including physical, emotional, social, and work/productivity impacts. A description of each impact is included in Table 1. These findings were consistent with prior work conducted with US-based experts. Further, experts reported that patients in different countries seem to experience ATTR similarly. When asked what type of scoring would be most useful as an output of a new PRO, experts agreed that in addition to a total score, the PRO should derive domain level scores which can help researchers/clinicians pinpoint problem areas and track disease progression. **Conclusion:** Experts provided valuable insights on the patient experience of ATTR and agreed that an ATTR-specific PRO is needed. Their insights have shaped the next step of PRO development, in which patients with ATTR will participate in interviews and share their perspective on the significant symptoms and impacts of ATTR.

Table 1. Impacts of ATTR reported by experts (n=13)

Impact	Description
Physical	Patients are dealing with <b>gross motor impairment and reduced activity</b> (for example walking or climbing stairs) as a result of peripheral neuropathy and cardiomyopathy.
Daily function	<b>Challenges with daily functioning</b> (including activities of daily living and instrumental activities of daily living) are internalized by patients as a loss of independence.
Emotional/mental	Patients often experience <b>anxiety, fear, and depression</b> related to their condition. Patients with hATTR also <b>experience guilt and worry</b> related to passing the condition on.
Social	Patients can become <b>socially isolated</b> , particularly if they suffer from GI issues. Patients with hATTR (asymptomatic, but genopositive) are <b>reluctant to engage in romantic relationships</b> .
Work, productivity	For younger patients, <b>productivity is often reduced and/or they need to stop working altogether</b> . For older patients who tend to be retired when diagnosed, work and productivity is not as impacted. Patients are <b>hesitant to join trials or try new medications</b> because it can mean <b>more time away from work</b> for clinical/study visits.

### (110.3) Using Patient-Reported Outcome measures (PROMs) to promote quality of care in the management of patients with established kidney disease requiring treatment with haemodialysis (PROM-HD): a qualitative study

Nicola Anderson, MSc, Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, United Kingdom; Christel McMullan, PhD, Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, United Kingdom; Melanie Calvert, PhD, Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, United Kingdom; Mary Dutton, MSc, Research and Development, University Hospitals Birmingham NHS Foundation Trust, Birmingham, United Kingdom; Paul Cockwell, PhD, Department of Nephrology, University Hospitals Birmingham NHS Foundation Trust, Birmingham, United Kingdom; Derek Kyte, PhD, Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, United Kingdom

**Aims:** Patients undergoing renal replacement therapy by haemodialysis (HD) commonly report high symptom burden, reduced quality of life (QOL) and often prioritize improvements in their QOL over long-term survival. Systematic collection/use of patient-reported outcome measures (PROMs) in these patients may help tailor care to their needs and improve outcomes. This study explored the views, perceptions and experiences of patients receiving HD and members of the multi-disciplinary team (MDT) on the implementation and use of PROM data **Methods:** Using qualitative methodology, semi-structured interviews were undertaken with 22 patients and 17 MDT members. Sample validated PROMs (IPOS-Renal, KDQOL-SF 1.3, KDQOL-36 1.0), details of core outcomes identified by the Standardized Outcomes in Nephrology (SONG-HD) initiative and a topic guide were used to inform discussion. Transcripts were analyzed deductively using the Consolidated Framework for Implementation Research (CFIR) and inductively using thematic analysis. The CFIR provided a pragmatic structure to report feasibility and acceptability of PROM use in HD settings. **Results:** Analysis identified key practical considerations: (i) frequency (of PROM completion); (ii) timing (around dialysis); (iii) setting (home or in-center); (iv) preferred mode of administration (electronic or paper versions); and (v) interpretation and feedback of the responses. Participants were keen to use PROMs to support the delivery of person-centered care through shared decision-making and management in all dialysis settings. A number of potential advantages of PROM use were highlighted, especially in research settings. However, the complexity associated with PROM interventions was recognized, in particular regarding patient safety

and the need for effective electronic systems. Possible barriers to implementation included: (i) lack of evidence base for use in routine kidney care; (ii) perceived time barriers for staff (work flow interruptions); (iii) patients being overburdened by questionnaires; (iv) risk of over-medicalizing the patient experience; and (v) health literacy issues for patients and less experienced staff. **Conclusion:** To assess whether PROMs can promote quality of care in HD settings, a comprehensive implementation strategy needs to be devised, considering best available measures and methodological considerations. The findings of this study can assist implementation; addressing the priorities and concerns of both patients and clinicians, including timely understanding of facilitators and barriers.

### (110.4) Novel use of creative elicitation tasks to further explore the patient experience of non-alcoholic steatohepatitis (NASH)

Nicola Williamson, Adelphi Values Ltd, Bollington, United Kingdom; Katie Tinsley, Adelphi Values Ltd, Bollington, United Kingdom; Euan McLeod, Pfizer Inc, Tadworth, United Kingdom; Naim Alkhouri, Texas Liver Institute, San Antonio, Texas, United States; Chloe Tolley, Adelphi Values Ltd, Bollington, United Kingdom; Adam Gater, Adelphi Values Ltd, Bollington, United Kingdom; Magdalena Harrington, Pfizer Inc, Cambridge, Massachusetts, United States

**Aims:** Prior qualitative concept elicitation (CE) interview and focus group studies in non-alcoholic steatohepatitis (NASH) have identified symptom and impact concepts experienced by NASH patients. However, the clinical presentation of NASH presents particular challenges for understanding the patient experience and monitoring treatment outcomes. This case study outlines the novel use of elicitation tasks to address the constraints of traditional CE interviews in the context of NASH; a chronic condition, characterized by non-specific symptoms, co-occurrence of comorbidities and complex medical history. **Methods:** Qualitative semi-structured CE interviews were conducted with biopsy-proven NASH patients ( $n = 20$ ). Interviews comprised open-ended CE questioning followed by interactive elicitation tasks (body map, diagnosis timeline, matching, patient journey timeline and ranking), to explore symptoms, impacts on health-related quality of life (HRQoL), symptom attribution to NASH or comorbidities, the chronology of the disease experience and interactions with clinicians and bothersome-ness of symptoms/impacts. Patients completed each elicitation task using the materials provided (i.e., worksheets and cards), then described their NASH experience using the completed worksheet for guidance. **Results:** Compared to patients' responses to open-ended CE questioning, the elicitation tasks elicited more in-depth information regarding patients' symptom presentation, relative to their comorbidities and long-standing medical history (onset, location, changes over time). The elicitation tasks enabled some patients to more clearly recall their long and complex medical history and make connections, for the first time, between their symptoms, NASH and comorbidities. The diagnosis and patient journey timeline tasks enabled patients to verify their order of events (diagnosis, symptoms, treatments, clinician interactions). The ranking task provided understanding of the relative impact of symptoms and HRQoL impacts on patients' lives due to NASH. **Conclusion:** Findings provided a greater understanding of the NASH patient experience beyond existing literature, exploring attribution of non-specific symptoms to NASH and/or comorbidities for the first time. Elicitation tasks enhanced recall of the patient experience of non-specific symptoms of NASH in the context of co-occurring comorbidities, allowing the lived experience of NASH to be explored comprehensively. This methodology is valuable when

exploring the experience of patients with complex and long-standing health problems, to generate richer and more complete patient experience data.

#### **(110.5) Comparison of qualitative methods to explore key symptom and functional impact concepts of presbyopia: literature review, social media listening, and qualitative interviews**

Amy Findley, Adelphi Values Ltd, Bollington, United Kingdom; Sarah Bentley, Adelphi Values Ltd, Bollington, United Kingdom; Rob Arbuckle, Adelphi Values Ltd, Bollington, United Kingdom; James Wolffsohn, Aston University, Birmingham, United Kingdom; Christel Naujoks, Novartis Pharma AG, Basel, Switzerland; Francesco Patalano, Novartis Pharma AG, Basel, Switzerland; Garima Sharma, Novartis Business Services, Hyderabad, India; Jyothi Kommineni, Novartis Business Services, Hyderabad, India; Nishith Tyagi, Novartis Business Services, Hyderabad, India; Sima Chivara-Razavi, Novartis Pharma AG, Basel, Switzerland

**Aims:** The patient experience of presbyopia (age-related impaired near-vision) was explored to support patient-reported outcome (PRO) development through literature review, social media listening (SML), and qualitative interviews with healthcare professionals (HCPs) and presbyopic individuals. The concepts identified, depth of data, and value of each method is compared. **Methods:** Keyword searches in bibliographic databases and review of abstracts identified 120 relevant publications; in-depth literature review of the qualitative studies identified key symptoms/functioning concepts. SML was conducted using publicly accessible social media sources with focus on ophthalmologic diseases using a pre-defined search string. Relevant posts ( $n = 1470$ ) were analyzed and key concepts identified. Semi-structured concept elicitation interviews were conducted with presbyopic individuals (US  $n = 30$ , Germany  $n = 10$ , France  $n = 10$ ), and HCPs (US  $n = 3$ , France  $n = 2$ , Germany  $n = 1$ , Japan  $n = 1$ ). Verbatim transcripts were coded using thematic analysis. A conceptual model summarized concepts identified across sources. **Results:** Overall, 158 concepts were identified. Qualitative interviews yielded most concepts ( $n = 151/158$ , 96%), with SML yielding a third of the concepts ( $n = 51/158$ , 32%) and the literature review yielding the least concepts ( $n = 33/158$ , 21%). The SML and literature review yielded fewer visual functioning symptoms (e.g., blurry vision:  $n = 2/7$  and  $n = 1/7$ , respectively) compared to the qualitative interviews ( $n = 7/7$ ). SML identified 5/9 secondary symptoms (e.g., headaches) whereas the literature review did not identify any. Proximal functional impacts (e.g., seeing objects in near vision, reading small print and daily living impacts) were almost all identified in qualitative interviews ( $n = 41/42$ , 98%) but less frequently through SML ( $n = 21/42$ , 50%) and the literature review ( $n = 13/42$ , 31%). SML identified more concepts related to distal impacts on quality of life (e.g., emotional, social and work concepts), but fewer impacts of correction aids, than the literature review. Interviews provided more in-depth exploration of subconcepts. **Conclusion:** Qualitative interviews identified more concepts and explored them in more depth than SML and reviewing literature. However, SML and literature reviewing are quicker, more cost-effective, and may provide early identification of relevant concepts to explore through interviews. Findings may differ for conditions with more qualitative literature/social media discussion. The resulting conceptual model will help determine which PROs cover significant patient experiences relevant for treatment outcomes.

## **111: Clinical Applications**

### **(111.1) Leveraging EPIC to Collect PROMIS CAT scores for Orthopedic Surgery Patients to Create Blueprint for System-wide Expansion**

Heidi McCoy, Master's in Science, Michigan Medicine, Ann Arbor, Michigan, United States; Ian Perry, Bachelor's of Industrial Operational Engineering, Michigan Medicine, Ann Arbor, Michigan, United States; Paul Talusan, Doctorate of Medicine, Michigan Medicine, Ann Arbor, Michigan, United States; Steve Bernstein, MD, MPH, Michigan Medicine, Ann Arbor, Michigan, United States; Amber Lopez, Michigan Medicine, Ann Arbor, Michigan, United States; Janet Rodriguez, Michigan Medicine, Ann Arbor, Michigan, United States

**Aims:** While implementing PROMIS CAT questionnaires has tremendous value for both clinicians and patients alike, the process to make it happen can be very daunting. Where to get started? What challenges can you expect? What are key elements to making implementation successful? These questions may arise and could derail the effort before it begins. We would like this opportunity to share Michigan Medicine's approach to implementation, so others may learn from our challenges and victories on this worthwhile journey. **Methods:** Michigan Medicine's Clinical Design and Innovation (CDI) team, division under the Quality Department, collaborated with three Orthopedic Surgery clinics to participate in a sequential pilot rollout to implement three PROMIS CAT questionnaires (physical function, pain interference, and depression) to pave the way for the entire organization. A project manager and industrial engineer from CDI partnered with physician champions, process owners, an EPIC-certified application coordinator, and clinic staff to create sustainable processes before, during, and after implementation. Three areas of focus were identified as key components to ensure a successful implementation; patient population selection, tablet workflow, and standardized communication. Multidisciplinary team members worked together every other week, over 7 months' time, to develop and test questionnaire completion processes to promote data-driven physician/patient conversations. **Results:** South Main Orthopedics, foot & ankle division, successfully launched and sustained the first wave of the pilot plan. Between 1/23/2019 and 1/23/2020, 8889 patient visits were assigned PROMIS CAT questionnaires. Of the two completion methods, 41% ( $n = 3648$ ) completed on the EPIC patient portal and 24% ( $n = 2133$ ) completed by MA in exam room. The remaining 35% ( $n = 3108$ ) were assigned and not completed, which was more than the 20% initially anticipated. **Conclusion:** Achieving 65% completion rate was just the beginning. Wave two, coming Fall 2020, launching WELCOME, an EPIC software enhancement, will support achieving nearly 100% completion. From the challenges and triumphs of these two pilot waves, a strategic scaling approach is in progress, connecting with Ophthalmology and Rheumatology as two potential areas of expansion.

### **(111.2) Mechanisms of change in patients with medically unexplained physical symptoms: a multi-site naturalistic study**

Jan Boehnke, School of Health Sciences, University of Dundee, Dundee, United Kingdom; Tomáš Řiháček, Department of Psychology, Faculty of Social Studies, Masaryk University, Brno, Czech Republic; Martina Pourová, Department of Psychology, Faculty of Social Studies, Masaryk University, Brno, Czech Republic

**Aims:** Process-outcome research in mental health settings investigates which therapeutic interventions work for whom and under which circumstances. Several psychological change mechanisms have been suggested to inform the treatment of patients suffering from medically unexplained physical symptoms (MUPS; see methods below). This study aimed to test the association of such mechanisms with therapeutic outcomes in a naturalistic sample of patients taking part in intensive group treatments. **Methods:** Across seven clinical sites providing multimodal group psychotherapy, weekly data on  $n = 291$  patients living with MUPS were gathered (72% female;  $M = 40.5$  years,  $SD = 11.1$ ). The diagnosis of MUPS was established based on the triangulation of patient self-report and expert evaluation. The target patient-reported outcomes were somatic symptoms (as measured by the Patient Health Questionnaire, PHQ-15) and mental distress (as measured by the Outcome Rating Scale, ORS). Based on a previous systematic reviews, the following proposed change mechanisms were assessed: somatic awareness, emotional regulation skills, acceptance of symptoms, satisfaction of a patient's relational needs, clarification of meaning, quality of the therapeutic alliance, and the quality of group cohesion. We used multilevel modeling to test whether these proposed mechanisms predict a time-lagged change in outcome. **Results:** The final assessments and a medical chart review for diagnoses have been finished. Preliminary results indicate that pre-post effect sizes are in line with expectations for naturalistic treatment settings (PHQ-15  $d = 0.42$ ; ORS  $d = 1.04$ ). Additionally, the therapeutic setting was successful in evoking the measured mechanisms. Most importantly, while these mechanisms predicted change in mental distress, they were largely unrelated to change in somatic symptoms. Subsequent analyses indicate the latter interacted with each other in a circular manner: improvement in somatic symptoms predicted time-lagged improvement in mental distress and vice versa. **Conclusion:** This is to our knowledge the first study investigating processes and outcomes for patients living with MUPS in the Czech Republic. The findings inform treatment strategies for patients with MUPS and support in particular that a change in somatic symptoms can be achieved indirectly by targeting patients' mental distress through psychotherapy.

### (111.3) Health-related quality of life and caregiver burden among relatives of out-of-hospital cardiac arrest survivors

Mattias Bohm, Department of Clinical Sciences, Neurology, Lund University, Skåne University Hospital, Malmö, Sweden, Malmö, Sweden; Tobias Cronberg, Department of Clinical Sciences, Neurology, Lund University, Skåne University Hospital, Lund, Sweden, Lund, Sweden; Hans Friberg, Department of Clinical Sciences, Anesthesiology and Intensive Care, Lund University, Skåne University Hospital, Malmö, Sweden, Malmö, Sweden; Niklas Nielsen, Department of Clinical Sciences Lund, Anesthesiology and Intensive Care, Lund University, Helsingborg Hospital, Helsingborg, Sweden, Helsingborg, Sweden; Susann Ullén, Clinical Studies Sweden, Forum South, Skåne University Hospital, Lund, Sweden, Lund, Sweden; Johan Undén, Department of Clinical Sciences Lund, Anesthesiology and Intensive Care, Lund University, Halmstad Hospital, Halmstad, Sweden, Lund, Sweden; Kristofer Årestedt, Linnaeus University, Faculty of Health and Life Sciences, Kalmar, Sweden, Kalmar, Sweden; Gisela Lilja, Department of Clinical Sciences, Neurology, Lund University, Skåne University Hospital, Lund, Sweden, Lund, Sweden

**Aims:** This study aimed to describe health-related quality of life (HRQoL) and caregiver burden among relatives of OHCA-survivors, in relation to cognitive impairments of the OHCA-survivors. Further, relatives' HRQoL and caregiver burden were compared with relatives of an ST-elevation myocardial infarction (STEMI) control group.

**Methods:** Data were taken from the cognitive substudy of the Targeted Temperature Management-trial. Face-to-face follow-up 6 months post-event was performed for relatives of 272 OHCA-survivors and 108 STEMI-controls, included at an intended ratio of 2:1. HRQoL was assessed with SF-36v2® and caregiver burden with the 22-item Zarit Burden Interview (ZBI-22). OHCA-survivors were categorized based on the results from cognitive assessments as having "no to mild cognitive impairment" (N-MCI) and "cognitive impairment" (CI). **Results:** The median age of relatives of the OHCA-survivors was 58 (IQR 18) years (83% females and 79% cohabited with the survivor). The overall scores for HRQoL were within average normative levels and there were no significant differences between the relatives of OHCA-survivors and the STEMI-controls (PCS mean 51.3 versus 50.5,  $p = 0.421$ ; MCS mean 48.4 versus 50.2,  $p = 0.085$ , respectively). When stratified for cognitive function, relatives of OHCA survivors with CI ( $n = 126$ ) versus N-MCI reported worse HRQoL in 5 of 8 domains, particularly in the domain of Role Emotional (mean 49.5 versus 45.7,  $p = 0.002$ ,  $ES = -0.19$ , respectively). In general, relatives to both OHCA-survivors and STEMI-controls reported low burden (median 11 versus 9.5,  $p = 0.099$ ). The most frequently reported aspect of burden regarded relatives' fears about further deterioration, where 23% of the relatives of the OHCA-survivors reported that they fear what the future may hold. The relatives of OHCA-survivors with CI reported higher levels of burden compared with N-MCI (median 18 versus 8,  $p < 0.001$ ,  $ES = 0.3$ , respectively), also 40% versus 17% had a score that was above cut-off ( $\geq 21$ ). **Conclusion:** In general, relatives to OHCA-survivors and STEMI-controls reported HRQoL comparable to a general population, and further low levels of burden. However, relatives of cognitively impaired OHCA-survivors report worse HRQoL and increased burden. This study adds important information about the situation for relatives of OHCA-survivors, and the results may be used during follow-up to identify those in need of support.

### (111.4) The creation of Voxe: an electronic patient-reported outcome measure platform to capture pediatric transplant patients' voices and inform clinical care

Samantha Anthony, PhD, The Hospital for Sick Children, Toronto, Ontario, Canada; Sarah Pol, MSc, The Hospital for Sick Children, Toronto, Ontario, Canada; Michael Brudno, PhD, The Hospital for Sick Children, Toronto, Ontario, Canada; Dorin Manase, BSc, The Hospital for Sick Children, Toronto, Ontario, Canada; Rulan Parekh, MD, The Hospital for Sick Children, Toronto, Ontario, Canada; Amanda Silva, BScH, The Hospital for Sick Children, Toronto, Ontario, Canada; Jennifer Stinson, PhD, The Hospital for Sick Children, Toronto, Ontario, Canada

**Aims:** Patient-reported outcome measures (PROMs) are vital to address the burden of disease, engage patients meaningfully, and capture their varied experience. Subjective evaluation of solid organ transplantation from the patient perspective is essential. The objective of this research program is to transform care delivery and improve health outcomes for pediatric transplant patients by implementing PROMs into clinical practice. Informed by the results of previous research—Phases 1 (Systematic Review), 2 (Key Stakeholder Interviews), and 3 (Consensus Workshop)—the aim of this study (Phase 4) is to design and develop an electronic PROM (ePROM) platform—Voxe—that will capture and integrate PROM data into the clinical care of pediatric transplant patients. **Methods:** The 'user-centric' approach, in which end-users (i.e., patients and healthcare providers) are central to the design process has guided the development of Voxe. Study participants include 12 heart, kidney, liver, or lung transplant recipients between 10 and 17 years of age and 12 members of their interdisciplinary healthcare teams. A rapid and iterative testing

methodology has been implemented to ‘test, learn and improve’ Voxe prior to coding and launch. The International Organization for Standardization (ISO) model is being utilized to validate each iteration for success. Results: In order to ensure success, ISO key performance indicators are being benchmarked and tracked. During each of the three iterations that include four transplant recipients and four healthcare providers, objective and subjective standards with metrics of (1) effectiveness—accuracy and completeness with which users achieve specific goals; (2) efficiency—resources used in relation to results achieved; and (3) satisfaction—extent to which the users’ physical, cognitive, and emotional responses that result from the use of Voxe meet the users’ needs and expectations are being collected. Conclusion: The ‘test, learn, and improve’ model will enable objective and subjective metrics from patients and healthcare providers to directly influence how Voxe looks and operates in order to drive adoption and success. Future phases will include usability testing and an implementation effectiveness evaluation of the Voxe ePROM platform. Ultimately, Voxe leverages eHealth technology as an innovative approach to capture and integrate patients’ voices into their care experience.

#### **(111.5) Barriers and benefits to the routine use of electronic patient-reported outcome measures in treating children with life-altering skin conditions**

Zephania Tyack, B Occ Thy (Hons.), PhD, The University of Queensland, Brisbane, Australia; Megan Simons, Queensland Children’s Hospital, Brisbane, Australia; Tania Zappala, Children’s Health Queensland, Brisbane, Australia; Gill Harvey, The University of Adelaide, Brisbane, Australia; Steven McPhail, Queensland University of Technology, Brisbane, Australia; Roy Kimble, Children’s Health Queensland, Brisbane, Australia

Aims: To advance understanding of the barriers and benefits to the use of electronic patient-reported outcome measures (ePROMs) in the routine treatment of children with life-altering skin conditions, from multiple stakeholder perspectives. Methods: Stakeholder groups were children with life-altering conditions (burn scars, infantile hemangiomas and dermatological conditions), receiving treatment at three outpatient clinics at a major metropolitan children’s hospital in Australia; their caregivers and treating clinicians. Data were collected using semi-structured interviews and field notes before and during a pragmatic pilot randomized controlled trial of the implementation of ePROMs of health-related quality of life (HRQoL). Qualtrics was used to administer the ePROMs. Barriers and benefits to the routine completion of ePROMs were mapped to the Consolidated Framework for Implementation Research pre- and post-implementation, for each stakeholder group and clinic. Results: Thirty interviews have been completed (14 children and caregivers, 16 clinicians) and field notes have involved 51 child and caregiver participants. Barriers at a clinic level included: safety and privacy concerns in two busy clinics; completing measures for initial consultations where natural communication was identified as a higher priority in one clinic; and a lack of capacity of the health care team to respond to some issues identified. Barriers to the completion of ePROMs included: a lack of appropriate technology at families’ homes; the need for assistance; and inability to prioritize ePROM completion due to the competing burden of caring for family members during COVID-19. Benefits included: being asked about topics that were important to families but not typically raised in consultations (i.e., financial impact of the condition, sleep); the perception of HRQoL as a relatively ‘safe’ topic; and the high value placed on ePROMs by caregivers who felt they would not typically feel comfortable raising the issues identified. Mapping of the barriers and benefits to the Consolidated Framework for Implementation Research covers leadership engagement and organizational

culture which will be discussed. Conclusion: Diverse methods are needed to overcome barriers and build on benefits to capture information on HRQoL routinely prior to consultations. This includes telephone and face-to-face assisted e-PROM completion, and paper and electronic methods of administering patient-reported outcome measures.

#### **112: Measurement development & evaluation I**

##### **(112.1) Development of the short form for chronic Hepatitis B quality of life instrument (CHBQOL-SF) using Rasch analysis and Delphi method**

Lin Zhu, PhD, Zhejiang University School of Public Health, Hangzhou, China; Li Zhang, MS, Zhejiang University School of Public Health, Hangzhou, China; Yingjing Zheng, PhD student, Zhejiang University School of Public Health, Hangzhou, China; Hongmei Wang, PhD, Department of Social Medicine and Department of Pharmacy of the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

Aims: The CHBQOL instrument was developed cultural-dependently as a specific measure for use in assessing the health-related quality of life of Chinese patients with chronic hepatitis B (CHB) and validated with classical psychometric methods. The aim of this study was to further refine the 23-item instrument using Rasch model analysis and Delphi method. Methods: A secondary data analysis was conducted on a sample of 578 CHB patients recruited from six hospitals. Item analysis with partial credit model was performed using RUMM2030 software on each domain of the CHBQOL instrument separately. The assessment included the evaluations of individual item fit, threshold ordering and differential item functioning (DIF), where the poorly performing items were identified and considered to be removed. In addition, the experts’ scores of the item’s importance collected by Delphi method were used to select more important items from the professional perspective. Results: The principal component analysis showed the four domains of CHBQOL were unidimensional. Disordered thresholds were initially found on 4 out of 6 items in Somatic symptoms domain, 1 out of 6 items in Emotional symptoms domain, 0 out of 2 items in Belief domain and 5 out of 9 items in Social stigma domain. Uniform DIF was observed for 4 items for age group, 2 items for gender and 1 item for different ALT levels. The results of Delphi method also suggested 6 items to be eliminated. The final CHBQOL-SF questionnaire with a total of 10 items retained the four dimension structure of the original instrument. The person separation index (PSI) of 0.76 showed a good reliability. The absolute values of individual item fit residuals were all smaller than 2.5 and there was no inverse threshold for each item, indicating that the items fit the Rasch model well and response options were set reasonably. Conclusion: The 10-item CHBQOL-SF questionnaire would reduce the measurement burden and offers an alternative to disease-specific self-reported outcome measures in clinical practice. However, its full psychometric properties and equivalence with the original instrument remain to be further examined in an independent sample. (This study was supported by National Natural Science Foundation of China, Grant Number 71573226.)

##### **(112.2) Development of a new outcome measure for use in forensic mental health services**

Howard Ryland, University of Oxford, Oxford, United Kingdom, United Kingdom; Raymond Fitzpatrick, University of Oxford, Oxford, United Kingdom; Jonathan Cook, University of Oxford,

Oxford, United Kingdom; Seena Fazel, University of Oxford, Oxford, United Kingdom

**Aims:** Systematic reviews have identified a lack of outcome measures that adequately capture the patient perspective in forensic mental health services. This presentation summarizes the development of a new outcome measure for use in these services, which is relevant to a range of stakeholders (including both patients and clinicians). The measure was designed to be quick and simple to use, in a way that is appropriate for routine clinical practice. **Methods:** A framework of candidate items was initially generated through thematic analysis of transcripts from interviews with patients and focus groups with multiple stakeholders. A process of prioritization was achieved through a two-round Delphi process, with stakeholders participating both directly online and via a researcher. Four consensus meetings were held with different combinations of stakeholders, including patients, clinicians, carers and commissioners. Delphi process results were discussed, to guide the research team in the development of a first draft of the new outcome measure. Further input was obtained from a dedicated patient and public advisory group. The patient-reported scale of the new outcome measure underwent two rounds of cognitive interviews with patients. Comments on the clinician-reported scale were obtained from a multidisciplinary team. The research team utilized this feedback to determine the final version of the measure. **Results:** A framework of 42 outcome statements across 6 domains was generated from the thematic analysis. 1 further statement was added to the second round of the Delphi process from participant suggestions. 8 out of the top 15 statements overlapped between the two stakeholder groups in the Delphi process of (1) patients and carers and (2) professionals. The iterative review process resulted in significant modifications to the initial draft scales, including the number of items, user instructions, response options and presentation. The final draft measure contained 20 items in the patient-reported scale and 23 in the clinician scale. **Conclusion:** A new outcome measure for use in forensic mental health services was developed using a multistage process. Further piloting is planned within this population to gain more information about its psychometric properties and guide additional refinement of the measure.

#### **(112.3) Extending PROMIS to children aged 1–5: a new suite of 12 measures for assessing pediatric health**

Jin-Shei Lai, Northwestern University, Chicago, Illinois, United States; Michael Kallen, Northwestern University, Chicago, Illinois, United States; David Cella, Northwestern University, Chicago, Illinois, United States; Courtney K. Blackwell, Northwestern University, Chicago, Illinois, United States; Lauren S Wakschlag, Northwestern University, Chicago, Illinois, United States; Richard Gershon, Northwestern University, Chicago, Illinois, United States

**Aims:** The Patient-Reported Outcomes Measurement Information System (PROMIS), funded by the US National Institutes of Health, includes over 300 measures of physical, mental, and social health for use with individuals age 5 and older. New evidence suggests the early expression of lifespan health and disease states can often be detected in early childhood. Therefore, assessments are needed that are developmentally appropriate, lifespan coherent, and universally applicable to children of all ages. This project aims to extend current PROMIS measures to children aged 1–5. Referencing the current PROMIS framework, and with input from experts and parents of 1–5 year-old children, we identified 12 domains important for assessing younger children. Domain items were constructed using the PROMIS methodology. This presentation reports the psychometric development of these 12 item banks. **Methods:** Twelve item pools were created: Family Relationship (FR), Peer Relationship (PR), Physical Activity (PA),

Sleep Disturbance (SleepDis), Sleep-related Impairment (SleepImp), Self-Regulation (SR), Anger/irritability (Ang), Anxiety (Anx), Depression (Dep), Engagement (Engage), Positive Affect (PosAffect), and Global Health (GH). Two data collection waves were conducted, using parents of children aged 1–5. With Wave-1 data, item pool unidimensionality was evaluated using confirmatory factor analysis (criteria: CFI  $\geq$  0.9; factor loading  $\geq$  0.3; RMSEA  $<$  0.1; residual correlation  $<$  0.15); the graded response model (GRM) was used for fit and parameter estimation. Final item parameters were obtained using multi-group GRM on the combined Wave-1 ( $n = 700$ ) and Wave-2 ( $n = 1057$ ) sample, with Wave 2 as the norming group. **Results:** Wave-1 data analyses supported the unidimensionality of PA, PosAffect, Ang, Anx, Dep, and GH. FR and PR were combined to form a “Relationship” bank; SleepDis and SleepImp were combined to form a “Sleep” bank. SR was divided into “Flexibility” and “Frustration Tolerance”; Engagement was divided into “Curiosity” and “Persistence.” Short forms (4–8 items) for all banks were constructed. **Conclusion:** This study supports the psychometric properties of 12 item banks that can be used by parents of children with aged 1–5. These new item banks were normed on a probability-based sample and can be administered using computerized adaptive testing or by short form. They will be publically available in the near future.

#### **(112.4) A new beginning: improving single-item-per-domain measures of generic functional health and well-being**

John Ware, PhD, John Ware Research Group, Bristol, Rhode Island, United States

**Aims:** Because all surveys begin with the first question, and responses to that question may be enough for many purposes and may determine the next question in CAT surveys, single-item-per-domain (SIPD) improvements are a high priority. The aims of this study were to test three approaches to improving the range and efficiency of SIPD measures of generic health-related quality of life (QOL) domains and to compare their performance in relation to widely used SIPD and multi-item measures of the same domains. **Methods:** Internet surveys were administered to representative samples of US adults ( $n = 4120$ ) and those chronically ill ( $n = 5418$ ), ages 19–97. Generic domain-specific item banks included comparator (SF-36v2; PROMIS-57, SF-8) SIPD and improved Quality of Life General (QGEN) SIPD measures based on: expanded domain content representation; increased response category range; and direct measurement of higher-order domains as opposed to specific symptoms or activities. Domain-specific comparisons of items addressed: face and content validity, descriptive statistics and response distributions ( $\chi^2$  tests of floor/ceiling effects), classical and modern item bank internal consistency criteria, correlations testing same (convergent) and different (discriminant) validity and validity for purpose of estimating higher-order physical (PCS) and mental (MCS) summaries. **Results:** In strong support of their validity, QGEN and comparator SIPD scores consistently correlated highly with same-domain multi-item scales and replicated the hypothesized pattern of correlations with PCS and MCS components. Significant observed differences, which were small, favored new QGEN item approaches. Comparisons between SIPD response distributions showed QGEN reductions ( $p < 0.001$ ) in ceiling effects, in relation to comparator SIPD measures, for six of eight domains and group means equivalent to those for multi-item measures in discriminating across groups differing in disease severity. **Conclusion:** Overall, results showed that new and comparator SIPD measures correlate equally with the same domains, with few small exceptions favoring the improved SIPD measures. Extending the SIPD range increased efficiency and reduced ceiling effects for common functional health and well-being domains and for estimating summary physical and mental component measures. The resulting



8-item, approximately 1-min, improved survey warrants further use and testing as a more efficient beginning or alternative to psychometric and utility surveys of the health domains and states studied.

### (112.5) Development of EQ-5D-5L Portuguese norms

Pedro Ferreira, PhD, CEISUC/FEUC, Coimbra, Portugal; Patrícia Antunes, PhD, CEISUC, Coimbra, Portugal; Lara Ferreira, PhD, CEISUC/ESGHT/CIEO, Faro, Portugal; Luís Pereira, PhD, CEISUC/ESGHT/CIEO, Faro, Portugal

**Aims:** HRQoL measure EQ-5D-5L has recently been validated in Portugal and the corresponding value set derived. However, no population norm data were available. The aim of this study was to accurately estimate the EQ-5D-5L mean index value for Portuguese subpopulations of interest, defined by gender, age group and region. **Methods:** The target population of this study was 8.7 million Portuguese adults, aged 18 and older. We used the stratified random sampling method to select a representative sample of the population. Between November 2015 and January 2016, 1006 individuals were surveyed by a market research company, using a CATI system. Each telephone call lasted 14 min and, after an eligibility check, it encompassed the Portuguese version of EQ-5D-5L, SF-12 and some sociodemographic questions. The quality control and monitoring of survey was conducted both through direct supervision and third-party phone call listening of 10% of the global sample size. **Results:** The majority of respondents were female (53.4%), aged between 30 and 49 years (35.1%), and three-quarters the participants were residents either in Lisbon and Oporto Metropolitan areas (43.3%) or in Northern and Center West coast (31.7%). The majority of respondents were married or living with a partner (57.9%), and 45.8% of the respondents got a low level of education. In terms of occupational status, 51.1% were employed and 48.7% of the respondents lived in a household with 3 or 4 elements. The majority of respondents did not report a chronic disease (53.5%) and 34.6% reported net monthly earnings between €1,000 and €1,999. The general population EQ-5D-5L norm score was 0.887 (index) and 76.0 (VAS). Looking at the Portuguese index, women provided lower scores (0.863) comparing to men (0.914), and the youngest obtained a higher score (0.961) comparing to the older (0.790). In mainland, the South obtained the highest score (0.910) and the Northern and central interior the lowest score (0.865). Single, higher educated and individuals without a chronic disease corresponded to higher scores. **Conclusion:** The obtained norms for the EQ-5D-5L index score may be used as reference values for comparative purposes in health economic studies.

## 113: Patient & clinician engagement

### (113.1) Multifaceted assessment of what matters to patients: Different connotations depend upon individuals' particular health concerns

Bruce Rapkin, Ph.D., Albert Einstein College of Medicine, Bronx, New York, United States

**Aims:** Assessment of what matters most to individuals seeking health care raises particular challenges because of the complexity of patient experience and motivation. Individuals' priorities and behaviors regarding health/healthcare will be influenced by attitudes and beliefs that pertain to specific health issues and episodes, and the nature of their health system encounters. Social and economic problems may impinge on health behavior and access to care. Individuals evaluate their current health states against salient social norms and reference groups, in light of anticipated impact on attaining personal goals. We

tested multifaceted assessment to determine what was most germane to individuals' health and well-being. **Methods:** The 2018 Bronx Community Health Survey included 1877 individuals sampled online and in-person to represent the sociodemographic diversity of the Bronx. Due to time constraints, participants were randomized to either complete the Dynamics of Care (DoC) Assessment or a quality of life (QOL)-Appraisal battery. The former probed individuals' recent health concerns in terms of decisions about seeking assistance, barriers to care, communication with providers, satisfaction with care, and problem resolution. The latter included PROMIS-10 Global Health measure, the Inventory of Urban Stressors, and the Brief Appraisal Inventory. **Results:** The study sample was 74% female, middle-aged, and ethnically diverse. Two-step cluster analysis of recent health concerns identified 8 patient groups. The DoC revealed differences in ways of valuing care depending on specific health needs, whereas the QOL Appraisal battery revealed group differences in the influence of social determinants and appraisal on overall QOL. Satisfaction with care and QOL were more strongly associated with family impact, stigmatizing comparisons, and norms for people dealing with psychological concerns and physical disabilities. Among individuals seeking to identify and prevent future health risks, resolution of specific concerns was strongly associated with satisfaction with providers, and not with social determinants and appraisal. **Conclusion:** The question, "What Matters to You?" has different connotations, depending upon individuals' particular health concerns. The two measurement approaches enabled examination of the dimensions of experience that matter most to different patients in different circumstances, and will likely benefit efforts to improve health communications and patient engagement along the continuum of care.

### (113.2) Providing research participants with information about their health: Results from burn survivor focus groups

Dagmar Amtmann, PhD, University of Washington, Seattle, Washington, United States; Kara McMullen, MPH, University of Washington, Seattle, Washington, United States; Alyssa M Bamer, MPH, University of Washington, Seattle, Washington, United States; Nicole Gibran, MD, University of Washington, Seattle, Washington, United States; Radha K Holavanahalli, PhD, University of Texas Southwestern, Dallas, Texas, United States; Jeffrey C Schneider, MD, Spaulding Rehabilitation Hospital, Dallas, Massachusetts, United States; Audrey Wolfe, MPH, Spaulding Rehabilitation Hospital, Boston, Massachusetts, United States; Adriana Sanchez, MA, University of Texas Southwestern, Dallas, Texas, United States

**Aims:** Participants often receive little to no feedback after participating in a research study even though they respond to numerous health instruments. We sought to better understand what information burn survivors might need, what formats might be most useful and what concerns about feedback they might have. **Methods:** Adult burn survivors and caregivers/partners participated in focus groups. Multiple formats of reports on health domains (e.g., pain, depression) were discussed. We explored whether participants wanted to receive a summary based on their responses, what should be on the summary and ways to handle reports that indicate problems (e.g., high depressive symptoms). Audio recordings were transcribed, anonymized and summarized. **Results:** A total of 11 burn survivors and 4 caregivers/partners participated in three focus groups at different locations in the US. Average age of the survivors was 49 years, 62% were male and 71% were white. All except one participant wanted to receive reports about their health after study participation, regardless of whether the feedback was positive or negative. Longitudinal line graphs were too complex; tables were preferred. The most preferred format listed health domains in which the burn survivor was doing

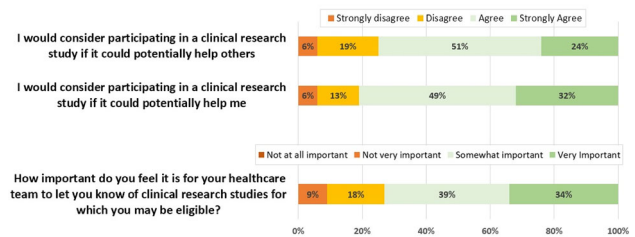
well and areas of concern, with optional links to more details. Survivors found the links to online resources helpful and suggested that information about drugs, alcohol, and PTSD be provided to all. Most would rather receive the potentially negative reports than not, but suggested that feedback should be provided no sooner than 1-year post injury. Most would share the results with their care partners and only a few would share it with care providers. Conclusion: Burn survivors are very interested in receiving summary reports of their responses to research surveys. Simple displays and messaging are essential for the reports to be useful. More research is needed to evaluate whether the reports can aid burn survivors in their recovery. Results: from the focus groups highlight the need for providing information that does not assume participants understand how health instruments are scored and the ability to interpret graphs. Clear guidance for participants on how to address issues identified in their reports is needed.

### (113.3) Patient-entered data as an aid to research recruitment

Irene Katzan, Cleveland Clinic, Cleveland, Ohio, United States; Wilson Tang, MD, Cleveland Clinic, Cleveland, Ohio, United States; Andrew Schuster, BA, Cleveland Clinic, Cleveland, Ohio, United States; Ryan Honomichl, PhD, Cleveland Clinic, Cleveland, Ohio, United States; Brittany Lapin, PhD, Cleveland Clinic, Cleveland, Ohio, United States

**Aims:** The ability to recruit patients into clinical research studies is a major factor in the success of clinical research trials, yet is often a significant challenge. The ability to use patient-reported questions to identify patients who are more likely to participate in clinical research studies could increase recruitment rates. The objectives of our study were to: (1) Develop patient-reported questions that reflect patients' perceptions about participating in clinical research studies; (2) Determine whether patient responses to these questions were predictive of patients' interest in participating in a precision medicine research study. **Methods:** Thirty-minute qualitative one-on-one patients interviews were conducted to develop self-reported questions that would reflect patients' likelihood of participating in a clinical research study. The candidate "research perception" questions identified from this process were added to a patient-reported questionnaire set routinely completed in a primary care clinic and which included the PROMIS Global Health, PHQ depression screen, and 3 social needs questions. As part of a separate online solicitation for research participation, patients also completed a "research recruitment" question regarding their interest in participating in an ongoing research study to identify genetic risk factors for cancer. A multivariable logistic regression model was constructed to evaluate the association of patient responses to the 3 "research perception" questions with a "yes" response to the "research recruitment" question after adjustment for demographics and PROM scores. **Results:** Three candidate "research perception" questions were identified based on findings from 32 qualitative interviews (Figure). Between 8/31/2018 and 4/19/2019, 908 patients (mean age 47.4 years; 59.5% female) completed these "research perception" and "research recruitment" questions and other PROMs. The majority of patients responded positively to the 3 "research perception" questions (Figure). Only one of them—I would consider participating in a clinical research study if it could potentially help others—was independently associated with a "yes" response to the research recruitment question: "Agree" OR 5.8 (95% CI 1.2–34.2), "Strongly Agree" OR = 7.5 (95% CI 1.5–45.0). **Conclusion:** A patient-reported question designed to reflect patients' likelihood of participating in a research study may help identify patients who are more likely to enroll in clinical research studies. Further validation of this approach is warranted.

Figure: Responses to "Research Perceptions" Questions



### (113.4) PRO workshops for multidisciplinary clinicians in home dialysis

Kara Schick-Makaroff, PhD, University of Alberta, Edmonton, Alberta, Canada; Joanna Czupryn, BSc, University of Alberta, Edmonton, Alberta, Canada; Jae-Yung Kwon, PhD(c), University of British Columbia, Vancouver, British Columbia, Canada; Scott Klarenbach, MD, University of Alberta, Edmonton, Alberta, Canada; Robin Cohen, PhD, McGill University, Montreal, Quebec, Canada; Loretta Lee, BScN, Patient Partner, University of Alberta, Montreal, Alberta, Canada; Elizabeth Kusi-Appiah, MN, University of Alberta, Edmonton, Alberta, Canada; Patricia Silbernagel, MSc, University of Alberta, Edmonton, Alberta, Canada; Rita Iradukunda, BScN, Alberta Health Services, Edmonton, Alberta, Canada; Richard Sawatzky, PhD, Trinity Western University, Langley, British Columbia, Canada; PhD, Trinity Western University, Langley, British Columbia, Canada

**Aims:** Our purpose was to understand the educational needs of a multidisciplinary team in home dialysis, develop and deliver workshops to support routine utilization of patient-reported outcomes (PROs). **Methods:** We developed PRO workshops for clinicians, informed by qualitative data from patients and clinicians and Bloom's Taxonomy, and then compared clinicians' perspectives on use of PROs with a pre-post test. Workshop development involved nurses, physicians, dietitians, social workers, and people on peritoneal or home hemodialysis at one large urban home dialysis clinic in Western Canada where PROs were collected as part of standard care. An interpretive description approach was used, and data were collected through six clinician and patient focus groups and interviews ( $n = 63$ ). Participants were asked about their current use of PROs, barriers in use, how PROs could be utilized, and areas they needed support. A series of 4 workshops were offered 12 times. Forty-one clinicians attended  $\geq 1$  workshop, of which 40 completed pre-post evaluation questionnaires. **Results:** Neither patients nor clinicians had previous training on how to use PROs and interpret results. The workshops addressed four areas of educational need: (1) PRO use and interpretation in practice (introduction of PROs to patients, workflow, interpretation of scores). (2) Patients valuing of, and relationship to, the use of PROs in their own care. (3) Strategies for PROs to support communication/coordination within the team (clinicians, patients, referrals). (4) Routine integration of PROs as a fundamental change to practice. Pre-post comparisons indicate that 35% of clinicians reported an increase in looking at PRO responses, 57% reported improvement in one's skill in explaining PRO completion to a patient, and 38% reported greater competence in follow-up. However, 33% of clinicians found less importance in PROs being used by kidney programs or practitioners, 25% found less enhancement of person-centered care with routine PRO use, and 35% reported a decrease in their responsibility to respond to PRO-identified results. **Conclusion:** Workshop development with stakeholders was a strength to the approach. Triangulation of data confirmed clinicians' uncertainty about use of PROs within their multidisciplinary team and how to approach PRO responses considered "out of scope," even after workshop delivery.

### **(113.5) Engaging patients and caregivers in the selection, development, implementation and interpretation of Clinical Outcome Assessment (COA) instruments: co-creation of practical ‘how to’ guidance**

Laure Delbecq, S.A. Eli Lilly Benelux N.V, Brussels, Belgium; Paola Kruger, EUPATI Italia, Rome, Italy; Ashley Duenas, Evidera, London, United Kingdom; Dominique Hamerlijnck, EUPATI Netherlands, Utrecht, Netherlands; Dagmar Kaschinski, Boehringer Ingelheim International GmbH, Ingelheim, Germany; Chi Pakarinen, Patient Focused Medicines Development, Ingelheim, Belgium; Ify Sargeant, Patient Focused Medicines Development, Brussels, Belgium; Anne-Marie Hamoir, Patient Focused Medicines Development, Brussels, Belgium

**Aims:** Although there is established guidance on Clinical Outcome Assessment (COA) from ISOQOL, FDA, EMA, COSMIN and others, there remains limited specific guidance on how (e.g., interviews, surveys, ad boards, financial compensation, contracting), when (i.e., what stage[s] of the drug development lifecycle) and who to engage (e.g., patient experts, patient representatives, ‘ordinary’ patients, caregivers) in COAs selection, development, implementation and interpretation. We describe application of the Patient Engagement Quality Guidance (PEQG) framework (co-created by Patient Focused Medicines Development [PFMD]) to develop tailored guidance on how to engage patients/caregivers in the COA process. **Methods:** A multi-stakeholder working group (WG) of 22 participants including patient experts, PFMD, pharmaceutical industry, clinical research organizations, and consultancy representatives was convened in February 2019. The WG reviewed published COA guidances and case examples provided by the WG, as well as other relevant publications, then applied the PEQG framework to identify gaps of PE within the COA process. Patient Engagement recommendations were synthesized and iteratively reviewed, followed by a public consultation validation. **Results:** Nine WG working sessions were held (February 2019–March 2020). The WG developed a preliminary draft of PE in COA guidance by applying PEQG to published recommendations and best practices. The preliminary guidance incorporates advice and examples covering PE in [1] identification of the burden of disease and relevant concepts for inclusion in COA questionnaires; [2] COA selection; [3] COA development/revision; [4] COA implementation; and [5] COA interpretation. For each section, the guidance summarizes the relevant process and describes incorporation of PE within the process (i.e., how, when, who). We will present the preliminary guidance and results from ongoing public consultation. **Conclusion:** The PEQG provides a practical and adaptable framework that can be used to co-create tailored PE guidance for specific activities. We have developed preliminary guidance for meaningful PE in COA process which—after further validation through public consultation—will provide a ‘How to’ module specific for PE in COA. Other PFMD working groups are also applying the PEQG to provide additional ‘How To’ modules for further specific PE activities with the aim of facilitating practical implementation of PEQG in diverse scenarios.

## **114: Pediatric Research**

### **(114.1) Associations among patient-reported outcomes measures of physical and psychological functioning and willingness to share social media data for research among adolescents with a chronic rheumatic disease**

Elissa R Weitzman, ScD, MSc, Harvard Medical School | Boston Children’s Hospital, Boston, Massachusetts, United States; Alexandra C Marin, PhD, Boston Children’s Hospital, Harvard Medical School, Boston, Massachusetts, United States; Rachele Cox, MPH, Boston

Children’s Hospital, Boston, Massachusetts, United States; Machiko Minegishi, MD MPH, Boston Children’s Hospital, Boston, Massachusetts, United States; Kara M Magane, MS, Boston Children’s Hospital, Boston, Massachusetts, United States; Lauren E Wisk, PhD, Boston Children’s Hospital, Harvard Medical School, Boston, Massachusetts, United States; submitted on behalf of for the CARRA PEPR Investigators

**Aims:** Adolescent social media (SM) use is ubiquitous. Information gleaned from SM may augment understanding of disease and treatment experiences and quality of life of youth living with a rheumatic disease (RD). Little is known about what whether youth with RD will share their SM for health research, and whether youth who will/will not share differ in health status, which may bias results from SM-derived information. **Methods:** We recruited adolescents in treatment for a rheumatic disease who were members of a US multisite clinical disease registry, collecting from them reports of mobility, pain interference, fatigue, depression, anxiety, and sense of meaning/purpose using Patient-Reported Outcomes Measurement Information System® Pediatric measures administered using computer-adaptive testing. Additionally, youth completed a survey about their SM use and willingness to prospectively share their SM data for health research. We compared PROMIS measures for sharing/non-sharing youth, using descriptive statistics and logistic regression. **Results:** Among  $n = 123$  participants (average age 15.6 years (SD = 1.6), 65.0% female),  $n = 117$  reported using SM of whom 43.6% view/read about other youth with RD, and 25.6% posted about their experience with RD. Of all SM users, 76.1% ( $n = 89$ ) shared their SM data—63.4% of males and 82.9% of females ( $p = 0.019$ ). Compared to non-sharers, the sharing cohort reported on average lower mobility (49.3 versus 54.2), greater pain interference (45.7 versus 39.6), more fatigue (48.8 versus 39.5), more depression (48.1 versus 42.0), and greater anxiety (45.1 versus 38.4) (all  $p$  values < 0.05). Higher levels of these factors were associated with sharing, in regression analyses controlling for age and gender (all  $p$  values < 0.05). **Conclusion:** High percentages of youth living with RD use SM including to read about others’ experiences with RD, while a smaller percentage posts about their RD. Most users shared access to their SM for research. The sharing cohort reported worse health than their non-sharing peers, across a range of measures. SM may offer a potent information source and engagement pathway for youth with RD, but differences between sharing/non-sharing cohorts merit consideration when designing studies and evaluating SM-derived findings.

### **(114.2) Creating meaningful outcome measurement strategies in rare diseases: learnings from Duchenne Muscular Dystrophy and Spinal Muscular Atrophy**

Hannah Staunton, Roche Products Ltd, Welwyn Garden City, United Kingdom; Pat Furlong, Parent Project Muscular Dystrophy, Hackensack, New Jersey, United States; Ryan Fischer, Parent Project Muscular Dystrophy, Hackensack, New Jersey, United States; Rosangel Cruz, Cure SMA, Elk Grove Village, Illinois, United States; Nicole Gusset, SMA Schweiz, Swiss Patient Organisation for Spinal Muscular Atrophy, Heimberg, Switzerland; SMA Europe, Freiburg, Germany, Heimberg, Switzerland; Inge Schwensen, SMA Europe and Deutsche Gesellschaft fuer Muskelkranke (DGM), Heimberg, Germany; Mencia de Lemus, SMA Europe, Freiburg, Germany and Fundación Atrofia Muscular Espinal, FundAME, Spain, Madrid, Spain; Fani Petridis, F. Hoffman-La Roche, Basel, Switzerland; Elena Zhuravleva, F. Hoffman-La Roche, Basel, Switzerland; Jessica Braid, Roche Products Ltd, Welwyn Garden City, United Kingdom; Dylan Trundell, Roche Products Ltd, Welwyn Garden City, United Kingdom; Dylan Trundell, Roche Products Ltd, Welwyn Garden City, United Kingdom

**Aims:** In rare diseases, there are challenges relating to disease heterogeneity, optimal sample sizes for qualitative research, and the identification of suitable outcome measures. In Duchenne Muscular Dystrophy (DMD) and Spinal Muscular Atrophy (SMA), there is a multitude of motor function and health-related quality of life measures that are widely used in clinical practice and clinical trials. Outlined here are key learnings from qualitative studies associated with the development of outcome measurement strategies in DMD and SMA. **Methods:** A series of qualitative studies including interviews, focus groups and surveys led to the creation of conceptual models in DMD and SMA, the creation of novel outcomes (a set of global impression items in DMD, and a novel patient and caregiver reported outcome called the SMA Independence scale), and assessment of existing motor function scales. Data were analyzed according to content and thematic analysis methods. **Results:** These studies provided valuable lessons for developing outcome measurement strategies in rare disease, including: Existing qualitative research in rare diseases can be limited. Multiple qualitative approaches are needed to allow the generation of a clinically and geographically diverse sample (e.g., literature sources, interviews, surveys and focus group data) to inform conceptual model development and outcome selection. Given the heterogeneity of disease experience, it is important to identify concepts of relevance to a diverse patient population (e.g., independence). Clinical development programs may be accelerated in rare diseases and therefore pragmatism is needed in adapting existing measures (e.g., adapting global impression items to be disease specific). Early and systematic involvement of patient organizations is necessary to facilitate incorporation of the patient voice into the holistic measurement strategy. This includes design, conduct and interpretation of results. These interactions need to start early in the drug lifecycle to ensure sufficient time for incorporation of insights to inform the overall outcome assessment strategy. **Conclusion:** Patient and caregiver input is critical in the development and selection of outcome assessments, in order to ensure the assessment of meaningful concepts which meet regulatory and health technology assessment standards. The learnings described here can serve as a methodological framework for qualitative research in rare diseases.

#### **(114.3) Agreement between child-self-report with caregiver-report and clinician-report for symptoms in pediatric cancer treatment**

Molly McFatrigh, MPH, Duke University School of Medicine, Durham, North Carolina, United States; David R. Freyer, DO, MS, Children's Hospital Los Angeles, Los Angeles, California, United States; Li Lin, MS, Duke University School of Medicine, Durham, North Carolina, United States; Pamela S. Hinds, PhD, RN, FAAN, Children's National Hospital, Washington, District Of Columbia, United States; Nicole R. Lucas, BS, Duke University School of Medicine, Durham, North Carolina, United States; Janice S. Withycombe, PhD, RN, MN, Clemson University School of Nursing, Durham, South Carolina, United States; Mia K. Waldron, MSN-Ed, RN-BC, CPN, Children's National Hospital, Washington, D.C., United States; Katie R. Villabroza, Children's Hospital Los Angeles, Los Angeles, California, United States; Deborah Tomlinson, MN, RN, The Hospital for Sick Children, Toronto, Ontario, Canada; Scott H. Maurer, MD, UPMC Children's Hospital of Pittsburgh, Pittsburgh, Pennsylvania, United States; Jennifer W. Mack, MD, UPMC Children's Hospital of Pittsburgh, Pittsburgh, Pennsylvania, United States; Jennifer W. Mack, MD, MPH, Dana-Farber Cancer Institute, Boston, Massachusetts, United States; Shana S. Jacobs, MD, Children's National Hospital, Washington, D.C., United States; Justin N. Baker, MD, FAAP, FAAHPM, St. Jude Children's Research Hospital,

Memphis, Tennessee, United States; Bryce B. Reeve, PhD, Duke University School of Medicine, Durham, North Carolina, United States

**Aims:** The clinician report of symptom adverse events (AE) is the standard in pediatric oncology clinical trials, despite the subjective nature of many symptoms such as fatigue. This study examines the agreement between the child and caregiver-report of symptoms on the Pediatric Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (Ped-PRO-CTCAE) for children undergoing cancer treatment, as well as the child's agreement with clinician's grading of the CTCAE. Additionally, we identify factors (e.g., child demographics) associated with better agreement. **Methods:** Participants, 7–18 years of age and undergoing cancer treatment and their caregiver completed the Ped-PRO-CTCAE (child and caregiver versions) and clinicians completed the CTCAE prior to treatment initiation (T1) and at follow-up (T2) approximately 7–17 days later for children receiving chemotherapy, and 4 + weeks later for those receiving radiation. Polychoric correlations were used to assess association and 95% confidence intervals around the means for each group to determine statistical significance. We used multivariable mixed effect models to identify factors associated with discrepancies in symptom AE reporting. **Results:** Four hundred eight-two child-caregiver dyads participated. The sample was diverse in terms of age, race/ethnicity and cancer type. At T2, correlations between child and caregiver-report ranged from  $r = 0.80$  (vomiting frequency) to 0.49 (fatigue interference). Mean scores differed significantly on symptoms such as nausea, fatigue and sadness, with the caregiver reporting consistently higher scores, indicating higher perception of symptom burden. Correlations between child and clinician-report ranged from 0.77 (cough frequency) to 0.26 (fatigue interference), and 345 child-clinician dyads participated. The child consistently reported higher mean scores than the clinician at T1 and T2, and was significantly different on symptoms such as fatigue, pain, and sadness. Factors that contributed to differences in agreement such as child age and caregiver health status will be reported at the ISOQOL conference. **Conclusion:** This study is one of the largest of its kind to collect symptom AE data from clinicians alongside child self-report and caregiver-report symptom data. We found caregivers overestimate and clinicians underestimate symptom frequency, severity and interference when compared to children themselves. Overall, our findings highlight the importance of the child's voice in oncology treatment.

#### **(114.4) Mothers and fathers provide different reports on aspects of the pediatric patient experience**

Paul Fairie, PhD, University of Calgary, Calgary, Alberta, Canada; Brian Steele, BA (Hon), University of Calgary, Calgary, Alberta, Canada; Kyle Kemp, MSc, University of Calgary, Calgary, Alberta, Canada; Maria Santana, PhD, University of Calgary, Calgary, Alberta, Canada

**Aims:** Measures of the pediatric patient experience often rely on surveys of parents and caregivers. While parents are excellent sources of information about their own children, it is important to understand how parental features can influence how the pediatric patient experiences are reported. Research shows that mothers and fathers differ in their relationships with their children, and the purpose of this study was to determine if patient experience survey results differ systematically between mothers and fathers. **Methods:** Caregivers (primarily parents) completed the Child-Hospital Consumer Assessment of Healthcare Providers and Systems (Child-HCAHPS) survey by telephone within 6 weeks of hospital discharge in Alberta, Canada. Surveys were subsequently linked with electronic medical records. We examined 46 patient experience measures including overall ratings as well as ratings of specific aspects of the hospitalization (such

as communication with providers or quality of the physical environment), and compared the responses of mothers to fathers. Results: A total of 7951 surveys were completed, with the large majority having been completed by mothers ( $n = 6770$ ) rather than fathers ( $n = 898$ ), with the remainder filled out by non-parent caregivers. Comparing the results of mothers to fathers, fathers rated the overall hospital experience more highly (8.9 out of 10 vs 8.7,  $p = 0.001$ ), felt more comfortable with the explanations provided by hospital staff (about medication, discharge, and other areas), and rated the environment (quietness, cleanliness, and availability of toys) of the hospitals more positively. Mothers and fathers did not differ on most ratings of provider communication. Conclusion: Mothers and fathers differ in their reports of their child's care. These findings can help analysts interpret survey results, especially when different mixes of mothers and fathers respond. The results align with earlier research on adult care experiences, which suggests that a gendered component of perceptions of care exists.

#### (114.5) Discrepancy in child self-reported and parent-proxy-reported vision-related quality of life and functional vision of visually impaired children and young people aged 8–17 years

Val Tadic, PhD, University of Greenwich, London, United Kingdom; Alexandra Robertson, PhD, UCL Great Ormond Street Institute of Child Health, London, United Kingdom; Mario Cortina-Borja, PhD, UCL Great Ormond Street Institute of Child Health, London, United Kingdom; Jugnoo Rahi, PhD FRCPH, UCL Great Ormond Street Institute of Child Health, London, United Kingdom

**Aims:** Ratings by children and their parents using child-reported outcome measures are often discordant. We investigated extent of, and factors associated with, agreement between children/young people with visual impairment (VI) and their parents using our two novel questionnaires measuring vision-related quality of life (VQoL)- the VQoL\_CYP and functional vision (FV)- the FVQ\_CYP. **Methods:** 152 children/young people aged 7-18 years with isolated VI (WHO criteria), and their parents, were recruited from 22 NHS Ophthalmology Departments (UK). Age-appropriate child and parent-proxy versions of the VQoL\_CYP and FVQ\_CYP were administered via post. Scores were calculated and transformed to 0-100 scale. Parent-child agreement, stratified by participants' age, gender and clinical characteristics (severity, timing of onset and stability of VI), was examined using the Bland-Altman (BA) method and intraclass correlation coefficients (ICC). **Results:** 56% children/young people were male; 53% aged 7–12 years; 21% had severe VI or blindness; 82% had early onset and 71% progressive VI. BA indicated a wide range of disagreement, with parents both under and overestimating their child's VQoL (mean-score-difference = 7.7, BA limits of agreement [LOA]: Lower = -12.81 [CI95% = - 14.04, - 11.57]; Upper = 17.33 [CI95% = 16.1, 18.57]), but consistently underestimating their FV (mean-score-difference = - 4.7, LOA: Lower = - 20.18 [CI95% = - 21.45, - 18.91]; Upper = 10.73 [CI95% = 9.46, 12]). ICCs for total sample indicated good parent-child agreement on both VQoL\_CYP (.708 [CI95% = .604, .786]) and FVQ\_CYP (.697 [CI95% = .449, .821]), but this varied by child's characteristics. Agreement was highest for children with more severe VI/blindness on both measures (VQoL ICC = .822 [.639, .914]; FV ICC = .814 [.63, .908]) and those for later (vs. early) VI onset on FV (.868 [.733, .937]). Additionally, agreement on VQoL varied by age, with moderate agreement for younger (595 [432, .721]), but good agreement for older children (764 [.599, .858]). **Conclusion:** Agreement between affected children/young people and their parents as proxy respondents using two complementary, but distinct self-report outcome measures for children/young people with VI varies meaningfully by age and key clinical characteristics. The differences in agreement are sufficient to advocate that self-reporting by

children/young people should remain the 'gold standard.' Where self-reporting by children/young people is not possible (e.g., due to cognitive impairment) parental-proxy reporting may be valuable in the context of severe VI or when assessing functional impact of VI.

## 115: PROs in Cancer Research II

### (115.1) Active use of patient-reported outcomes during chemotherapy or immunotherapy for bladder cancer—preliminary data from a national randomized trial

Gry Assam Taarnhøj, MD PhD, Department of Oncology, Rigshospitalet, Copenhagen, Denmark; Christoffer Johansen, Prof. MD, PhD, DMSc, Department of Oncology, Rigshospitalet, Denmark, Copenhagen, Denmark; Henriette Lindberg, MD, PhD, Department of Oncology, Herlev & Gentofte Hospitals, Herlev, Denmark, Herlev, Denmark; Andreas Carus, MD, PhD, Department of Oncology, Aalborg University Hospital, Aalborg, Denmark, Aalborg, Denmark; Rikke Hedegaard Dahlrot, MD, PhD, Department of Oncology, Odense University Hospital, Odense, Denmark, Odense, Denmark; Helle Pappot, Prof., MD, DMSc, Department of Oncology, University of Copenhagen, Rigshospitalet, Denmark, Odense, Denmark

**Aims:** The active use of patient-reported outcomes during cancer treatment has the potential to improve clinical outcomes. This is especially important for bladder cancer patients for whom prognosis is poor and comorbidities trouble completion of treatment. These patients need extended supportive care to improve clinical outcomes. The aim of this study is to test the impact of active use of PROs in the bladder cancer population during chemo- or immunotherapy. **Methods:** This randomized study at four university hospitals in Denmark started enrollment in January 2019 and continues until 230 patients are enrolled, expected in October 2020. All Danish speaking patients with urothelial cell carcinoma of the urinary tract initiating standard treatment with chemo- or immunotherapy and access to electronic communication with health authorities are eligible. Patients are allocated 1:1 to the intervention or control arm. The intervention arm receives weekly electronic PRO-CTCAE questionnaires at home with built-in alerts to patients if a given symptom exceeds the predefined threshold. Clinicians view PRO reports at every clinical visit in the hospital. Co-primary endpoints are completion of treatment and hospital admissions. Secondary endpoints are quality of life, overall survival, and dose reductions. Endpoints will be tested with Fisher's exact test, multivariate linear regression models and Kaplan-Meier survival analysis. **Results:** As of 04.02.2020 131 patients have been enrolled with characteristics according to Fig. 1. Data will be updated at time of the ISOQOL conference including patient adherence data. **Conclusion:** This study will evaluate the impact of the active use of PROs in the bladder cancer population receiving chemo- or immunotherapy with the aim of improving clinical outcomes.

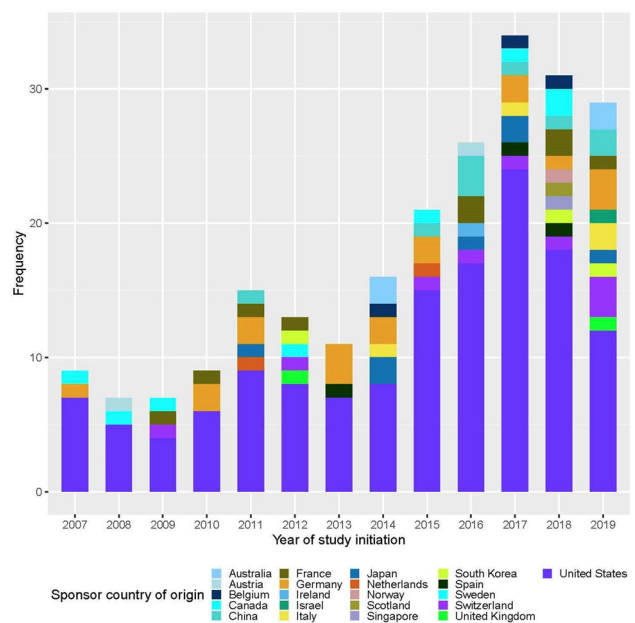
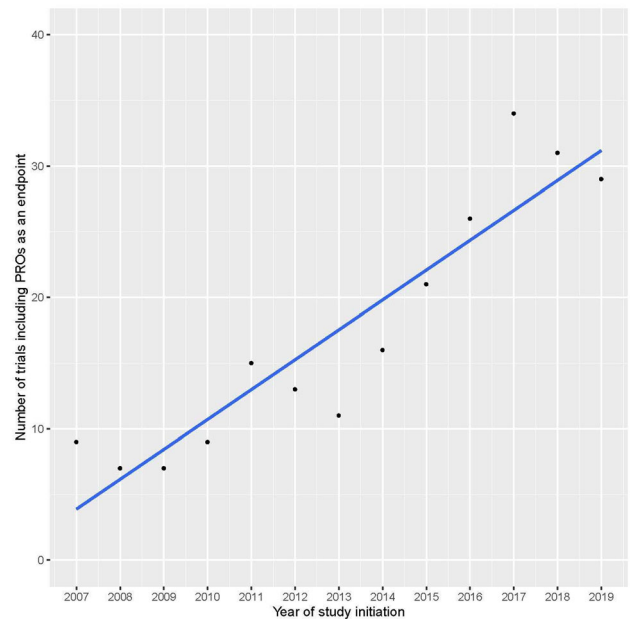
Figure 1. Clinical characteristics of enrolled patients, n=131.

Clinical data	All patients N=131	Intervention arm N=75	Control arm N=56
Gender			
Male	101	57	44
Female	30	18	12
Median age (years, range)	69 (40-87)	69 (45-87)	69 (40-86)
Disease stage			
Locally advanced	49	25	24
Metastatic	82	50	32
Treatment			
Cisplatin+gemcitabine	72	42	30
Carboplatin+gemcitabine	17	9	8
Vinflunine	1	1	0
Pembrolizumab	41	23	18

### (115.2) Trends in patient-reported outcome (PRO) use in early-phase oncology trials

Julia Lai-Kwon, MBBS, Drug Development Unit, The Institute of Cancer Research and the Royal Marsden Hospital, London, United Kingdom; Zhulin Yin, Clinical Trials and Statistics Unit, The Institute of Cancer Research, London, United Kingdom; Anna Minchom, MD, Drug Development Unit, The Institute of Cancer Research and the Royal Marsden Hospital, London, United Kingdom; Christina Yap, PhD, Clinical Trials and Statistics Unit, The Institute of Cancer Research, London, United Kingdom

**Aims:** Early-phase oncology trials establish the safety and tolerability of novel anti-cancer agents. However, clinician-assessed toxicity gradings may miss up to half of adverse events compared to patient-reported events, leading to an incomplete picture of a drug's tolerability. There is growing interest in PROs to enhance toxicity reporting and improve patient representation in drug development. However, little is known about PRO use in this setting. The aim of this study was to describe trends and characteristics of PRO use in early-phase oncology trials. **Methods:** Trials with a dose escalation component registered on ClinicalTrials.gov to commence from 01/01/2007 to 20/01/2020 with 'PROs' or 'health-related quality of life' as an outcome was extracted. Search results were screened to confirm inclusion criteria were met. Study and PRO characteristics were extracted. Descriptive analysis was performed to describe trends in PRO usage. **Results:** 548 studies were identified. 231 (42.2%) were eligible: adult (224, 97%), pediatric (7, 3%), solid tumor (176, 75.9%), hematology (56, 24.1%), seamless phase 1/2 (108, 46.8%). Maximum tolerated dose (MTD) (107, 35%) and safety (95, 31%) were the most common primary endpoints. The majority involved drug combinations (119, 51.5%) and most common therapies were targeted therapy (94, 40.7%), immunotherapy (33, 14.3%) and radiotherapy (33, 14.3%). PRO endpoints were identified in more studies (2.3 increase/year, 95% CI 1.6-2.9, Fig. 1) from an increasing variety of countries (Fig. 2) over time. PROs were typically secondary endpoints (209, 89.7%). The median number of PRO measures was 1 (range: 1-7). PROs were most frequently implemented in the dose escalation phase (114, 49.1%) and phases 1/2 (54, 23.3%). The most commonly used PROs were the EORTC-QLQ-C30 (81, 21.3%) and EQ-5D-5L (19, 5%). **Conclusion:** PRO use has increased significantly over time in a wider variety of settings. This will inform a survey of trialists from academia and industry to assess attitudes towards PROs and their potential to define tolerable doses and regimens. Further methodological work is necessary to determine how to integrate PRO data into traditional early-phase endpoints (e.g., MTD). Guidelines are required to standardize the use and reporting of PROs in early-phase trials to maximize their utilization.



### (115.3) Implementation of the Was It Worth It scale to measure perception of large panel germline genetic testing in cancer patients

Katie Kunze, Ph.D., Mayo Clinic, Scottsdale, Arizona, United States; Michael Golafshar, M.S., Mayo Clinic, Scottsdale, Arizona, United States

States; N. Jewel Samadder, M.D., Mayo Clinic, Phoenix, Arizona, United States; Cindy Azevedo, M.S., Mayo Clinic, Phoenix, Arizona, United States; Douglas Riegert-Johnson, M.D., Mayo Clinic, Jacksonville, Florida, United States; Lisa Boardman, M.D., Mayo Clinic, Jacksonville, Minnesota, United States; Deborah Rhodes, M.D., Mayo Clinic, Rochester, Minnesota, United States; Myra Wick, M.D., Ph.D., Mayo Clinic, Rochester, Minnesota, United States; Scott Okuno, M.D., Mayo Clinic, Rochester, Minnesota, United States; Keith Stewart, M.B. Ch.B., Mayo Clinic, Rochester, Minnesota, United States; Amylou Dueck, M.B. Ch.B., Mayo Clinic, Rochester, Minnesota, United States; Amylou Dueck, Ph.D., Mayo Clinic, Scottsdale, Arizona, United States

**Aims:** The Was It Worth It questionnaire (WIWI) was developed for Mayo Clinic phase 1 trials (Sloan et al. 2011) and has been adapted for clinical trial applications in cancer and other research areas to collect the patient view of treatment and trial experiences. We adapted the WIWI for assessing cancer patient perceptions of their participation in a large scale germline genetic testing study **Methods:** 2983 patients across three large cancer treatment centers underwent 84-gene panel germline genetic testing to assess whether their cancer was likely due to genetic mutation. Patients were administered patient-reported outcome (PRO) measures including the WIWI to assess their symptoms and experiences. The WIWI was modified to focus on participation in genetic testing and was administered after patients received results. **Results:** were analyzed by patient characteristics of interest. **Results:** The current response rate for the WIWI is 61% (1772/2889). A majority of respondents viewed their study participation positively with 78% stating it was worthwhile to participate, 92% would participate in the genetic evaluation again, and 87% would recommend this evaluation to others. Only 0.5% indicated their quality of life (QOL) “got worse” by participating with most (1567/1772, 89%) stating their QOL stayed the same, and 11% reporting that their QOL improved due to their participation. Roughly 29% of respondents reported that their experience was better than expected. Most (70%) reported their experience was as expected, and only 1% ( $n = 25$ ) reported the experience was worse than expected. 283 individuals provided a free-text response containing a suggested improvement to testing. Chi square tests showed statistically significant ( $p < .05$ ) differences in rates of endorsement by presence of pathogenic mutation, gender, race, ethnicity, cancer type, age, cancer stage, and family history of cancer for many of the WIWI items. **Conclusion:** The WIWI was key in assessing the experience of patients undergoing germline genetic testing. It provided a view of patient perspectives not observed in other administered PRO measures. The WIWI allowed for assessment of differences in patient experience and perception of genetic testing by characteristics of interest. **Results:** of the WIWI highlighted the perceived improvement in patient QOL by participating in germline genetic testing.

#### Was It Worth It (WIWI) Questionnaire

**Participating in genetic evaluation is a personal choice and an individual experience. We would like to get your feedback on your experience with genetic evaluation.**

**Directions:** Please answer each question by circling Y (for yes), N (for no), or U (for uncertain).

Was it worthwhile for you to participate in this genetic evaluation? Y N U

If you had to do it over, would you participate in this genetic evaluation again? Y N U

Would you recommend participating in this genetic evaluation to others? Y N U

Overall, did your quality of life change by participating in this genetic evaluation (circle one response)?

It improved                      It stayed the same                      It got worse

Overall how was your experience of participating in this genetic evaluation (circle one response)?

Better than I expected                      The same as I expected                      Worse than I expected

If there was **one thing** that could have been done to improve your experience in this genetic evaluation, what would it be?

\_\_\_\_\_

\_\_\_\_\_

Would you like to talk to someone about your concerns (circle one response)? Yes No

#### **(115.4) Factors in randomized controlled trials reported to impact the implementation of patient-reported outcome measures (PROMs) into routine care: a systematic review**

Natasha Roberts, PhD, Queensland University of Technology, Kelvin Grove, Australia

**Aims:** This systematic review aimed to evaluate the fidelity of PROMs interventions and the impact of fidelity study outcomes on clinical trial success. The secondary aim was to identify any intrinsic factors associated with the fidelity of PROMs interventions. **Methods:** The systematic review protocol was developed using the Cochrane Effective Practice and Organization of Care (EPOC) Guidelines. It underwent peer review for registration with the PROSPERO International Prospective Register of Systematic Reviews. A search strategy was developed for electronic databases and the grey literature, which was complemented by asking field experts and handsearching reference lists for clinical trials investigating the use of PROMs intervention. Two reviewers were used for screening and data extraction. Implementation science theory was used to identify intrinsic factors contributing to processes and outcomes for patient care. Data were synthesized to identify factors relating to the context, participants, the

intervention and overall study fidelity. Results: The review found that many studies had investigated PROMs interventions in clinical practice using clinical trials designs. Overall, 34 randomized controlled trials were identified, 15 specifically focussing on the oncology setting. The PROMs intervention design was reported to address identified contextual and participant requirements in 28 studies. Despite this, 32 studies mentioned barriers to delivering the intervention as intended and directly impacting the research findings. These included difficulties integrating PROMs due to the local hospital or clinic setting, clinician engagement and patient engagement. Strategies used by research teams to ensure fidelity of intervention delivery included education, coordination of care, support of clinicians and management of technology. Conclusion: Despite the planning and resources available for clinical trials, factors in the context and from participants impacted the fidelity of the PROMs intervention. Research staff facilitated adherence to the study protocol overcome these, but they are not extensively reported. Such strategies can be used to inform implementation of PROMs into routine care in the future.

**(115.5) A pilot randomized trial of online self-monitoring of adverse events during pelvic radiotherapy using eRAPID (electronic patient self-reporting of Adverse-events: Patient Information and aDvice)**

Patricia (Trish) Holch, PhD, Leeds Beckett University, Centre for Psychological Research, Leeds, United Kingdom; Kate Absolom, PhD, University of Leeds, Patient Centered Outcomes Research Group, Leeds, United Kingdom; Andrea Gibson, BSc RN, University of Leeds, Patient Centered Outcomes Research Group, Leeds, United Kingdom; Anthony Murphy, RN, The Chrisite, NHS Trust Manchester, Manchester, United Kingdom; Zoe Rogers, MSc, University of Leeds, Patient Centered Outcomes Research Group, Leeds, United Kingdom; Marie Holmes, MSc, University of Leeds, Patient Centered Outcomes Research Group, Leeds, United Kingdom; Rosemary Peacock, PhD, University of Leeds, Patient Centered Outcomes Research Group, Leeds, United Kingdom; Ann Henry, PhD MD, Leeds Teaching Hospitals NHS Trust, Leeds Cancer Centre, Bexley Wing, Leeds, United Kingdom; Sarah Dickinson, BSc, University of Leeds, Patient Centered Outcomes Research Group, Leeds, United Kingdom; Kevin Franks, MD, Leeds Teaching Hospitals NHS Trust, Leeds Cancer Centre, Bexley Wing, Leeds, United Kingdom; Katrina Walker, MD, Leeds Teaching Hospitals NHS Trust, Leeds Cancer Centre, Bexley Wing, Leeds, United Kingdom; Katrina Walker, MSc, University of Leeds, Clinical Trials Research Unit, Leeds, United Kingdom; Eleanor Hudson, MSc, University of Leeds, Clinical Trials Research Unit, Leeds, United Kingdom; Lucy Mc Parland, MSc, University of Leeds, Clinical Trials Research Unit, Leeds, United Kingdom; Julia Brown, PhD, University of Leeds, Clinical Trials Research Unit, Leeds, United Kingdom; Galina Velikova, PhD MD, University of Leeds, Patient Centered Outcomes Research Group, Leeds, United Kingdom

**Aims:** Pelvic radiotherapy (RT) for cancer is potentially curative increases survival but can result in serious short and longer-term adverse effects (AEs). Strategies to enhance patient monitoring and management of acute problems may improve outcomes. The eRAPID approach uses a secure online system for patients to self-report AEs from home; delivering immediate self-management advice or initiating hospital contact. In a randomized feasibility/pilot study in two major UK sites (Leeds Cancer Centre and The Christie Hospital, Manchester) we aimed to refine the intervention, establish feasibility and recruitment/attrition rates and select a suitable quality of life (QOL) outcome measure for a future definitive trial. **Methods:** A prospective two-center parallel RCT (1:1 intervention (eRAPID) vs usual care (UC)) with repeated-measures and mixed-methods was

conducted. Eligible patients included those undergoing pelvic RT for prostate (RT only) and lower gastrointestinal and gynecological cancers (chemo-radiotherapy). eRAPID participants reported AE from home weekly for 12 weeks and at 18 & 24 weeks. We measured and analyzed descriptively: QOL (FACT-G, EORTC-QLQ-C30), patient engagement (self-efficacy scale, patient activation measure), process of care (hospital records of patient contacts/admissions) and economic measures (EQ5D-5L). Semi-structured interviews with staff and patients were analyzed thematically. Results: Between 2016 and 2018, 502 patients were screened, 228 approached, 167 consented (73.2%) and randomized (83-eRAPID, 84-UC). Sixteen (9.6%) subsequently withdrew. Patient adherence with online reporting was 82% of expected at week 1 and 63% at week 12. Return rates of outcome measures were 99.8% at baseline, 77.8% at 18 & 73.7% at 24 weeks. eRAPID patients reported less deterioration in QOL (FACT-G, QLQ-C30 and EQ5D scores) across the study compared to UC, particularly for chemo-radiotherapy patients. Patient and staff endorsed eRAPID as easy to use and providing beneficial support. Conclusion: This pilot randomized study confirmed eRAPID is acceptable to both patients and staff and recruitment feasible (consent rate of > 70%, withdrawal < 10%; online completions 60–80%). Further we have robust data from which to select a suitable outcome measure for a definitive trial. Outcome measures highlighted a potential QOL benefit in the chemo-radiotherapy groups. A formally powered multicenter trial is currently being developed to fully explore the potential of eRAPID in pelvic radiotherapy.

**116: Response Shift**

**(116.1) Response-shift effects in a Neuromyelitis Optica Spectrum Disorder clinical trial: a novel application of random-effects modeling and equating for small samples**

Carolyn Schwartz, Sc.D., DeltaQuest Foundation, Inc., Concord, Massachusetts, United States; Roland Stark, M.Ed., DeltaQuest Foundation, Inc., Concord, Massachusetts, United States; Brian Stucky, Ph.D., DeltaQuest Foundation, Inc., Los Angeles, California, United States; Yuelin Li, Ph.D., Memorial Sloan Kettering Cancer Center, New York, New York, United States; Bruce Rapkin, Ph.D., Albert Einstein College of Medicine, Bronx, New York, United States

**Aims:** Researchers have long posited that response-shift effects may obfuscate treatment effects but, to our knowledge, no one has yet empirically tested this hypothesis in clinical-trial data using multivariate statistical methods. The present work investigated possible response-shift effects in a recent clinical trial testing a new treatment for Neuromyelitis Optica Spectrum Disorder (NMOSD). This pivotal trial provided impressive support for the drug Eculizumab in preventing relapse (primary outcome) and for the more objective evaluative outcomes, but less strong or null results as the indicators became more subjective. This pattern of results suggests that response-shift effects are present. **Methods:** This secondary analysis utilized data from a randomized, double-blind trial evaluating the impact of Eculizumab in preventing relapses in 143 people with NMOSD. Treatment arm and then relapse status were hypothesized ‘catalysts’ of response shift in two series of analyses. Because the study sample was too small for Oort structural-equation modeling, we devised a “de-constructed” version using random-effects models (REMs). Beginning by testing an omnibus response-shift hypothesis, REMs then elucidate specific response-shift types by focusing on a global outcome (EQ-5D Visual Analogue Scale (VAS)) that is likely subject to response-shift effects. The predictors (SF36<sup>TM</sup>v2 mental and physical component scores (MCS and PCS)) helped us to detect response-shift effects in VAS. We then “back-translated” the VAS into the MCS and PCS scores that would have been observed if



response shift had not been present. Results: The omnibus test revealed treatment- and relapse-related response shifts. REMs revealed recalibration and reconceptualization response-shift effects for treatment, and recalibration, reprioritization, and reconceptualization response-shift effects for relapse. Equating was done using raw scores from the VAS, MCS, and PCS, and for computing scores that removed response-shift effects. Correlation analysis and descriptive displays provided a more comprehensive examination of response-shift effects. Conclusion: This secondary analysis of clinical-trial data revealed that not receiving Eculizumab and, more specifically, the experience of relapse made people change their thinking about QOL. Thus, the QOL impacts of placebo/relapse on mental health in particular were under-estimated by the usual analyses. This novel application of REM and equating provides a small-sample method for better estimating treatment effects in clinical trials.

### **(116.2) Response shift in self-reported depression outcomes during treatment-resistant depression**

Myriam Blanchin, PhD, University of Nantes, Nantes, France; Samuel Bulteau, CHU Nantes and University of Nantes, Nantes, France; Morgane Péré, CHU Nantes, Nantes, France; Anne Sauvaget, CHU Nantes and University of Nantes, Nantes, France; Véronique Sébille, CHU Nantes and University of Nantes, Nantes, France

Aims: Major depressive disorder is known to affect patients' self-referential thoughts. Most psychological treatments for depression aim at influencing self-referential processing and their efficacy is often measured with patient-reported outcomes. Hence, treatment-induced changes in self-perception of a construct as depression over time (i.e., response shift) may impact the self-reported measure of depression and may lead to erroneous conclusions regarding treatment efficacy. It may also be of interest to measure changes in self-perception as it can be a goal of therapy. Response shift analysis seems essential to quantify changes in patients' self-perception and to adequately assess changes in self-reported depression outcomes. This study aimed at investigating response shift in patients with unipolar treatment-resistant depression in a clinical trial comparing the effects of three treatment strategies on depression outcomes. Methods: Response shift was investigated on data from a multicenter randomized trial involving 170 patients receiving either an antidepressant treatment (venlafaxine) or repetitive transcranial magnetic stimulation or both. The Oort's procedure based on structural equation models (SEM) was performed in each group on the depression outcome measured by the dimensions of the Beck Depression Inventory: negative self-reference, performance impairment and sad mood. Treatment effects on depression change were studied in a longitudinal multi-group SEM including all response shift effects previously detected according to group membership. Results: Response shift was only detected in patients receiving antidepressant alone. After 4 weeks of treatment, these patients tended to report higher scores of negative self-reference on average compared to baseline, given similar depression levels over time. This could reflect a treatment-induced increased awareness of negative self-reference. Furthermore, the mean level of depression significantly decreased in all groups but has significantly more decreased in the antidepressant alone group. In contrast, the clinical trial, from which the data originated and in which response shift was not taken into account, had concluded to the same decrease of depression in each group. Conclusion: Response shift effect has led to underestimate the improvement of depression in the antidepressant treatment group and has biased the measure of treatment efficacy in this clinical trial. However, it may also have revealed increased awareness of negative self-reference in these patients.

### **(116.3) "I knew who I was this morning, but I've changed a few times since then": A register-based cohort study on response shift in quality of life following injury**

Ritva Rissanen, PhD, Karolinska Institutet, Stockholm, Sweden; Erik Eriksson, Karolinska Institutet, Stockholm, Sweden; Marie Hasselberg, Professor, Karolinska Institutet, Stockholm, Sweden

Aims: During the past decades, several studies have investigated quality of life following traumatic events with ambiguous and paradoxical results with retrospective pre-injury levels being higher than the general population norms. A plausible explanation for this is the phenomenon of response shift. Hence, the aim of this study was to determine the magnitude of response shift in a self-reported measure of quality of life over time among people who have suffered an injury. Methods: A total of 2512 participants who had suffered an injury and were registered in the LifeGene database were included. By using data collected in LifeGene we had a unique opportunity to study quality of life following injury as the register contained a 'true' preinjury assessment of quality of life, i.e., assessed before the injury occurred. In order to analyze the response shift a "Then-test" was applied, which involves the analysis of the traditional measure of change in quality of life, response shift and actual change. In addition, we assessed response shift by using Structural Equation Models. Results: The results indicate that people with injuries report a larger loss of quality of life than indicated by retrospective pre- and post-measure, which can be explained by that a person re-evaluates the concept of quality of life and what it means to have good quality of life. Preliminary results from the Structural Equation model-analysis support the above-mentioned findings. Final results will be presented. Conclusion: In conclusion, the findings of this study seem to support the notion of response shift in quality of life following injury. Hence, if response shift is not considered, the actual impact of injury on quality of life may be underestimated because the person has re-evaluated what quality of life means and the internal benchmark of it has changed.

### **(116.4) Operationalization and detection of response shift: synthesis, limitations, and research opportunities**

Veronique Sebille, Professor, INSERM U1246 SPHERE—University of Nantes, Nantes, France; Lisa M. Lix, Department of Community Health Sciences, University of Manitoba, Winnipeg, Canada, Winnipeg, Manitoba, Canada; Olawale F. Ayilara, Department of Community Health Sciences, University of Manitoba, Winnipeg, Canada, Winnipeg, Manitoba, Canada; Tolulope T. Sajobi, Department of Community Health Sciences and O'Brien Institute for Public Health, University of Calgary, Calgary, Alberta, Canada, Calgary, Alberta, Canada; A. Cecile J.W. Janssens, Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, USA, Atlanta, Georgia, United States; Richard Sawatzky, School of Nursing, Trinity Western University, Langley, BC, Canada, Atlanta, British Columbia, Canada; Mirjam A. G. Sprangers, Department of Medical Psychology, Amsterdam UMC, Location AMC, Amsterdam, Amsterdam, Netherlands; Mathilde G.E. Verdam, Department of Methodology and Statistics, Institute of Psychology, Leiden University, The Netherlands, Leiden, Netherlands; submitted on behalf of Response Shift—in Sync Working Group

Aims: Over the years, a range of quantitative and qualitative methods to detect response shift (RS) have been developed. These methods differ in their conceptual and operational definitions of RS, underlying assumptions, interpretation and alternative explanations that may account for their results. Our objectives are to: (a) compare current RS methods and critically review their strengths and limitations, and

(b) identify future research opportunities to enrich the investigation and interpretation of RS. **Methods:** This work is part of an international, collaborative, interdisciplinary initiative involving experienced and new researchers. **Methods:** to detect RS were critically appraised, including design-based, qualitative, individualized, preference-based and statistical methods. This assessment was used to describe their similarities and differences in terms of conceptual and operational definitions, and their ability to detect, adjust for, and explain RS. We focused on explicating their underlying assumptions, alternative explanations, and implied conceptual and operational definitions of RS. Following the review of current methodologies, areas for future research were identified and discussed. **Results:** Ten major methods were identified. All methods aim to detect RS, but not all can adjust for or explain RS once detected. **Methods:** use different level of analysis (group or individual level) and hence share some underlying assumptions regarding homogeneity (group level) or self-reflection (individual level). All methods can refer to the definition related to a change in the meaning of one's self-evaluation of a target construct although some do not operationalize the discrepancy between observed and target change (e.g., individualized and preference-based methods). For each method there are different alternative explanations that may account for the results. The limitations identified include the interpretation of results without further substantiation, ignoring inter-individual variation, and limited number of assessments. New avenues are suggested for interpretation of detected RS, handling heterogeneity and multiple time points, individual vs group analyses, and investigations of item vs domain level RS. **Conclusion:** This comprehensive overview of RS detection methods describes how each method defines and operationalizes RS, and how this affects the interpretation of results. The importance of substantive interpretation of detected RS to make alternative explanations less likely is emphasized.

### (116.5) Assessment of the link between response shift and the change over time in the number of Guttman errors via a simulation study

Yseulys Dubuy, UMR INSERM 1246 SPHERE, Nantes, France; Véronique Sébille, UMR INSERM 1246 SPHERE, Nantes, France; Myriam Blanchin, UMR INSERM 1246 SPHERE, Nantes, France; Jean-Benoit Hardouin, UMR INSERM 1246 SPHERE, Nantes, France

**Aims:** **Methods:** for response shift (RS) detection at the individual level could be of great interest when analyzing changes in PRO data. The Guttman errors (GE), which measure discrepancies in each respondent's answers compared to the average sample responses, might be useful for this purpose. Indeed, changes in the individual number of GE could allow identifying, at individual level, patients who might perceive the questionnaire differently than the majority of the sample over time. This study aims at assessing the link between RS and the change in the number of GE over time (denoted IG) via simulations and explores the discriminating ability, the sensitivity, and the specificity of this change. **Methods:** Responses of patients (affected or not by RS) were simulated to determine whether patients with RS had larger changes in their numbers of GE over time than patients without. Effects on IG of factors related to the sample (sample size, proportion of patients affected by RS and average change in the latent trait over time), the questionnaire structure (number of items and number of response categories), and RS (manifestation, number of items affected by RS and position of these items along the latent trait continuum) were investigated. **Results:** As expected, patients affected by RS had, on average, higher changes in the number of GE over time than patients without. The following parameters showed substantial effects on the performances of IG: the

number of items affected by RS, the position of these items along the latent trait continuum (a parameter rarely considered in simulation studies), the number of response categories per item and the number of items. In our simulation framework, IG performed well when items affected by RS were located in the lower tail of the latent trait continuum and when the number of response categories and the number of items affected by RS were large. **Conclusion:** The link between RS and the change in the number of GE was established and assessed. GE could be a valuable non-parametric tool for RS detection at a more individual level.

Table 1: Values used to interpret individual patient-level change on PROs in RA

PRO	Values applied to individual patient-level change in RA clinical trials of JAKs			Alternative values identified in the RA literature		
	Value applied to individual patient-level change (and reference to source)	Sufficient evidence to define individual patient-level change?	Methods used to derive value	Value applied to individual patient-level change (and reference to source)	Sufficient evidence to define individual patient-level change?	Methods used to derive value
HAQ-DI	≥0.22 or ≥0.3 (Wells et al., 1993)	No	Anchor-based, but not an established PRO of the same concept (i.e. physical functioning and disability)	≥0.14 to 0.32 (Strand et al., 2014)	No	Anchor-based, but not an established PRO of the same concept (i.e. physical functioning and disability)
Pain VAS (0-100mm)	≥10mm (ref not identified)	No	Unknown method	≥ 11.9mm (Pope et al., 2009)	No	Patient anchor, but relying on recall of previous health status
PGA (0-100mm)	≥10mm (ref not identified)	No	Unknown method	≥ 9.1mm (Wells et al., 1993)	No	Patient anchor, but relying on recall of previous health status
SF-36	Component summary scores: ≥2.5, Domain scores: ≥5 (ref not identified)	No	Unknown method	PCS: > 3.4, MCS: > 4.6, Domain ranges: > 3.4 to > 7.2 (Mansuish et al., 2011) PCS: ≥1.9 to 4.6, MCS: ≥1.1 to 4, Domain: ≥0 to 35.1 (Strand et al., 2014)	No	Improved with use of modified reliable change index, but no patient anchors
FACT-Fatigue (0-52)	≥3-4 or 3.56 (Cella et al., 2005)	No	Distribution-based methods only	≥ 3.9-10 (Cella et al., 2005) 14.2 (Pouchot et al., 2008) 2.5-5.9 (Strand et al., 2014)	No	Cella et al used anchor-based methods with the most relevant patient anchor (SF-36 Vitality scale), but this is still not an established measure of the concept of interest (i.e. fatigue)
Insomnia Severity Index (ISI)	≥ 8.4 decrease (Morin et al., 2011)	No	Anchor-based, but not an established PRO of the same concept	None identified	N/A	N/A
Morning stiffness severity (0-10)	≥10mm (Strand et al., 2019)	No	Unknown method	None identified	N/A	N/A
Morning stiffness duration (minutes)	Half a standard deviation (Strand et al., 2019)	No	Distribution-based methods only	None identified	N/A	N/A

## 117: Policy

### (117.1) Scoping review of patient-reported outcome used by insurance providers in the United States

Amy Cizik, PhD, MPH, University of Utah, Salt Lake City, Utah, United States; Bernice Gulek, MS, ACNP, University of Washington and Washington State University, Seattle, Washington, United States; Suzanne Kinsky, PhD, MPH, UPMC Center for High-Value Health Care, Pittsburgh, Pennsylvania, United States; Polly McCracken, MA, University of Pittsburgh, Pittsburgh, Pennsylvania, United States; Elizabeth C.S. Swart, BS, UPMC Center for High-Value Health Care, Pittsburgh, Pennsylvania, United States; Rose L. Turner, MLIS, University of Pittsburgh, Pittsburgh, Pennsylvania, United States; Emre Yucel, PhD, MS, MIBS, Pharm, Amgen, Thousand Oaks, California, United States; Janel Hamner, MD, PhD, University of Pittsburgh, Pittsburgh, Pennsylvania, United States

**Aims:** To perform a scoping review of insurance providers in the United States (US) using Patient-Reported Outcomes (PROs). **Methods:** Medline (PubMed), Web of Science (Clarivate Analytics), the Cochrane central register of controlled trials (Wiley), Business Source Complete (EbscoHost) and EconLit (EbscoHost) were searched electronically for the period of 1999 through 2019 for studies

related to health insurers using patient-reported outcomes to manage population health. Articles were evaluated first by abstract and then by full text. Each article was evaluated by two reviewers who then reconciled any conflicts. The manuscripts which pass through abstract and full-article review will be synthesized to understand the uses of PROs in US insurance providers. Results: The literature search identified 14,248 abstracts. Abstracts were excluded for being outside the US (4282), not sampling from a health-insured population (6474), and other reasons (e.g., digital health) (783). 2709 manuscripts are undergoing full article screening. At the time of submission, 244 articles have been evaluated by full text. Manuscripts were excluded for not sampling from a health-insured population (42), not focusing on insurer population health management (125), not including PROs (25), and other reasons (14). If proportions hold, 421 articles will be in the final report. In the first 38 manuscripts synthesized, emerging themes include: the most common PRO used by health insurers is a single self-rated health question that is used as an independent variable (e.g., as an adjustor) in population health outcome models; PROs have been used to evaluate population health strategies such as clinical interventions and payment strategies; PROs have been used to evaluate and monitor quality of care; PROs have been used to predict future health status, utilization, and/or costs. Many areas for population health management, such as lifestyle management—reducing obesity, smoking, alcohol misuse insurers seem to be missing opportunities to use validated PRO measures. Conclusion: PROs are being used by US health insurers, but the most common use is as an independent variable in population health outcome data models. There is an enormous opportunity to develop expanded use cases and methodologies for PROs use by US health insurers and determine unmet needs within patient-centered insurance design.

### **(117.2) Integrating policies to address the use of street medicine in Haiti: a qualitative study**

Farrah Pompilus, PhD, MCPHS University; Modus Outcomes, Boston, Massachusetts, United States

Aims: The purchase of medicine through street vendors are often used in Haiti because of the government's challenges with supplying medicine to its citizens. The goal of this research was to create policy recommendations to improve patients' access to medicine. Methods: The research consisted of a policy review, conceptual literature review, the development of a conceptual framework, consultations with pharmacists, interviews with patients, the development of a PRO instrument, and health policy recommendations. Search terms were created to conduct the policy review on the ministry of health document center and the concept review on PubMed. Data were extracted, analyzed, and used to develop into a conceptual framework of barriers to access and adhere to medicine. Consultations with pharmacist and patients living in Haiti were used to further inform the conceptualization and develop a patient-reported outcome instrument. Consolidated results from the research activities were used to create a set of policy recommendations to improve Haiti's medicine access. Results: A review of the policies revealed a lack of appropriate regulation and inadequate pharmaceutical services. A total of 82 unique concepts emerged from the literature review and organized into a 13-domain conceptual framework. Consultations with four practicing pharmacists and interviews with  $n = 19$  patients (mean age = 34.05, standard deviation = 13.51) led to the revisions of the conceptual framework. The barriers patients' faced were extensive and included issues around the supply of medicine within the pharmacy, medicine

effectiveness, staff, physiological distress, logistical issues, attitudes and beliefs, financial issues, family influence, environmental factors, memory, education, political, and discrimination issues. Six policy improvements were recommended to combat some barriers, including: standardize policy prices, incentivize pharmacy students, implement mobile pharmacies, develop a pharmaceutical manufacturing facility, and monitor patients access to medicine. Policy improvements can be monitored by the newly developed survey, Haitian Access and Adherence to Medicine (HAAM) scale. Conclusion: Policy recommendations were developed to help improve the access of medicine for patients living in Haiti. By having this information, the MSPP can better structure their policies with considerations on the informal medicine sector based on the patient's experience.

### **(117.3) An overview of self-reported measures used in patients with renal replacement therapy: Are existing instruments appropriate for healthcare quality assessment in Germany?**

Gregor Liegl, IQTIG, Berlin, Germany; Christopher Kienle, IQTIG, Berlin, Germany; Julia Böttcher, IQTIG, Berlin, Germany; Julia Ginkel, IQTIG, Berlin, Germany; Tobias Mertzig, IQTIG, Berlin, Germany; Carsten Volland, IQTIG, Berlin, Germany; Mandy Wagner, IQTIG, Berlin, Germany; Konstanze Blatt, IQTIG, Berlin, Germany

Aims: The German Federal Joint Committee (G-BA) commissioned the Institute for Quality Assurance and Transparency in Healthcare (IQTIG) to develop a patient survey based on patient-reported experience measures (PREMs) and patient-reported outcome measures (PROMs) for quality assessment of renal replacement therapy (RRT) in Germany. In a first step, quality-related aspects of care relevant to dialysis patients and renal transplant recipients were identified. The aim of this subproject was to examine whether existing patient-reported instruments are suitable for measuring the set of quality-related aspects or whether a specific instrument for RRT quality assessment in Germany needs to be developed. Methods: Four bibliographic databases (MEDLINE, Embase, CINAHL, and the Cochrane Library) were searched systematically. Review articles on PREMs and PROMs used in patients with RRT that were published between 2013 and 2019 in English or German language were included. Two reviewers performed title-abstract screening and subsequent full-text evaluation independently. In addition, we searched the Patient-Reported Outcome and Quality of Life Instruments Database (PROQOLID) for further instruments. We explored the agreement of each identified instrument with the construct definitions as prespecified for the quality-related aspects of RRT in Germany. Results: From 962 screened abstracts, 33 full-texts were retrieved of which 24 met our inclusion criteria. In sum, 113 patient-reported instruments were identified. Most instruments were generic, only 30 instruments were disease and/or treatment specific (kidney disease: 7; dialysis: 16; transplantation: (7)). The majority of instruments were classified as PROM, only 7 instruments were classified as PREM or a combination of both. While most relevant RRT outcome domains could potentially be addressed by existing PROMs, none of the identified PREMs were rated to satisfy the requirements for being used as patient-reported healthcare quality assessment tool in Germany. Conclusion: Most of the predefined quality-related aspects of RRT in Germany are related to processes of care (e.g., 'patient-centered decision-making'), which are measured by PREMs. However, existing PREMs were rated to be not specific enough to validly assess the quality of RRT in Germany.

Consequently, a targeted patient-reported RRT quality assessment tool with consistent layout and structure will be newly developed.

#### (117.4) Are healthcare organizations in Canada ready to use patient-reported data to improve person-centered care?

##### A system-level perspective on the feasibility of implementing person-centered quality indicators

Kimberly Manalili, PhD Candidate, University of Calgary, Calgary, Alberta, Canada; Cathie Scott, PhD, University of Calgary, Calgary, Alberta, Canada; Brenda Hemmelgarn, MD, PhD, University of Alberta, Edmonton, Alberta, Canada; Maeve O’Beirne, MD, PHD, University of Calgary, Calgary, Alberta, Canada; Paul Fairie, PhD, University of Calgary, Calgary, Alberta, Canada; Brian Steele, BSc, University of Calgary, Calgary, Alberta, Canada; Maria Jose Santana, MPharm, PhD, University of Calgary, Calgary, Alberta, Canada

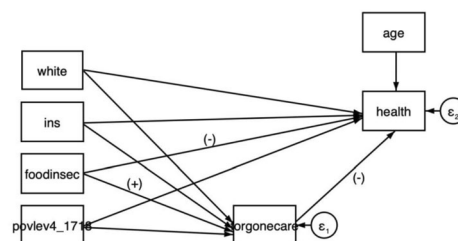
**Aims:** Despite efforts to use patient-reported data to improve person-centered care, the extent to which a healthcare organization is “ready” to implement these measures influences the feasibility of implementation, and consequently, their effectiveness in driving real improvements in care. To ensure optimal implementation of Person-Centered Quality Indicators (PC-QIs), we assessed how feasible it would be to implement these indicators in various healthcare jurisdictions across Canada, based on their reported readiness to use PC-QIs to make system-level improvements in person-centered care (PCC). **Methods:** A web-based survey was conducted with representatives of healthcare delivery and coordinating organizations that guide the development and/or implementation of person-centered care measurement in Canada between November 2019 and March 2020. The survey was comprised of two sections testing organizational readiness theory. In the first section, participants assessed each PC-QI on interest in implementing the PC-QIs, measurability (validity), and whether data can be interpreted and used as part of quality improvement processes to improve PCC. The second component was adapted from the validated Organizational Readiness for Change tool, to assess motivational factors and general capacity for implementation. **Results:** There were 33 regional healthcare organizations represented, covering all 13 provinces/territories across Canada (60% response rate). Across all 26 PC-QIs, more than 85% of organizations indicated interest in implementing the indicators. However, there were only four PC-QIs that were considered highly feasible to implement—where 75% of organizations indicated that the data were already being collected for that particular indicator and quality improvement processes were in place to make changes. These PC-QIs included: structures to report person-centered performance; communication between the healthcare provider—nurse; coordination of care; patient and caregiver decisions about their care and treatment; and overall experience. **Limitations in resources** (e.g., time constraints and data systems) were seen as the most challenging aspects of readiness for implementation. **Conclusion:** Despite high motivation and interest in using PC-QIs to improve PCC, most organizations across Canada are not “ready” to implement them. Efforts are needed to ensure that organizations have the capacity to collect, use, and report data on PCC in order to make the needed improvements that matter to patients.

#### (117.5) Disproportionate impact of food insecurity on forgone healthcare and health of US adolescents with special healthcare needs

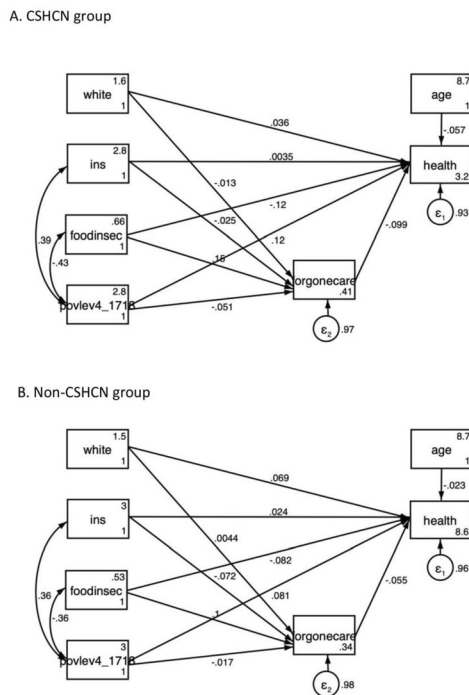
Nalin Payakachat, PhD, University of Arkansas for Medical Sciences, Little Rock, Arkansas, United States; J Mick Tilford, PhD, University of Arkansas for Medical Sciences, Little Rock, Arkansas, United States

**Aims:** Food insecurity is a hardship among American families due to limited or lack of resources to buy food. Families with food insecurity may choose to forgo healthcare for a child, which in turn negatively impacts child health. This study evaluates direct and indirect associations between food insecurity, foregone healthcare, and health among adolescents aged 12–17 years who have special healthcare needs (CSHCN). **Methods:** A cross-sectional study was conducted using the combined 2017–2018 US National Survey of Children’s Health. Only adolescents between 12–17 years old were included. CSHCN were identified by responses of yes to any of the 5 screener questions. Food insecurity status was determined by whether the child’s household was able to afford the food they need during the past 12 months (0 = always afford to eat good nutritious meals; 1 = always afford enough to eat, but not always the kinds of food we should eat; 2 = sometimes/often we could not afford enough to eat). Health was measured using the overall health status question (0 = fair/poor; 1 = good; 2 = excellent/very good). A structural equation modeling with a group analysis (CSHCN vs. non-CSHCN) was employed to determine the association of food insecurity and foregone healthcare on child health, adjusted by age, sex, insurance type, and poverty levels (Fig. 1). **Results:** Of 21,496 adolescents, 28.5% (n = 6,127) were in the CSHCN group. Both groups were similar in average age ( $14.7 \pm 1.7$  years), race (70% white), and sex (52% male). The CSHCN group reported poorer health ( $p < 0.001$ ) and higher rate of forgone healthcare (6.6% vs. 1.8%) than the non-CSHCN group. Food insecurity directly impacted health of CSHCN in a higher degree than non-CSHCN (std.coef =  $-0.12$  vs.  $-0.082$ ,  $p < 0.001$ ) (Fig. 2). The total effect of food insecurity on health, based on direct and indirect effects through forgone healthcare, was significantly higher in the CSHCN group than the non-CSHCN group (mean difference =  $-0.083$ ,  $p < 0.001$ ). **Conclusion:** Food insecurity increases the probability that families forgo healthcare needs for adolescents, which negatively impacts child health. The effect of food insecurity is especially pronounced among CSHCN. If the findings are causal, policy interventions that alleviate food insecurity (such as Supplemental Nutrition Assistance Program benefits) have the potential to improve health of adolescents with special healthcare needs.

**Disproportionate impact of food insecurity on forgone healthcare and health of US adolescents with special healthcare needs**



**Figure 1.** A conceptual model with the assumption that food insecurity increases forgone healthcare and negatively impact health of adolescents aged 12–17 years, adjusted for age, sex, race, insurance types, and poverty levels.



**Figure 2.** Group estimations with standardized coefficients using a structural equation modeling approach for CSHCN and non-CSHCN groups

## 118: Measurement Development & Evaluation II

### (118.1) Adaptation of patient-reported outcome measures in a low-literacy non-western population: an example of the Patient-Specific Functional Scale (PSFS)

Anupa Pathak, PhD Student, University of Otago, Dunedin, New Zealand; Saurab Sharma, PhD, Kathmandu University School of Medical Sciences, Dhulikhel, Nepal; Allen Heinemann, PhD, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Daniel Cury Ribeiro, PhD, University of Otago, Dunedin, New Zealand; J Haxby Abbott, PhD, University of Otago, Dunedin, New Zealand

**Aims:** Individuals from diverse cultural backgrounds, including indigenous populations and immigrants, often struggle to comprehend patient-reported outcome measures (PROMs). Response options presented as a numeric rating scale create particular problems. The Patient-Specific Functional Scale (PSFS) is a widely used PROM with a 0–10 response format. We aimed to develop verbal response options for the PSFS for use in a low literacy country in South Asia—Nepal. We also sought to assess if error rates were affected by age, education, language, or previous experience using numeric rating scales. We hypothesized that respondents would prefer a verbal rating scale for the PSFS to a numeric rating scale and have fewer errors. **Methods:** The study was conducted in two phases. First, we interviewed 42 individuals with musculoskeletal problems, chronic obstructive pulmonary disease, spinal cord injury, and stroke to understand how they describe varying levels of abilities. Then, we developed verbal response options for the PSFS. In Phase 2, we pre-tested the scales on 118 participants using the three-step test interview and paired comparison survey. We asked participants to indicate which response option they preferred and coded error rates qualitatively as a logical inconsistency, missing response, and/or multiple responses. **Results:** Participants most commonly described their

ability in terms of the quality (96%) and quantity of task performance (88%). We developed two sets of verbal responses for the PSFS and pre-tested them. Although respondents preferred the verbal (50%) over the numeric rating scale (12%), error rates were similar between numeric (34%) and verbal scales (31%, and 36%). Error rates were associated with previous use of a numeric scale, age, and years of education, with some groups displaying up to 80% error. **Conclusion:** While the PSFS is recommended for use in clinical practice, the scale can have high error rates among Nepalese patients, especially those who are older, have less education, and no prior rating scale experience. Errors are not related to the type of response options. Patients may benefit from an interview format, explanatory prompts, fewer response options, and continued use of PROM along with observational measures.

### (118.2) A PROMIS physical function clinical outcome assessment (COA) designed for use in advanced cancer

David Cella, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Sara Shaunfield, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; John Devin Peipert, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Michael Anthony Kallen, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Robert Chapman, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Katie Benjamin, Abbvie, Chicago, Illinois, United States; Linda Deal, Pfizer, New York, New York, United States; Daniel Eek, AstraZeneca, Gothenburg, Sweden; Laurie Eliason, GlaxoSmithKline, Collegeville, Pennsylvania, United States; Kelly McQuarrie, Janssen, Horsham, Pennsylvania, United States; Allison Martin Nguyen, Janssen, Horsham, Pennsylvania, United States; Allison Martin Nguyen, Merck, Kenilworth, New Jersey, United States; James W. Shaw, Bristol Myers Squibb, Lawrenceville, New Jersey, United States; Susan Vallow, Novartis, East Hanover, New Jersey, United States

**Aims:** To develop a PROMIS® Physical Function (PF) short form for qualification by the Food and Drug Administration (FDA) as an approved Drug Development Tool (DDT) in advanced cancer. **Methods:** In partnership with the pharmaceutical industry and FDA, we identified PROMIS PF items suited to the outcome assessment of patient-reported PF across a range of treatments for advanced cancer. Beginning with cancer-targeted PROMIS PF short forms and items judged to be relevant from the 165-item PROMIS PF bank, we identified a subset of 31 highly relevant and responsive items. These were evaluated using mixed methods, including cognitive debriefing with 31 patients from 5 representative countries/languages, the NCI-funded MY Health Study, and a panel survey ( $n = 2400$ , including 1,000 cancer patients), to assess differential item functioning (DIF), and reliability and criterion (convergent and known groups) validity. Cognitive interviews and item-level analyses provided data to evaluate COA performance vis-à-vis translation quality and requirements for FDA approval for limited context of use. **Results:** Cognitive interview results revealed good respondent understanding of the 8 items (Table). Cross-sectional analyses revealed strong psychometric properties for the PF 8c. Floor ( $< 0.001\%$ ) and ceiling ( $< 5\%$ ) effects were minimal. Reliability was  $> 0.90$  over much of the score range without much precision loss vs. the original 31 items (Figure). Spearman's correlations between the PF 8c and criterion variables ranged between 0.47 and 0.94, supporting convergent validity. Known-groups validity was supported across clinically distinct subgroups (adjacent category effect size range = 0.30–0.50). No appreciable DIF was found. **Results:** were reviewed in a joint meeting of investigators from Northwestern University, FDA, National Cancer Institute, the pharmaceutical

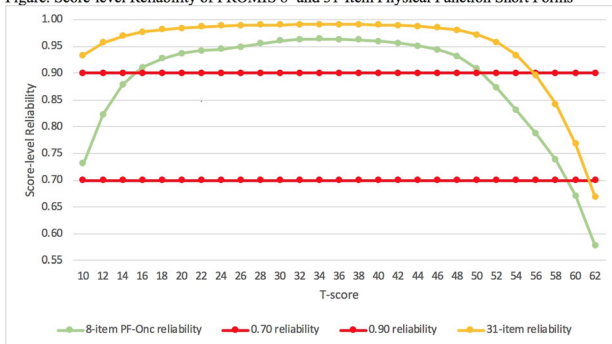
industry, and the patient community. This produced an 8-item short form (PF 8c) that addressed the input of all constituents. Conclusion: The PROMIS Short Form v2.0—PF8c, a robust, 8-item short form derived from the PROMIS PF item bank, was developed collaboratively among a diverse set of stakeholders. It is under review by the USA FDA for approval as a COA for limited context of use in advanced cancer clinical trials. Further research using the PF8c in industry clinical trials will evaluate longitudinal validity (e.g., responsiveness to change, responder definition).

Table. Participant perceptions of item relevance and confidence when responding to the PROMIS Short Form v2.0—Physical Function 8c

Item ID	Item Stem	When responding, how confident did you feel about your answer to this question?	Is this question relevant to your experiences with cancer? <sup>1</sup>
PFA21	Are you able to go up and down stairs at a normal pace?	Very confident = 22 (71%) Confident = 8 (26%) Not at all confident = 1 (3%)	Yes = 26 (84%) No = 5 (16%)
PFA9	Are you able to bend down and pick up clothing from the floor?	Very confident = 22 (71%) Confident = 8 (26%) Not at all confident = 1 (3%)	Yes = 22 (71%) No = 9 (29%)
PFA16r1	Are you able to dress yourself, including tying shoelaces and buttoning your clothes?	Very confident = 24 (77%) Confident = 7 (23%) Not at all confident = 0	Yes = 24 (77%) No = 7 (23%)
PFA55	Are you able to wash and dry your body?	Very confident = 22 (71%) Confident = 8 (26%) Not at all confident = 1 (3%)	Yes = 22 (71%) No = 9 (29%)
PFA23	Are you able to go for a walk of at least 15 minutes?	Very confident = 22 (71%) Confident = 8 (26%) Not at all confident = 1 (3%)	Yes = 22 (71%) No = 9 (29%)
PFA15	Are you able to stand up from an armless straight chair?	Very confident = 20 (65%) Confident = 11 (35%) Not at all confident = 0 (0%)	Yes = 22 (71%) No = 9 (29%)
PFA1	Does your health now limit you in doing vigorous activities, such as running, lifting heavy objects, participating in strenuous sports?	Very confident = 19 (61%) Confident = 11 (35%) Not at all confident = 1 (3%)	Yes = 28 (90%) No = 3 (10%)
PFB50	How much difficulty do you have doing your daily physical activities, because of your health?	Very confident = 21 (68%) Confident = 10 (32%) Not at all confident = 0 (0%)	Yes = 23 (74%) No = 8 (26%)

<sup>1</sup>For each item, ~4 participants that reported the questions were not relevant to cancer were from the Chinese sample, in which the interviewer instead asked if cancer impacts their ability to perform the activities described in each item, to which these participants reported their answers would be the same whether or not they had cancer and thus recorded as not relevant.

Figure. Score-level Reliability of PROMIS 8- and 31-Item Physical Function Short Forms



### (118.3) Prevalence of gas-related symptoms and impact on quality of life in the French general population using the newly validated Intestinal Gas Questionnaire (IGQ)—a 1543 participants online survey

Martin Duracinsky, MD, PHD, CTU Health Economics, Hopital Hotel-Dieu, AP-HP, Paris, France; Boris Le Nevé, PhD, Danone Nutricia Research, Palaiseau, France; Laurent Quinquis, Danone Nutricia Research, Palaiseau, France; Bénédicte Monnerie, Danone Nutricia Research, Palaiseau, France; Chloé Six, CREDOC, Paris, France; Gabriel Tavoularis, CREDOC, Paris, France; Olivier

Chassany, MD, PhD, Patient-Centered Outcomes Research, Université de Paris, Paris, France

Aims: Available data on prevalence of gastrointestinal symptoms in general population (GP) are limited and often using non-validated tools. Aim of this study was to generate first data on the frequency of gas-related symptoms and their impact on quality of life (QoL) in a representative sample of French population using the newly validated Intestinal Gas Questionnaire (IGQ). Methods: 1543 adults from “Behaviors and food consumption in France” (CCAF) survey 2019 were recruited by phone to complete an online survey between January and July 2019. Participants provided socio-demographic characteristics and completed the 17-item IGQ questionnaire (7 symptom severity items and 10 impact on QoL items) and a lifestyle questionnaire. To ensure representativity of the sample, a quota-based method sampling used gender, age, employment, familial status, education level, region and size of urban area. IGQ global score and a score for each of the 6 dimensions (bloating, flatulence, belching, bad breath, stomach rumbling, difficult gas evacuation (range: 0-100) was computed. Descriptive statistics and non-parametric tests investigated associations between symptoms, socio-demographic characteristics and lifestyle parameters. Subjects rating at least one IGQ symptom severity item AND one IGQ impact item over the middle of the response scale (i.e., > 5 on 0-10 numeric scale) were considered bothered by digestive problems. Results: Mean  $\pm$  SD global IGQ score was  $11.2 \pm 10.8$ . Individuals mostly impacted by their symptoms represented 22% of the cohort ( $n = 388$ ) with a mean total IGQ score of  $25.9 \pm 11.9$ . Only 7% of individuals declared being free of any digestive symptom (IGQ total score = 0). Following parameters were associated with significantly higher (worse) IGQ global scores: age 18–24 ( $p < 0.001$ ), sedentary lifestyle ( $p < 0.001$ ), unemployed ( $p < 0.001$ ), bigger size urban area ( $p = 0.042$ ), underweight and obese ( $p = 0.005$ ), being on a diet ( $p = 0.031$ ), having allergy ( $p = 0.006$ ) and smoking ( $p = 0.003$ ). Female gender was associated with significantly higher bloating and difficult gas evacuation scores ( $p < 0.001$ ). Among symptomatic individuals, 21% took an action including medication. Conclusion: Using the newly validated IGQ, we quantified for the first-time the prevalence of symptoms related to intestinal gas in a representative sample of the French GP. Age, sedentary lifestyle and BMI are associated with severity of symptoms and impact on QoL.

### (118.4) Regulatory development of a novel outcome measure to assess communication ability in children with Angelman syndrome

Jennifer Panagoulas, RAC, GeneTx Biotherapeutics, Sarasota, Florida, United States; Allyson Berent, DVM, GeneTx Biotherapeutics, Sarasota, Florida, United States; Christy Zigler, PhD, Duke University, Durham, North Carolina, United States; Nicole Lucas, Duke University, Durham, North Carolina, United States; Li Lin, Duke University, Durham, North Carolina, United States; Molly McFatrigh, Duke University, Durham, North Carolina, United States; Kelly Gordon, Duke University, Durham, North Carolina, United States; Harrison Jones, Duke University, Durham, North Carolina, United States; Paula Evans, GeneTx Biotherapeutics, Sarasota, Florida, United States; Bryce Reeve, PhD, Duke University, Durham, North Carolina, United States; PhD, Duke University, Durham, North Carolina, United States

Aims: Angelman syndrome (AS) is a rare neurogenetic disorder caused by the loss of expression/function of the maternally inherited UBE3A gene in neurons. Among many severe symptoms including developmental delay, apraxia, and seizures, a hallmark feature of individuals with AS is a profound lack of speech. No treatments are

currently approved by the US FDA to treat AS. A survey sponsored by a non-profit patient foundation was administered to caregivers through social media in order to identify the most relevant functional domains where a positive change could have a substantial impact on quality of life. Existing communication measures were generally inadequate to evaluate communication in non-verbal individuals with AS due to large floor effects. In a patient focused listening session with FDA to discuss endpoints that were meaningful to caregivers, it was suggested that a novel communication measure relevant to AS could be developed. Methods: A caregiver-reported questionnaire was developed to evaluate expressive, receptive and pragmatic communication in children with AS for use in clinical trials of potential disease modifying therapies. This Observer-Reported Communication Ability (ORCA) measure allows examination of changes in communication ability over time but does not rely on speech. ORCA development followed established best practices for outcome measure development; however, the process was unique as the study team incorporated insights from academic experts, patient advocates, regulatory specialists, and community stakeholders at multiple stages of development. Results: A novel observer reported outcome measure was successfully created in 1 year to assess a meaningful concept of interest for caregivers of individuals with AS: communication ability. The questionnaire assesses 22 conceptual areas of communication ability in individuals with AS. The ORCA is being piloted in an ongoing Phase 1/2 clinical trial in AS. Conclusion: Qualitative and quantitative methodology and careful consideration of regulatory feedback gave rise to a new measure intended to broadly evaluate changes in communication in individuals with AS. This work was successfully achieved by uniquely incorporating the regulatory and advocacy perspectives from the beginning and establishing links with community stakeholders. Discussion with FDA will continue to establish its use as a fit-for-purpose assessment of communication ability.

#### **(118.5) Development of an individualized symptom tracking tool for people with Down syndrome and dementia to facilitate Goal Attainment Scaling**

Kari Knox, BScN, DGI Clinical Inc., Halifax, Nova Scotia, Canada; Justin Stanley, BScE, DGI Clinical Inc., Halifax, Nova Scotia, Canada; James Hendrix, PhD, LuMind IDSC Foundation, Burlington, Massachusetts, United States; Hampus Hillerstrom, MSc, MSc, MBA, LuMind IDSC Foundation, Burlington, Massachusetts, United States; Chere Chapman, MPH, MBA, DGI Clinical Inc., Halifax, Nova Scotia, Canada; Kenneth Rockwood, MD, Dalhousie University, Halifax, Nova Scotia, Canada

**Aims:** The increasing life expectancy of people with Down syndrome comes with an increased, age-related risk of Alzheimer disease and other forms of dementia. Identifying symptoms and tracking disease progression in this setting can be difficult due to varying levels of function even before the onset of dementia. Goal Attainment Scaling (GAS) is an individualized patient-reported outcome that may be applicable to monitor disease progression and treatment effectiveness in this population. Here, we revised a dementia symptom menu to facilitate the use of GAS in people with Down syndrome. Methods: A validated dementia symptom menu was revised by four Down syndrome experts. We then recruited 10 caregivers of people affected by both Down syndrome and suspected dementia to participate in individual semi-structured interviews to review the menu. Each participant reviewed 9–15 goal areas to assess the clarity and comprehensiveness of each item. Responses were systematically coded by two researchers as “clear,” “unclear,” or “remove.” Participants were encouraged to suggest additional items and recommend changes to items that were unclear. Results: The median caregiver age was

65 years (range 54–77). Most were female (9/10) with  $\geq 15$  years of education (10/10). The person for whom they cared had a median age of 58 years (range 52–61) with Down syndrome and either a formal diagnosis (6/10) or clinical suspicion (4/10) of dementia. We revised a dementia symptom menu consisting of 58 goal areas each with 4–17 descriptors (580 total). Of the 580 descriptors, 37 (6%) were unclear and were reworded; one goal area (4 descriptors) was removed. A further 47 descriptors were added (including one goal area) to include participant-identified concepts. The final menu contained 58 goal areas, each with 7–17 descriptors (623 total). Conclusion: A comprehensive symptom menu for people with Down syndrome and dementia was developed to facilitate GAS. Incorporating both expert clinician opinion and input from caregivers of people with Down syndrome and dementia identified meaningful items that incorporate patient/caregiver perspectives.

#### **(118.6) The Endometriosis Daily Diary: Modifying An Existing PRO for endometriosis-related pain**

Yanfen Guan, M.S., Merck & Co., Inc., Kenilworth, NJ, USA, Kenilworth, New Jersey, United States; Charlotte Panter, Adelphi Values, Bollington, Cheshire, United Kingdom; Samantha Wratten, Adelphi Values, Bollington, Cheshire, United Kingdom; Sharan Randhawa, Adelphi Values, Bollington, Cheshire, United Kingdom; Rob Arbuckle, Adelphi Values, Bollington, Cheshire, United Kingdom; Arnaud Fauconnier, Université Paris-Saclay, UVSQ, RISCQ, Service de gynécologie et obstétrique, Centre Hospitalier Intercommunal de Poissy-Saint-Germain-en-Laye, Bollington, Cheshire, France; Linda Giudice, University of California, San Francisco, California, United States; Felipe Arbelaez, Merck & Co., Inc., Kenilworth, NJ, USA, Kenilworth, New Jersey, United States; Jessica Weaver, Merck & Co., Inc., Kenilworth, NJ, USA, Kenilworth, New Jersey, United States; Allison M Nguyen, Merck & Co., Inc., Kenilworth, NJ, USA, Kenilworth, New Jersey, United States; Merck & Co., Inc., Kenilworth, NJ, USA, Kenilworth, New Jersey, United States

**Aims:** Endometriosis-related pain (ERP) is the hallmark symptom of endometriosis (a disease affecting 6–10% of women of reproductive age) and has a substantial impact on health-related quality of life (HRQoL). This research involved modification of a 7-item dysmenorrhea daily diary (DysDD) to develop a fit-for-purpose PRO measure, the Endometriosis Daily Diary (EDD), for use in ERP clinical trials. Methods: In line with current regulatory guidance and consistent with guidelines in recent publications, the EDD was developed by: (1) drafting an endometriosis disease conceptual model based on review of qualitative literature and on-line patient forums; (2) conducting 60-min concept elicitation (CE) interviews with 30 United States (US) females with ERP; (3) updating conceptual model and modifying DysDD to form draft EDD; (4) refining EDD items based on 2 rounds of cognitive debriefing (CD) (15 US females each round) and a translatability assessment. Two clinical experts were consulted at key stages throughout the study. Results: For CE, 20 adults (aged 18–49) and 10 adolescents (aged 12–17) with mild to severe ERP were recruited. Conceptual saturation was achieved in the CE sample; thus the updated conceptual model provides a comprehensive summary of endometriosis to inform modification of the DysDD. Across 2 rounds of CD, a total of 20 adults and 10 adolescents (aged 12–49) were recruited. Most items were considered relevant for the majority of participants. All instructions, items, response scales, and the recall period were generally well understood and interpreted as intended. The EDD is a 28-item, 24-h recall daily eDiary that assesses both symptoms and impact of ERP on patients' HRQoL, including severity of cyclic and non-cyclic pelvic pain (3 items), dyspareunia (4 items), impact of ERP on functioning and daily

life (14 items), symptoms associated with ERP (4 items), and bowel symptoms (3 items). Conclusion: An existing PRO, the DysDD, was successfully leveraged to develop a new fit-for-purpose measure specific to ERP. The EDD demonstrated content validity in both adults and adolescents experiencing ERP. Following usability testing and psychometric validation, the EDD is expected to be a reliable and valid PRO for use in clinical trials of novel ERP treatments.

## 119: Research in Neurological & Mental Health

### (119.1) Clinician Perspective on Clinical Outcome Assessments in Neurosarcoidosis

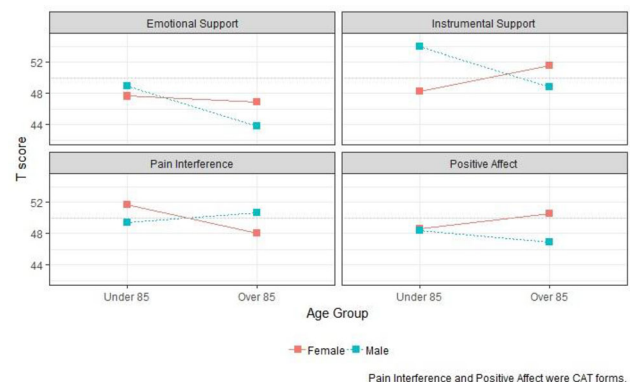
Brandon Moss, MD, Cleveland Clinic Foundation, Cleveland, Ohio, United States; Deborah Miller, PhD, Cleveland Clinic Foundation, Cleveland, Ohio, United States; Daniel Culver, DO, Cleveland Clinic Foundation, Cleveland, Ohio, United States

**Aims:** Neurosarcoidosis (NS) is a rare, costly, and often disabling disease, characterized by its heterogeneity and variability. The ability to track disease progression and treatment response in clinical practice has been limited by a lack of standardized, validated clinical outcome assessments (COAs). We conducted a survey to describe what COAs clinicians are currently using for patients with NS and to understand clinicians' needs regarding future COA development. **Methods:** Surveys were sent to members of the NS Consortium and a list of neurologists and clinic directors for Sarcoidosis Centers of Excellence maintained by the Foundation for Sarcoidosis Research. Survey questions were developed with the assistance of a PhD-level sociologist. The questionnaire was pilot tested for face validity and clarity with 5 participants. **Results:** There were 43 respondents out of 156 surveys sent. 58% were neurologists, 37% pulmonologists. 86% were from the United States, 14% international (Canada, Germany, Italy, Spain, United Kingdom). 47% were center directors. Among clinicians, there was a median of 9.5 years of experience treating patients with NS. 19% of clinics used patient-reported outcome measures (PROMs), 21% clinician-reported outcomes, and 21% performance outcomes. Fatigue and health-related quality of life (HRQoL) PROMs, the Expanded Disability Status Scale, and the Timed 25-Foot Walk were the most commonly used instruments for each COA type, respectively. Neuropathic pain, problems with lower extremity function, headaches, and numbness were the most frequently encountered manifestations of disease. Cognitive deficits, visual impairment, neuropathic pain, problems with lower extremity function, and bowel/bladder dysfunction were judged to have the most impact on HRQoL. Tracking disease progression and treatment response was ranked as the primary reason to use COAs. Time and lack of disease-specific measures were ranked as the primary barriers to using COAs. **Conclusion:** Our results suggest there is a low frequency of COA use in clinical practice, partly due to time requirements and lack of disease-specific measures. Validation of measures for NS will be important to move towards standardized assessments in the field. Screening tests may be useful to limit time requirements for COA administration given the heterogeneity of disease manifestations.

### (119.2) Differences in emotional health across age groups and gender in cognitively healthy adults and with Mild Cognitive Impairment and Alzheimer's Disease: results from Advancing Reliable Measurement in Alzheimer's Disease and Cognitive Aging (ARMADA)

Emily Ho, Northwestern University, Chicago, Illinois, United States; Cindy Nowinski, Northwestern University, Chicago, Illinois, United States; Sandra Weintraub, Northwestern University, Chicago, Illinois, United States; Richard Gershon, Northwestern University, Chicago, Illinois, United States

**Aims:** The ARMADA study is evaluating use of the NIH Toolbox for Assessment of Neurological and Behavioral Function (NIHTB) in older adults (age 65 and older) with normal to impaired cognition. The NIH TB assesses cognitive, motor, sensory and emotional health. Goals of this report are to: (1) compare self-reported emotional health and functioning across cohorts of cognitively healthy adults and those diagnosed with Mild Cognitive Impairment (MCI) and Alzheimer's Disease (AD), and (2) compare the same measures across gender and cognitively healthy age groups: young old (below 85 years) and oldest old (above 85 years). **Methods:** The NIHTB Emotion Battery, comprised of measures of negative affect, psychological well-being, stress and self-efficacy, and social functioning, was administered to 656 US participants (young old = 307, oldest old = 106, MCI = 176, AD = 67). The mean age was 76.85, SD = 7.44, with 61% female). One-way ANOVAs were conducted on cognitive health groups to examine potential differences in emotional health. Additionally, two-way ANOVAs with interaction terms for age groups and gender were conducted. **Results:** Based on preliminary data collection, there were few differences in reported emotional health outcomes across cognitive health groups, with the exception that cognitively healthy and MCI adults reported significantly more self-efficacy than AD participants ( $p < 0.01$ ). Looking only at cognitively healthy adults, older participants reported significantly lower levels of anger affect, fearfulness, negative affect, anger hostility, and perceived hostility. There were also main effects of gender on emotional outcomes: males reported more anger and physical aggression, though less sadness and friendships. Significant interactions between age groups and gender were found in pain interference, positive affect, emotional support, and instrumental support (Fig. 1). **Conclusion:** These early findings suggest that those with MCI generally enjoy similar emotional outcomes compared with cognitively healthy adults. Comparing emotional health outcomes across age groups corroborates prior research that the oldest old experience less negative emotion, with better emotional regulation. However, oldest old males experienced less emotional and instrumental support and less positive affect, suggesting differential experiences in emotional well-being across the latter life span.





### (119.3) Estimating health-related quality of life for adults with learning disabilities

John O'Dwyer, MSc, University of Leeds, Leeds, United Kingdom; David Meads, PhD, University of Leeds, Leeds, United Kingdom; Louise Bryant, PhD, University of Leeds, Leeds, United Kingdom; Paul Kind, PhD, University of Leeds, Leeds, United Kingdom; Claire Hulme, PhD, University of Exeter, Exeter, United Kingdom

**Aims:** Over 1.2 million people in the UK have a mild or moderate learning disability (LD), living on average 16 years less than the general population. Research with this population is constrained with evidence suggesting a difficulty completing research materials, including the EQ-5D. The main difficulty experienced is with understanding the meaning of the wording/language of the statements. Rephrasing and explanation of terms make completion easier, however this invalidates the questionnaire. The aim of this study is to develop a self-report version of the EQ-5D for adults with mild or moderate LD. **Methods:** To inform the development of the adaptation, the study brings together evidence from a systematic review on self-reported quality of life (QoL) measures that have been used with this population to identify potential adaptations that might be made to the EQ-5D. In-depth semi-structured interviews and focus groups were conducted with carers/supporters of adults with LD. A further focus group was conducted with a group of adults with LD. This qualitative phase was used to explore and understand the key difficulties experienced by people with LD when completing and understanding the EQ-5D. Using the findings from the review, interviews and focus groups, an adapted self-reported version of EQ-5D for adults with a mild or moderate LD was developed. **Results:** From the systematic review, 47 self-reported generic QoL/HRQoL measures were identified; 13 have been validated for use by adults with mild to moderate LD. Adaptations such as pictograms, contextualization of language and longer completion times are reported. Following a framework analysis of the qualitative data, dimensions included in the EQ-5D were deemed appropriate for the measurement of HRQoL in adults with LD. Amendments to dimension titles are suggested. The adapted version includes pictograms and simplified wording suitable for adults with mild or moderate LD. **Conclusion:** The new version of EQ-5D will be tested with a general population sample in order to assess the extent to which valuations using the adapted EQ-5D correspond to the previously established measure. The final phase focuses on testing the adapted version of EQ-5D in terms of ease of completion, internal consistency and reliability.

### (119.4) How the COVID-19 pandemic impacts the psychosocial well-being of children and adolescents in the Netherlands

Michiel Luijten, MSc, Emma Children's Hospital Amsterdam UMC, Amsterdam, Netherlands; Maud van Muilekom, MSc, Emma Children's Hospital Amsterdam UMC, Amsterdam, Netherlands; Lorynn Teela, MSc, Emma Children's Hospital Amsterdam UMC, Amsterdam, Netherlands; Hedy van Oers, PhD, Emma Children's Hospital Amsterdam UMC, Amsterdam, Netherlands; Kim Oostrom, PhD, Emma Children's Hospital Amsterdam UMC, Amsterdam, Netherlands; Lotte Haverman, PhD, Emma Children's Hospital Amsterdam UMC, Amsterdam, Netherlands

**Aims:** Recent measures of implementing social isolation and physical distancing as governmental reactions to the COVID-19 outbreak profoundly impact daily life, including that of children and adolescents. The one on the other day children and adolescents were not allowed to go to school or participate in sports or other social settings. It is therefore relevant to investigate the impact of these measures on psychosocial outcomes in children and adolescents in the general

population. In this study we surveyed how the COVID-19 outbreak impacts on the psychosocial functioning in a sample of Dutch children and adolescents during the first months of the largest public health crises of our time. **Methods:** In April 2020, children and adolescents aged 8-18 years, representative of the Dutch population on key demographics, were asked to complete the following Patient-Reported Outcomes Measurement Information System (PROMIS®) pediatric item banks as computerized adaptive test (CAT): anger, anxiety, depressive symptoms, peer relationships, sleep-related impairment and global health, online using the KLIK PROM portal ([www.hetklikt.nu](http://www.hetklikt.nu)). In addition, parents were asked to complete sociodemographic questions about themselves (age, ethnicity, education level) and their child (age, gender, education level and presence of chronic conditions). Finally, both children and parents answered COVID-19 specific questions such as consequences for employment, school and the atmosphere at home. Using independent sample T-tests, PROMIS COVID-19 T-scores will be compared to normative control data that were collected in the general population pre-COVID (2018;  $n = 1098$ ). **Results:** In total, 1067 children and parents completed all questionnaires. Data management and analyses are currently being carried out. **Conclusion:** Results: will be shown at the conference.

### (119.5) Impact of frontline medical workers' self-leadership on acute stress reaction(ASR) and quality of life during the COVID-19 outbreak in China

Yuan He, MS, Université Paris-Saclay, Paris, France; Yanbo Zhu, PhD, Beijing University of Chinese Medicine, Beijing, China; Haitao Xu, MD, Changzheng Hospital, Second Military Medical University, Shanghai, China; Lin Lin, PhD, China-Japan Friendship Hospital, Beijing, China; Cuiping Xu, PhD, The First Affiliated Hospital of Shandong First Medical University, Jinan, China; Huimei Shi, PhD, Shanghai Ninth People's Hospital, Jinan, China; Hui Zhang, MS, Beijing University of Chinese Medicine, Beijing, China

**Aims:** This study examines how frontline medical workers' self-leadership impacts on their acute stress reaction (ASR) and quality of life (QOL) during the COVID-19 outbreak in China. **Methods:** 187 valid samples of frontline medical workers were collected from 10th February to 16th 2020 through a set of internet-based questionnaires, which includes general information, the Revised Self-Leadership Questionnaire (RSLQ), the Stanford Acute Stress Response Questionnaire (SASRQ), and the WHOQOL-BREF. **Results:** (1)The average scores of the nine dimensions of self-leadership are  $(3.11 \pm 0.86)$  to  $(3.72 \pm 0.62)$  points, which are in the upper-middle level; the score of ASR for COVID-19 was  $(29.27 \pm 25.87)$  points, and 53 respondents appeared with acute stress disorder (ASD) (total SASRQ score  $\geq 40$ , accounting for 28.34%); the total QOL score assessed by WHOQOL-BREF is  $(63.14 \pm 12.53)$  points.(2)The correlation analysis showed that self-punishment ( $r = 0.188$ ,  $p < 0.05$ ) and visualizing successful performance ( $r = 0.167$ ,  $p < 0.05$ ) were positively correlated with the ASR; self-goal setting ( $r = 0.300$ ,  $p < 0.001$ ), self-observation ( $r = 0.244$ ,  $p < 0.001$ ), self-reward ( $r = 0.203$ ,  $p < 0.01$ ), focusing thoughts on natural rewards ( $r = 0.344$ ,  $p < 0.001$ ), self-talk ( $r = 0.160$ ,  $p < 0.05$ ), evaluating beliefs and assumptions ( $r = 0.292$ ,  $p < 0.001$ ) were positively correlated with the QOL.(3)The multiple regression model that further controlled the interrelationship between confounding factors and the 9 dimensions of self-leadership revealed that self-punishment ( $b' = 0.272$ ,  $p = 0.007$ ), visualizing successful performance ( $b' = 0.269$ ,  $p = 0.012$ ) and focusing thoughts on natural rewards ( $b' = -0.301$ ,  $p = 0.035$ ) are the influencing factors of ASR; self-punishment ( $b' = -0.327$ ,  $p = 0.001$ ), focusing thoughts on natural rewards ( $b' = 0.516$ ,  $p = 0.000$ ) were the factors that influence the QOL.

Conclusion: Facing COVID-19, frontline medical workers as the main body of the pandemic prevention and control were in a state of higher stress and lower QOL. Research suggests that medical practitioners can apply self-leadership theory to construct their mental health and positive behaviors, through focusing thoughts on natural rewards as well as alleviating self-punishment and successful performance visualizing, to release acute stress responses and improve quality of life.

## 120: Application of PROMIS

### (120.1) Common patient-reported outcomes across ICHOM Standard Sets—the value of PROMIS

Caroline Terwee, PhD, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, Netherlands; Marloes Zuidgeest, National Health Care Institute, Diemen-Zuid, Netherlands; Harold Vonkeman, PhD, Medisch Spectrum Twente and University of Twente, Enschede, Netherlands; David Cella, PhD, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Lotte Haverman, PhD, Emma Children's Hospital Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands; Leo Roorda, PhD, Amsterdam Rehabilitation Research Center | Reade, Amsterdam, Netherlands

Aims: The International Consortium for Health Outcomes Measurement (ICHOM) develops condition-specific Standard Sets of outcomes to be measured in clinical practice for value-based health-care evaluation. There are, however, large differences and inconsistencies between sets in selected patient-reported outcomes (PROs), terms and definitions used, and recommended patient-reported outcome measures (PROMs), even for the same PROs, which threatens the validity and practical applicability of the ICHOM Standard Sets. It would be ideal if common PROs would be named and defined similarly and measured with the same PROMs across conditions. PROMIS® offers an evidence-based conceptual framework of commonly relevant PROs and validated PROMs that are applicable across patient populations and medical specialties. The aim of this study was to identify shared PROs across ICHOM Standard Sets and to examine to what extent these PROs can be measured with PROMIS. Methods: All individuals PROs and recommended PROMs were extracted from all available ICHOM Standard Sets in January 2020. Similar PROs were categorized into unique PRO concepts. Subsequently, it was examined which of these PRO domains can be measured with PROMIS. Results: In 28 ICHOM Standard Sets, 182 PROs were identified. A total of 96 different PROMs are recommended for measuring these PROs. The 182 PROs were categorized into 21 unique PRO concepts. More than half (12/21) of these PRO concepts (covering 74% of the 182 PROs and 79% of the 96 PROMs) can be measured with a PROMIS measure. Furthermore, inconsistencies were found in the selected PROs and PROMs across Standard Sets. It is unclear why some PROs are included in some Standard Sets, but not in others. Conclusion: Considerable overlap was found in PROs across ICHOM Standard Sets, and large differences in terms used and recommended PROMs, even for the same PROs. Inconsistencies in the selected PROs and PROMs across Standard Sets questions the validity of the Standard Sets. We recommend a more universal and standardized approach to PRO and PROM selection, using a common measurement system such as PROMIS, to improve the validity of outcome measurements in clinical practice, and facilitate benchmarking, learning and improve quality of care across patient groups.

### (120.2) Estimating Power for Clinical Trials with PROMIS endpoints using Item Response Theory

Jinxiang Hu, University of Kansas Medical Center, Kansas City, Missouri, United States; Yu Wang, University of Kansas Medical Center, Kansas City, Missouri, United States

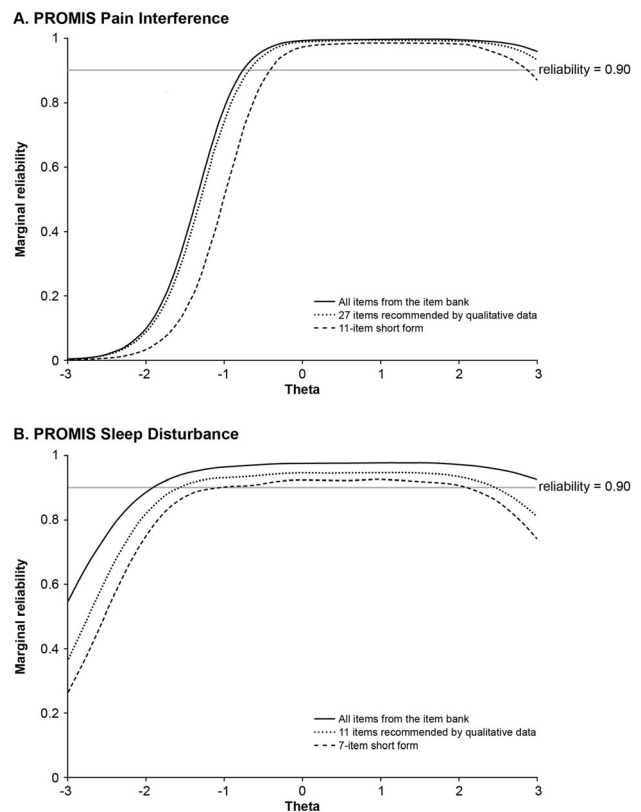
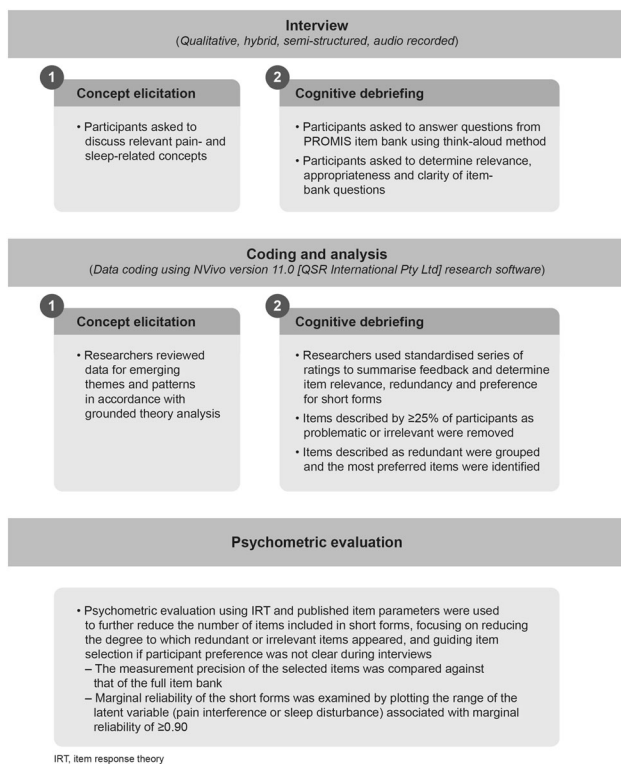
Aims: Patient-reported outcomes (PRO) are important in patient-centered health outcomes research, epidemiological studies, quality of life (QOL) studies, and clinical trials. Patient-Reported Outcomes Measurement Information System (PROMIS) is a set of standardized, generic PRO questionnaires developed for clinical and research purpose. In clinical trials, it is crucial to estimate power to avoid waste of resources while still able to detect the treatment effect. However, for clinical trials with PRO as end points, Classical Test Theory (CTT) using observed scores (e.g., total/average scores) are routinely used for power estimation. The purpose of this project is to provide guidance for power and sample size estimate for clinical trials with PROMIS measures as endpoints using IRT, especially for early stage trials. Methods: Motivated from PROMIS depression scales (4a, 6a, 8a), we conducted a simulation study in order to estimate power differences between IRT- and CTT-based scoring for a two-armed prospective randomized clinical trial (control vs active arm). We simulated data using various sample size, allocation ratio, number of items, effect sizes, and missing data. Three models were fit to each simulation: IRT with MLE, IRT with Bayesian estimator, and CTT. Results: Our results showed missing data, effect size, and sample size are important indicators of IRT power. Number of items is not significantly associated with power. Conclusion: For rare diseases or early stage trials, it is important to use IRT framework for accurate power estimation. IRT and CTT both provide good power with large sample size and effect size. Future work can examine the IRT power for detecting change over time and non-normal distribution of latent scores.

### (120.3) Qualitative and quantitative psychometric approaches to evaluate the PROMIS Pain Interference and Sleep Disturbance item banks for use in patients with rheumatoid arthritis

Brandon Becker, GlaxoSmithKline, Upper Providence, Pennsylvania, United States; Kimberly Raymond, Optum, Johnston, Rhode Island, United States; Carol Hawkes, GlaxoSmithKline, London, United Kingdom; April Mitchell Foster, Optum, Johnston, Rhode Island, United States; Andrew Lovley, Optum, Johnston, Rhode Island, United States; Cory Saucier, Optum, Johnston, Rhode Island, United States; Avery A. Rizio, Optum, Johnston, Rhode Island, United States; Jakob Bue Bjorner, Optum, Copenhagen, Denmark; Mark Kosinski, Optum, Johnston, Rhode Island, United States

Aims: Patient-reported outcome measures focusing on pain severity may provide limited insight into the impact of rheumatoid arthritis (RA) on patients' lives. The wide-ranging PROMIS Pain Interference and Sleep Disturbance item banks may provide RA-relevant content for reporting pain and sleep in clinical trials. This study evaluated the content validity of the PROMIS Pain Interference and Sleep Disturbance item banks in a population of patients with moderate-to-severe RA to develop RA-specific short forms that demonstrated content validity and adequate measurement precision. Methods: Qualitative, semi-structured, hybrid interviews comprising concept elicitation and cognitive debriefing methods were conducted with patients with moderate-to-severe RA (Fig. 1). Findings from the interviews were used to identify relevant candidate items for short forms.

Psychometric evaluation, which employed established item response theory (IRT)-derived item parameters, was used to develop final recommended short forms with high measurement precision across the full range of pain interference and sleep disturbance. Results: Thirty-two adults with RA were interviewed. Participants reported that pain and sleep disruptions from RA impacted multiple aspects of daily life. Cognitive debriefing revealed both item banks to be easily understood and the 7-day recall period to be appropriate. In total, 27/40 pain interference and 11/27 sleep disturbance items were identified by participants as being most appropriate for capturing the pain- and sleep-related impacts of RA. Psychometric evaluation identified 11 Pain Interference and 7 Sleep Disturbance items for inclusion in short forms. The short forms were associated with marginal reliability  $\geq 0.90$  across a broad measurement range, indicating that they can detect differences or changes in scores at the individual patient level with a high degree of certainty (Fig. 2). Conclusion: These findings support a mixed-method approach to developing RA-specific short forms for PROMIS Pain Interference and Sleep Disturbance item banks in an RA population. The recommended short forms demonstrated content validity, reliability and strong measurement precision. Future research should confirm the cross-sectional and longitudinal psychometric properties of these short forms in patients with RA and identify meaningful within-patient change thresholds. Study funded by GSK (206578/HO-18-17125). Medical writing support provided by Eithne Maguire, Fishawack Indicia Ltd, UK, funded by GSK.



#### (120.4) Validation of the pediatric Patient-Reported Outcomes Measurement Information System (PROMIS®) Anger Item bank in the general Dutch population

Maud van Muilekom, MSc, Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Psychosocial Department, Amsterdam, Netherlands; Michiel Luijten, MSc, Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Psychosocial Department, Amsterdam, Netherlands; Raphaële van Litsenburg, PhD, Princess Máxima Center for Pediatric Oncology, Utrecht, Netherlands; Lotte Haverman, PhD, Emma Children's Hospital, Amsterdam UMC, University of Amsterdam, Psychosocial Department, Amsterdam, Netherlands; Caroline Terwee, PhD, Amsterdam UMC, Vrije Universiteit, Epidemiology and Biostatistics, Amsterdam, Netherlands; Martha Grootenhuis, PhD, Princess Máxima Center for Pediatric Oncology, Amsterdam, Netherlands

**Aims:** To reduce the burden of completing questionnaires, Patient-Reported Outcomes Measurement Information System (PROMIS®) item banks, facilitating Computerized Adaptive Testing (CAT), were introduced. This study aimed to assess validity and reliability of the pediatric Dutch-Flemish v2.0 PROMIS Anger item bank, short form and CAT in the general Dutch population and to provide normative data. **Methods:** Children (8–18y,  $n = 1328$ ), representative of the Dutch population, were asked to complete the PROMIS v2.0 Anger item bank (PROMIS Anger, Nitens = 9) and Pediatric Quality of Life Inventory (PedsQLTM, Nitens = 23). The following assumptions were assessed: Unidimensionality by CFA (CFI  $> .95$ , TLI  $> .95$ , RMSEA  $< .10$ ), local independence by residual correlations ( $r < .20$ ), and monotonicity by Mokken analysis ( $H > .50$ ,  $Hi > .30$ ). A graded response model (GRM) was fit to the data and structural validity was assessed by looking at item-fit statistics ( $S-X2$ ,  $p < .001 = \text{misfit}$ ). Standard error of measurement (SEM) was used

to calculate reliability ( $SEM < .32 = .90$  reliability). Relative efficiency was calculated ( $1 - SEM^2/nitems$ ) to compare how good the PROMIS Anger forms and PedsQL emotional functioning (EF) subscale perform relative to the amount of items administered. Dutch mean T-scores based on the US model were calculated to provide normative data. Correlations were assessed between PROMIS Anger and PedsQL subscales based on US parameters, where a moderately high correlation ( $r > .50$ ) was expected between T-scores and EF subscale and lower correlations ( $\Delta r > .10$ ) for other subscales. Results: Data from 527 children (response-rate = 39.7%) was used for analyses. All assumptions were met. Structural validity of the GRM model was sufficient as no items displayed misfit ( $S-X^2 = 22.9-40.3$ ,  $pS-X^2 > .001$ ). The model provided reliable measurements at the population mean and  $> 2SD$  in the clinically relevant direction. CAT outperformed all other measures in efficiency. Dutch mean T-score was 44.20 ( $SD = 11.39$ ). Finally, PROMIS Anger correlated moderately high ( $r = .64$ ) with the PedsQL EF subscale and lower with other subscales. Conclusion: The pediatric PROMIS Anger item bank was successfully validated for use within the Dutch population and normative data are now available. It was therefore implemented in the KLIK PROM portal ([www.hetklickt.nu](http://www.hetklickt.nu)) as CAT for use in clinical practice. To explain PROMIS and CAT to patients and facilitate use in clinical practice, an educational video has been developed, which will be shown during the conference.

**(120.5) Standardizing the measurement of fatigue in Multiple Sclerosis: extensive evidence on the content validity and measurement properties of the PROMIS Fatigue (MS) short form 8b**

Paul Kamudoni, Global Evidence & Value Development, Merck Healthcare KgaA, Darmstadt, Germany; Jeffrey Johns, Institute of Medicines Development, Cardiff, United Kingdom; Karon Cook, Feral Scholars, Broadus, Texas, United States; Rana Salem, Department of Rehabilitation Medicine, University of Washington, Seattle, Washington, United States; Sam Salek, School of Life and Medical Sciences, University of Hertfordshire, Cardiff, United Kingdom; Jana Raab, Global Evidence & Value Development—R&D, Merck Healthcare KgaA, Cardiff, Germany; Rod Middleton, UK MS Register, Swansea Medical School, Swansea, UK, Swansea, United Kingdom; Christian Henke, Global Evidence & Value Development—R&D, Merck Healthcare KgaA, Darmstadt, Darmstadt, Germany; Dagmar Amtmann, Department of Rehabilitation Medicine, University of Washington, Seattle, Washington, United States

**Aims:** The aim of this research was to generate additional validity evidence for the PROMIS Fatigue (MS) 8b, including responsiveness and meaningful score interpretation criteria, across the US and UK populations. **Methods:** A mixed-methods, two-step design was followed in this research. Step 1 involved cognitive debriefing (CD) the PROMIS-Fatigue (MS) 8b with MS patients from the US ( $n = 29$ ), to confirm content validity. Step 2 included a cross-sectional study in two tertiary MS centers in the US ( $n = 296$ ) [US sample] and a 96-week longitudinal study in the UK MS register cohort (still ongoing) ( $n = 384$ ) [UK sample] to evaluate measurement properties. Psychometric analyses examined reliability, validity, responsiveness and meaningful score change criteria (over 52 weeks of follow-up in the UK sample). Results: The CD interviews confirmed the comprehensiveness and appropriateness of the PROMIS-Fatigue(MS)8b in covering fatigue experience and impacts related to MS. The 7-day recall period and 5-point Likert scale were well understood and judged to be appropriate. In the observational studies (UK MS sample, US Sample), the mean PROMIS-Fatigue(MS) T-score at baseline was 57.4–59.9. Internal consistency (Cronbach's  $\alpha > 0.9$ ) and

test-retest reliability at 5–27 days follow-up (Intraclass correlations  $\geq 0.9$ ). PROMIS Fatigue (MS) T-scores significantly discriminated (i.e.,  $p < 0.001$ ) between severe and mild-moderate levels of fatigue (PROMIS GHS fatigue global question, Fatigue Severity Scale), physical health (GHS GPH summary score, GHS physical health global question). PROMIS Fatigue (MS) 8b scores were sensitive to worsening ( $ES = -0.44/-0.22$ ) as well improvements ( $ES = 0.5/0.34$ ) in fatigue over a 52-week follow-up duration (anchors: GHS fatigue global question/Fatigue Severity Scale-FSS). Mean score change was 3.86/3.46 in the minimally improving group and 3.37/1.17 in the minimal worsening groups, respectively. SEM was 2.8–3. Thus, we propose a score change of 3.4–3.9 as cut-off for important individual-level improvement or worsening. Conclusion: This research extends the evidence supporting the content validity and the robust psychometric performance of the PROMIS Fatigue (MS) 8b across populations (USA, UK). Importantly, data supporting the measure's integration in clinical practice and research, including meaningful score interpretation, are now available.

## Oral Briefs

### 201: Clinical research I

**(B201.1) Associations between PROMIS pediatric measures and physical and neurocognitive performances among childhood cancer survivors: a report from the PEPR Consortium**

Jin-ah Sim, PhD, Department of Epidemiology & Cancer Control, St. Jude Children's Research Hospital, Memphis, Tennessee, United States; Justin N. Baker, MD, Department of Oncology, St. Jude Children's Research Hospital, Memphis, Tennessee, United States; Kirsten K. Ness, PhD, Department of Epidemiology & Cancer Control, St. Jude Children's Research Hospital, Memphis, Tennessee, United States; Kevin R. Krull, PhD, Department of Epidemiology & Cancer Control, St. Jude Children's Research Hospital, Memphis, Tennessee, United States; Deo Kumar Srivastava, PhD, Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, Tennessee, United States; Christopher B. Forrest, MD, PhD, Department of Pediatrics, Children's Hospital of Philadelphia, Memphis, Pennsylvania, United States; Leslie L. Robison, PhD, Department of Epidemiology & Cancer Control, St. Jude Children's Research Hospital, Memphis, Tennessee, United States; Melissa M. Hudson, MD, Department of Oncology, St. Jude Children's Research Hospital, Memphis, Tennessee, United States; I-Chan Huang, PhD, Department of Epidemiology & Cancer Control, St. Jude Children's Research Hospital, Memphis, Texas, United States

**Aims:** PROMIS measures are widely used to assess patient-reported outcomes (PROs) in children and youth. However, validity of the PROMIS measures in childhood cancer survivors who experience a high burden of physical and/or neurocognitive late effects is understudied. We aimed to evaluate the validity of PROMIS measures by comparing clinically assessed physical and neurocognitive performances among child and youth survivors. **Methods:** Participants included 293 individuals who took part in the Pediatric Patient-Reported Outcomes in Chronic Diseases Consortium (PEPR). Inclusion criteria included survivors of pediatric malignancies who were 8–18.9 years of age at the time of study. PROs included PROMIS depression, fatigue, pain intensity, sleep-related impairment, perceived cognitive dysfunction, and mobility measures. Physical performance was tested using a pediatric-modified total neuropathy measure. Neurocognitive performance was tested using academic performance, attention, memory, processing speed, and executive function batteries. Survivors were classified as having physical impairment if scored  $\geq 5$  on the physical performance test, and as

having neurocognitive impairment if 40% of specific batteries were impaired. They were further classified as (1) no impairment, (2) impairment on physical performance alone, (3) impairment on neurocognitive performance alone, and (4) impairment on both. Multivariate linear regressions tested the associations of physical/neurocognitive impairment with PRO domain scores adjusting for age, sex, and years after diagnosis. Results: Mean age of survivors was 14.2 years (SD = 2.9); 50.2% were male; 60.2% had neither physical nor neurocognitive impairment, 22.0% had neurocognitive impairment, 11.0% had physical impairment, and 6.8% had both. Compared to having no impairment, survivors having both physical and neurocognitive impairments had more depression ( $\beta = 9.71$ ,  $p < 0.001$ ), fatigue ( $\beta = 8.57$ ,  $p = 0.002$ ), sleep-related impairment ( $\beta = 6.03$ ,  $p = 0.01$ ), perceived cognitive dysfunction ( $\beta = 8.06$ ,  $p < 0.001$ ), and poorer mobility ( $\beta = 7.98$ ,  $p < 0.001$ ) domains. Compared to having no impairment, survivors having physical impairment alone had more fatigue ( $\beta = 6.66$ ,  $p = 0.003$ ), sleep-related impairment ( $\beta = 5.34$ ,  $p = 0.004$ ), perceived cognitive dysfunction ( $\beta = 3.78$ ,  $p = 0.02$ ), and poorer mobility ( $\beta = 4.49$ ,  $p = 0.001$ ). However, compared to having no impairment, survivors having neurocognitive impairment alone had comparable PROs (Ps > 0.05). Conclusion: PROs correlated well with survivors having both physical and neurocognitive impaired performances or physical impaired performance alone. PROs should be used to screen at-risk survivors who may benefit from interventions

### (B201.2) Major stressors in the year prior to disease onset affects RA characteristics at baseline and 12 months: results from the Canadian Early Arthritis Cohort (CATCH)

Nicole Andersen, McGill University, Montreal, Quebec, Canada; Orit Schieir, University of Toronto, Montreal, Quebec, Canada; Marie-France Valois, McGill University, Montreal, Quebec, Canada; Kathleen Andersen, Johns Hopkins University, Baltimore, Maryland, United States; Gilles Boires, Université de Sherbrooke, Sherbrooke, Quebec, Canada; Louis Bessette, Université de Laval, Sherbrooke, Quebec, Canada; Diane Tin, Southlake Regional Hospital, Newmarket, Ontario, Canada; Carter Thorne, Southlake Regional Hospital, Newmarket, Ontario, Canada; Janet Pope, Western University, London, Ontario, Canada; Carol Hitchon, University of Manitoba, Winnipeg, Manitoba, Canada; Glen Hazlewood, University of Manitoba, Winnipeg, Manitoba, Canada; Glen Hazlewood, University of Calgary, Calgary, Alberta, Canada; Edward Keystone, University of Toronto, Toronto, Ontario, Canada; Vivian Bykerk, Weill Cornell Medical College Hospital for Special Surgery, New York, New York, United States; Susan Bartlett, McGill University, Montreal, Quebec, Canada

**Aims:** Although many people with RA link disease onset to recent stressful life events, results from retrospective studies are unclear. The objectives were to describe the incidence of major stressors (+STRESS) in year prior to diagnosis and compare characteristics and patient-reported outcomes (PROs) of newly diagnosed RA patients with and without +STRESS at 0 and 12 months. **Methods:** Data were from early RA patients (symptoms < 1 year) enrolled in Canadian Early Arthritis Cohort (CATCH) from 2007 to 2017, who had  $\geq 12$  months of follow-up. Patients reported major psychological (death, divorce/separation, family, financial, other) and physical (motor vehicle accident, surgery, major illness/infection, other) stressors in previous year. We used independent t-tests and Chi square to compare characteristics by stressors at baseline, and multivariable regression to examine the impact of +STRESS on disease activity and PROs at 1 year, adjusting for age, sex, education, fibromyalgia, and SJC. **Results:** The 1933 adults were mostly female (72%), with a mean (SD) age of 55 (15) years. 52% reported 1 + stressors in

previous year; family (48%), financial stress (36%), death (35%), surgery (28%), and major illness (26%) were the most common stressors. Patients with +STRESS were more likely to be women, younger, have more comorbidities including fibromyalgia, and higher mean DAS28. Patients with +STRESS also had significantly higher mean pain, fatigue, depression, sleep disturbance, patient global, and HAQ scores at baseline. At 1 year, SJC and the proportion in DAS28 REM was similar between groups. However, PROs (pain, HAQ, Fatigue, Pt Global, Depression, Poor Sleep) remained higher in +STRESS, with evidence of an additive effect for number of stressors and having both physical and psychological stressors (Table). The greatest impacts were on mood, sleep disturbance, and fatigue. **Conclusion:** In this pan-Canadian early RA cohort, more than half reported 1 + stressful life events in year prior to diagnosis. Individuals reporting major stressors had significantly worse pain, patient global, disability, depression, fatigue, and sleep disturbance at diagnosis; 1 year later, though disease activity was similar between groups, the effects of +STRESS on PROs persisted. Early RA patients with recent major stressors may benefit from emotional support to optimize how they feel and function.

Mean (SD) or N (%)	No Stress (N=928; 48%)	Physical (N=131; 7%)	Psychological (N=658; 34%)	Both (N=216; 11%)
Age	56 (15)	56 (15)	53 (14)	52 (15)
Women	622 (67%)	82 (63%)	512 (78%)	174 (81%)
College Education	464 (50%)	76 (58%)	345 (52%)	126 (58%)
Rheum Dis Comorbid Index	1.1 (1.2)	1.4 (1.4)	1.1 (1.3)	1.4 (1.3)
OA or Spinal pain	168 (18%)	35 (27%)	117 (18%)	55 (25%)
Fibromyalgia diagnosis	15 (2%)	2 (2%)	13 (2%)	11 (5%)
Symptom duration (months)	5.6 (3.0)	5.7 (3.0)	5.9 (3.0)	5.9 (3.0)
DAS28 – mean	5.0 (1.4)	5.1 (1.5)	5.0 (1.5)	5.2 (1.4)
MTX $\pm$ csDMARDs	679 (73%)	100 (76%)	489 (74%)	166 (77%)
Oral Steroids	295 (32%)	40 (31%)	215 (33%)	55 (25%)
Pain (0-10)	5.3 (2.8)	5.5 (2.9)	5.7 (2.8)	6.2 (2.8)
HAQ-DI	1.0 (0.7)	1.2 (0.7)	1.1 (0.7)	1.3 (0.7)
Fatigue (0-10)	4.7 (3.1)	5.0 (3.0)	5.7 (2.9)	5.9 (2.9)
Patient Global (0-10)	5.6 (2.9)	6.0 (2.9)	6.0 (2.9)	6.4 (3.0)
Depression (SF12 MCS < 45.6)	329 (35%)	54 (41%)	356 (54%)	123 (57%)
Poor sleep (0-10)	4.5 (3.4)	4.8 (3.3)	5.3 (3.2)	6.0 (3.1)

Effect of stress in year prior to diagnosis on outcomes at 12 months (N=1933).\*

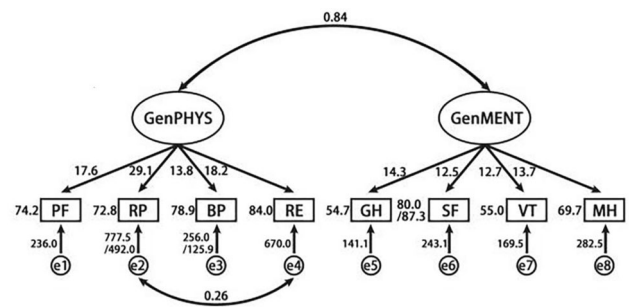
	Pain	Fatigue	Patient Global	Depression	Poor Sleep	HAQ-DI	DAS28 REM	Swollen Joints
Any Stress	0.27 (0.02, 0.53)	0.39 (0.12, 0.66)	0.25 (0.01, 0.51)	1.55 (1.24, 1.95)	0.49 (0.21, 0.76)	0.05 (0.01, 0.11)	1.02 (0.82, 1.27)	0.12 (-0.19, 0.42)
Total stressors	0.24 (0.12, 0.37)	0.32 (0.18, 0.45)	0.24 (0.11, 0.37)	1.30 (1.17, 1.45)	0.37 (0.24, 0.51)	0.05 (0.02, 0.08)	1.00 (0.90, 1.12)	0.05 (-0.10, 0.20)
0 (n=928, 48%)	REF	REF	REF	REF	REF	REF	REF	REF
1 (n=575, 30%)	-0.02 (-0.31, 0.28)	0.14 (-0.17, 0.46)	-0.04 (-0.34, 0.26)	1.31 (1.00, 1.71)	0.18 (-0.14, 0.50)	0.00 (-0.07, 0.06)	1.12 (0.87, 1.43)	0.05 (-0.29, 0.42)
2 (n=295, 15%)	0.54 (0.17, 0.92)	0.59 (0.19, 0.99)	0.54 (0.16, 0.92)	1.79 (1.29, 2.48)	0.80 (0.40, 1.20)	0.10 (0.02, 0.18)	0.99 (0.72, 1.35)	0.20 (-0.25, 0.65)
3 (n=88, 5%)	1.06 (0.45, 1.66)	0.73 (0.08, 1.39)	0.81 (0.19, 1.44)	1.90 (1.15, 3.16)	0.72 (0.06, 1.38)	0.18 (0.04, 0.32)	0.90 (0.53, 1.52)	0.23 (-0.50, 0.96)
4+ (n=47, 2%)	0.75 (-0.11, 1.61)	1.78 (0.84, 2.72)	1.02 (0.12, 1.91)	3.45 (1.70, 6.99)	2.10 (1.16, 3.04)	0.19 (0.00, 0.38)	0.48 (0.21, 1.03)	0.03 (-0.99, 1.05)

\*Multivariable linear regression coefficient adjusted for age, sex, education, baseline fibromyalgia diagnosis and swollen joints

### (B201.3) Detecting response shift in quality of life measurement among patients with hypertension using then-test and structural equation modeling

Hao Chen, MSc, Department of Social Medicine of School of Public Health, Zhejiang University School of Medicine, Hangzhou, China; Lin Zhu, PhD, Department of Social Medicine of School of Public Health, Zhejiang University School of Medicine, Hangzhou, China; Rui Zhou, PhD, Department of Social Medicine of School of Public Health, Zhejiang University School of Medicine, Hangzhou, China; Panpan Liu, MSc, Department of Public Health, Hangzhou Medical College, Hangzhou, China; Xiaoyang Lu, MSc, Department of Pharmacy of the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China; Donald Patrick, PhD, Department of Health Services, University of Washington, Hangzhou, Washington, United States; Todd Edwards, PhD, Department of Health Services, University of Washington, Seattle, Washington, United States; Hongmei Wang, PhD, Department of Social Medicine of School of Public Health, Zhejiang University School of Medicine, Hangzhou, China

**Aims:** Outcomes derived from longitudinal self-reported quality of life measures can be confounded by response shift (RS). The primary aim of this study was to detect RS among patients with hypertension attending a community-based disease management program, and to explore possible predictors of the occurrence of RS. The second aim was to test the agreement of the then-test approach and the structural equation modeling (SEM) approach to RS detection. **Methods:** 240 consecutive consulting or followed up patients with diagnosed hypertension were recruited in a community health service center. A SF-36 instrument was self-administrated at baseline (pre-test), and the other two sets of SF-36 instruments were completed 4 weeks later to elucidate how respondents perceived their health status at 4 weeks ago (then-test) and how they felt currently (post-test). RS was assessed by the then-test approach and a 4-step SEM approach. By integrating the then-test with SEM, underlying assumptions of then-test design were examined. Partial correlations and hierarchical multiple regressions were used to detect predictors of RS. **Results:** Data from 211 (87.9%) patients were eligible for analyses. Mean age of the participants was 66.1 years (SD 10.8). 65.4% respondents had experienced recalibration in at least one scale. More than half participants reported recalibration in BP, GH, SF, VT, MH scales. Recalibration at the group level was detected in RP, BP and SF scales. After accounting for RS effects, the score changes in the quality of life of respondents became generally insignificant, except for slight improvement in the PF scale. The consistency assumption between then-test and post-test was verified, while recall bias caused divergent outcomes from two approaches in the PF scale. The regressions models found that education was a positive predictor of RS in nearly all scales, while older age and severe illness experience were negative predictors of RS in physical related scales and mental related scales respectively. **Conclusion:** Recalibration existed among patients with hypertension, and had effect on observed HRQOL change, which suggests RS should be considered in hypertension researches with longitudinal HRQOL data. Results: showed generally good agreements of the then-test approach and the SEM approach, and combined detection methods are recommended.



**(B201.4) The impact of circadian rhythm on the clustering of fatigue, depression and insomnia in breast cancer survivors: a latent class analysis**

Belle de Rooij, PhD, Netherlands Comprehensive Cancer Organisation (IKNL), Eindhoven, Netherlands; Imogen Ramsey, Msc, Rosemary Bryant AO Research Centre, University of South Australia Cancer Research Institute, University of South Australia, Adelaide, Australia; Felix Clouth, Msc, Department of Methodology and Statistics, Tilburg University, Tilburg, Netherlands; Brigid M Lynch, PhD, Cancer Epidemiology Division, Cancer Council Victoria, Melbourne, Australia; Jeff K. Vallance, PhD, Faculty of Health Disciplines, Athabasca University, Athabasca, Alberta, Canada; Nadia Corsini, PhD, Rosemary Bryant AO Research Centre, University of South Australia Cancer Research Institute, University of South Australia, Athabasca, Australia; Terry Boyle, PhD, Australian Centre for Precision Health, School of Health Sciences, University of South Australia Cancer Research Institute, Adelaide, Australia

**Aims:** Breast cancer survivors’ day-and-night bodily processes (i.e., circadian rhythms) are often misaligned due to the cancer and/or cancer treatments. Through neuroendocrine pathways, circadian rhythm may influence the prevalence and clustering of behavioral symptoms that often co-occur in breast cancer survivors including fatigue, insomnia and depression. We aimed to (1) identify subgroups of breast cancer survivors based on the severity of symptoms of fatigue, insomnia and depression, and (2) assess whether circadian rhythm is associated with these subgroups. **Methods:** Among 265 breast cancer survivors, circadian rhythm (circadian phase, amplitude and stability; Horne-Ostberg Morningness/Eveningness scale and Circadian Type Inventory) was assessed at 3–4 months after diagnosis (T0), and symptoms of fatigue (FACIT-Fatigue), depression (PHQ9) and insomnia (PSQI) were assessed after 2–3 years (T1) and 6–8 years (T2). We applied latent class analysis to classify survivors in unobserved groups (‘classes’) based on self-reported symptoms of fatigue, depression and insomnia at T1. The impact of circadian rhythm on class allocation was assessed using multinomial logistic regression analysis. Changes in class allocation from T1 to T2 were assessed using latent transition models. **Results:** We identified 3 latent classes: (1) low symptom burden (38%), (2) moderate symptom burden (41%) and high symptom burden (21%). Survivors with high symptom burden were younger, more often unmarried, unemployed and less often received chemotherapy. After adjustment for covariates, survivors with a late circadian phase (‘evening types’) were more likely to have medium (OR 3.38, 95% CI 2.62–4.14) or high (OR 5.12, 4.16–6.08) symptom burden compared to survivors with an early circadian phase (‘morning types’). Further, survivors with a languid circadian amplitude were more likely to have medium (OR 2.44, 95% CI 1.71–3.18) or high (OR 5.56, 95% CI 4.64–6.49) symptom burden compared to survivors with a vigorous circadian amplitude. The majority of survivors with moderate or high symptom burden at T1 had persistent symptom burden at T2 (59% and 64% respectively). **Conclusion:** A delayed circadian phase and languid

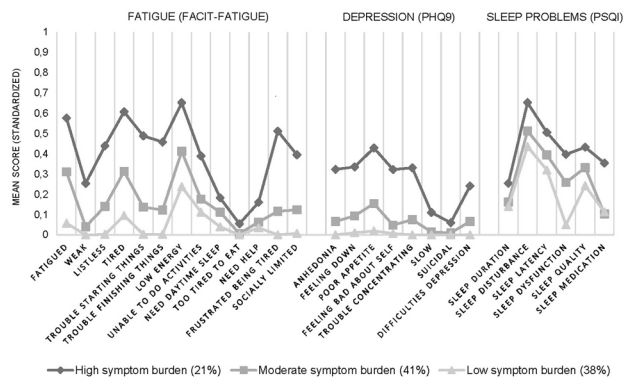
Significance tests of RS for SEM approach

Scale	RS	Significance test	
		$\chi^2_{SBDiff}(df=1)$	Prob.
PF			
RP	Non-uniform recalibration	8.84	0.003
BP	Non-uniform recalibration	17.41	<0.001
RE			
GH			
SF	Uniform recalibration	22.98	<0.001
VT			
MH			

Pairwise comparisons of pre-test, then-test and post-test HRQoL scores and effect-sizes of observed change, RS and true change

Scale	pre-test mean (SD)	then-test mean (SD)	post-test mean (SD)	Effect-sizes		
				Observed change <sup>a</sup>	RS <sup>b</sup>	true change <sup>c</sup>
PF	74.2(24.5)	77.9(24.3)	77.5(24.1)	0.14*	0.15*	-0.01
RP	72.0(41.4)	80.1(37.3)	79.5(37.7)	0.18*	0.20*	-0.02
BP	77.8(21.4)	83.9(18.0)	82.8(18.1)	0.23*	0.28*	-0.06
RE	84.3(33.7)	86.7(32.0)	88.3(30.2)	-0.12	-0.06	-0.06
GH	54.3(19.5)	52.7(17.0)	52.4(16.7)	0.12	0.07	0.05
SF	80.5(20.4)	85.0(19.5)	85.1(18.8)	0.22*	0.22*	0.00
VT	55.1(18.5)	53.9(16.1)	52.9(16.1)	-0.10	-0.09	-0.02
MH	69.9(20.6)	67.7(20.3)	67.3(21.1)	-0.13	-0.11	-0.02

circadian amplitude after breast cancer treatment was associated with more symptoms of fatigue, depression and insomnia at follow-up. Psycho-oncological or pharmaceutical interventions that expedite and stabilize survivors' sleep/wake cycle may help alleviate breast cancer survivors' symptom burden.



#### (B201.5) Validation of the PROMIS® pediatric item banks Anxiety and Depressive Symptoms in a general Dutch population

Leonie Klaufus, MSc, Public Health Service Amsterdam, Department of Epidemiology, Health Promotion, and Health Care Innovation, Amsterdam, Netherlands; Michiel Luijten, MSc, Amsterdam UMC, Emma Children's Hospital, Psychosocial Department, Amsterdam, Netherlands; Eva Verlinden, PhD, Public Health Service Amsterdam, Department of Epidemiology, Health Promotion, and Health Care Innovation, Amsterdam, Netherlands; Marcel van der Wal, PhD, Public Health Service Amsterdam, Department of Epidemiology, Health Promotion, and Health Care Innovation, Amsterdam, Netherlands; Caroline Terwee, PhD, Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Epidemiology and Biostatistics, Amsterdam Public Health research institute, Amsterdam, Netherlands; Pim Cuijpers, PhD, Vrije Universiteit Amsterdam, Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health research institute, Amsterdam, Netherlands; Mai Chinapaw, PhD, Amsterdam UMC, Vrije Universiteit Amsterdam, Department of Public and Occupational Health, Amsterdam Public Health research institute, Amsterdam, Netherlands; Lotte Haverman, PhD, Amsterdam UMC, Emma Children's Hospital, Psychosocial Department, Amsterdam, Netherlands

**Aims:** This study aims to validate the Dutch-Flemish PROMIS pediatric item banks v2.0 Anxiety and Depressive Symptoms in a general Dutch population. **Methods:** Participants ( $n = 2893$ , aged 8–18), recruited by two certified internet panel agencies, completed the PROMIS pediatric item banks v2.0 Anxiety and Depressive Symptoms online. Both item banks were assessed on unidimensionality, local dependence, monotonicity, Graded Response Model (GRM) item fit, and differential item functioning (DIF) across gender, age groups, region, ethnicity, and language. The PROMIS pediatric Anxiety and Depressive Symptoms short forms 8a and simulated computerized adaptive testings (CATs) were assessed on reliability and construct validity compared to the Revised Child Anxiety and Depression Scale short version (RCADS-22) subscales. **Results:** The PROMIS pediatric item banks v2.0 Anxiety and Depressive Symptoms showed sufficient unidimensionality (Omega  $H = 0.83, 0.95$ ; ECV = 0.79, 0.93, respectively), local independence (residual correlations  $< 0.2$ ), and monotonicity ( $H = 0.61, 0.69$ , respectively). Both

item banks showed sufficient GRM item fit ( $S-X^2 p$  value  $< 0.001$ ), except for the Depressive Symptoms items 2697R1r "I wanted to be by myself," 7010 "I felt sad for no reason," and 9001r "I felt too sad to eat." No DIF was found for gender, age groups, region, ethnicity, and language, except for the Depressive Symptoms items 2697R1r "I wanted to be by myself" and 488R1r "I could not stop feeling sad" that showed uniform DIF for language (McFadden pseudo  $R^2$  change  $> 2\%$ ). Based on US parameters, the PROMIS pediatric Anxiety and Depressive Symptoms short forms 8a showed a reliability of  $> 0.90$  in 2% and 34%, and the CATs in 26% and 41% of the participants, respectively. Both short forms and CATs revealed high positive correlations ( $r > 0.70$ ) with the corresponding RCADS-22 subscales and slightly lower correlations with the non-corresponding RCADS-22 subscales ( $r \leq 0.70$ ) as expected. **Conclusion:** The Dutch-Flemish PROMIS pediatric item banks v2.0 Anxiety and Depressive Symptoms show sufficient psychometric properties, with the exception of four Depressive Symptoms items that show DIF for language or poor GRM item fit; the short forms 8a and CATs seem valid, but reliable for a small percentage of children.

#### (B201.6) An international qualitative study informing the development of a patient-reported outcome instrument for adults receiving gender-affirming treatments: the GENDER-Q

Manraj Kaur, PhD, Brigham and Women's Hospital, Boston, Massachusetts, United States; Shelby Deibert, McMaster University, Hamilton, Ontario, Canada; Thomas Satterwhite, Align Surgical Associates Inc., San Francisco, California, United States; Richard Santucci, Align Surgical Associates Inc., San Francisco, California, United States; Giancarlo McEvenue, Top Aesthetics, Fort Lauderdale, Florida, United States; John Semple, Women's College Hospital, Fort Lauderdale, Ontario, Canada; Natasha Johnson, McMaster Children's Hospital, Hamilton, Ontario, Canada; Shane Morrison, University of Washington Medical Center, Seattle, Washington, United States; Margriet Mullender, Amsterdam University Medical Center, Amsterdam, Netherlands; Lotte Poulsen, Odense University Hospital, Odense, Denmark; Maeghan Ross, Odense University Hospital, Odense, Denmark; Maeghan Ross, Amsterdam University Medical Center, Amsterdam, Netherlands; Andrea Pusic, Brigham and Women's Hospital, Boston, Massachusetts, United States; Anne Klassen, McMaster University, Hamilton, Ontario, Canada

**Aims:** The goal of the gender-affirming treatments (GATs) is to align the gender role and its expression with the experienced gender, ultimately resulting in improved gender dysphoria and quality of life (QOL). There is a lack of rigorous, validated, and specific patient-reported outcome (PRO) instrument for assessing outcomes in individuals receiving GATs. The aim of this study was to develop a conceptual framework for a patient-reported outcome (PRO) instrument for adults receiving gender-affirming treatments (GENDER-Q) by developing a comprehensive understanding of issues that individuals consider to be important. **Methods:** An interpretive description approach was used to conduct in-depth interviews with 79 adults who were seeking or receiving GATs between October 2018 and March 2020 across six centers in four countries. The interviews were used to explore the impact of gender-affirming treatments on the individual's QOL, and satisfaction with appearance and process of care. Interviews were transcribed and coded using line-by-line approach. Constant comparison was used to develop and refine the conceptual framework. The interviews and the analyses are ongoing and will conclude when data redundancy is achieved. **Results:** We have completed 79 interviews to-date, with participants aged 18 to 62 years (mean  $34.5 \pm 13$  years). The participants were identified as transmale ( $n = 38, 48\%$ ), transfemale ( $n = 37, 47\%$ ), or non-

binary ( $n = 4$ , 5%), and had hormone replacement therapy ( $n = 68.86\%$ ), top surgery ( $n = 40$ , 51%), bottom surgery ( $n = 44$ , 56%) or other GAT surgeries ( $n = 8$ , 10%). Preliminary analyses suggest that the participants described concepts of interest in three top-level domains, each of which included subdomains: appearance (face, chest, torso, upper and lower extremities, overall), QOL (physical well-being, psychological well-being, social well-being, sexual well-being), and process of care (satisfaction with healthcare team). The participants were able to describe the impact of the GATs on their QOL from pre-treatment to post-treatment in all of these domains. Conclusion: Rich, in-depth concept elicitation interviews are key to ensuring the PRO instrument measures what matters to patients. The conceptual framework developed as a result of this study will form the basis of the scales in the GENDER-Q, which will be refined with the help of further patient and clinician feedback, in the form of interviews and field-test.

#### **(B201.7) Remote monitoring of head and neck cancer patients during radiotherapy with Apple Watch; Design of the OncoWatch feasibility study**

Cecilie Holländer-Mieritz, MD, Rigshospitalet, Department of Oncology, Copenhagen, Denmark; Allan Green, The Capital region of Denmark, Copenhagen, Denmark; Claus Andrup Kristensen, MD PhD, Rigshospitalet, Department of Oncology, Copenhagen, Denmark; Ivan Richter Vogelius, Professor, Rigshospitalet, Department of Oncology, Copenhagen, Denmark; Helle Pappot, Professor, Rigshospitalet, Department of Oncology, Copenhagen, Denmark

**Aims:** Patients with head and neck cancer (HNC) experience severe side effects during radiotherapy (RT). Ongoing technological advances in wearable sensors allow for real-time collection of objective data, e.g., physical activity and heart rate. A smartwatch such as the Apple Watch allows for objective health data monitoring outside hospital with minimal effort for the patient. We here describe the design of the OncoWatch feasibility study. **Methods:** A prospective, single cohort trial will be conducted at Rigshospitalet, Department of Oncology, Denmark. Patients  $\geq 18$  years planned to primary or post-operative curatively intended radiotherapy for HNC are eligible. Consenting patients will be asked to wear an Apple Watch continuously during radiotherapy and until 2 weeks after end of RT. The study will include 20 patients. Demographic data, objective toxicity scores, and hospitalizations will be documented. Enrollment is expected to begin April 2020. **Results:** The primary outcome is to determine if it is feasible for the patients to wear a smartwatch continuously (minimum 16 h/day) for up 12 weeks with description of the data completeness. Secondly, we will explore how the heart rate and physical activity change over the treatment course. **Conclusion:** The study will assess the feasibility of using the Apple Watch for home monitoring of patients with HNC. Remote monitoring with an Apple Watch may lead to early identification of symptoms and secure timely intervention for symptom management.

#### **(B201.8) Body image scale: Evaluation of the psychometric properties in three Indian head and neck cancer populations**

Chindhu Shunmuga Sundaram, The University of Sydney, Sydney, Australia; Haryana Dhillon, PhD, The University of Sydney, Sydney, Australia; Phyllis Butow, PhD, The University of Sydney, Sydney,

Australia; Puma Sundaresan, The University of Sydney, Sydney, Australia; Mahati Chittem, PhD, IIT Hyderabad, Hyderabad, India; Nagaraj Huilgol, Dr. Balabhai Nanavati Hospital, Hyderabad, India; Niveditha Akula, IIT Hyderabad, Hyderabad, India; Surendran Veeraiah, PhD, Cancer Institute (WIA), Chennai, India; Claudia Rutherford, PhD, The University of Sydney, Sydney, Australia

**Aims:** Body image is a subjective concept encompassing individuals' views and emotions about their body. Head and neck cancer (HNC) diagnoses and treatment affect several psychosocial concepts including body image. Large numbers of HNC patients are diagnosed each year in India but no suitable measures in regional languages assessing body image exist. Translating existing valid and reliable measures, rather than developing new ones, reduces effort, saves time and speeds up acquisition of knowledge related to cultural differences. This study assessed the psychometric properties of three translated versions of the Body image scale (BIS), a measure suitable for clinical and research use in HNC populations, in Indian populations speaking Tamil, Telugu, and Hindi. **Methods:** Translated versions of BIS were completed by HNC patients (approximately 200 per language group sufficient to explore treatment differences) recruited from three cancer centers across India. Eligibility criteria included: 18 years of age and above, any gender, any diagnosis from the HNC spectrum except thyroid cancer, and no exclusions based on treatment type or stage. Psychometric analyses were conducted including factor analysis using principal component analysis to explore scaling, internal consistency reliability using Cronbach's alpha, and convergent and discriminant validity. Patients also completed the EORTC QLQ C-30 and EORTC QLQ HN-35 measures to enable exploration of convergent and discriminant validity. **Results:** Participants included 205 Tamil speaking, 216 Telugu speaking and 200 Hindi speaking HNC patients (total  $n = 621$ ). Factor analysis revealed a one-factor solution with cumulative variances of 78.66%, 65.91%, and 71.42% for Tamil, Telugu, and Hindi populations, respectively. Cronbach's alpha coefficients ranged between 0.891 and 0.969 indicating good reliability. Hypothesized correlations between similar and different constructs were as expected, indicating construct validity. **Conclusion:** The reliability and validity of the three translated versions of the BIS were maintained, providing a method for assessing body image of HNC populations speaking Tamil, Telugu and Hindi across the illness trajectory. The translated versions were suitable for routine clinical and research use in Indian HNC populations.

#### **(B201.9) Decisional Regret Five Years After Couples Experiencing Infertility Seek Consultation from Reproductive Endocrinologist: A Mixed Methods Study**

Rachel Cusatis, Medical College of Wisconsin, Milwaukee, Wisconsin, United States; Colin Johnson, MD, Medical College of Wisconsin, Wauwatosa, Wisconsin, United States; Kate Schoyer, MD, Medical College of Wisconsin, Wauwatosa, Wisconsin, United States; Judith Myers, MS, Medical College of Wisconsin, Wauwatosa, Wisconsin, United States; Jay Sandlow, MD, Medical College of Wisconsin, Wauwatosa, Wisconsin, United States; Kathryn E. Flynn, PhD, Medical College of Wisconsin, Wauwatosa, Wisconsin, United States

**Aims:** To understand the extent to which men and women who sought a consultation from a reproductive specialist for infertility express decisional regret about their family-building choices 5 years later. **Methods:** We enrolled couples and individuals ( $n = 156$ ) with an



initial consultation with a reproductive specialist. Participants completed questionnaires prior to their first consultation, at 12 months, and 5 years. On the 5-year assessment, we administered the Decisional Regret Scale, a 5-item scale assessing regret (0–100), referencing “the decision you made to add a child to your family.” A score of 25 indicates moderate to severe regret. We used linear regression to assess the relationship between regret and path to parenthood, i.e., biological child(ren), adoption/other, or no child(ren). We considered a two-tailed alpha of 0.05 significant. We asked an open-ended question: “Please describe how you feel about the decisions you made to try to add a child to your family.” Responses were grouped into common themes. Results: 46 men and 73 women responded to the 5-year survey, more than half of whom did not express any regret (Fig. 1). Regret scores ranged from 0 to 90; the average score was 12.0 (2.9) for men and 10.9 (2.0) for women. Individuals who did not have a child at 5 years expressed significantly higher regret [(29.3 (6.8)] compared to those who were pregnant or parenting a child at 5 years [8.4 (1.3);  $p < .01$ ]. Adjusted for gender, patients who had a biological child or had a child through another path had significantly lower regret, by 20 points or more, compared to those who did not have a child (Table 1). Major qualitative themes include acknowledging physical and mental difficulties of treatment, disappointment/grief in unsuccessful attempts and miscarriages, finances, and, for those who had child(ren), the feeling that ‘it was all worth it.’ Conclusion: These data support the importance of attending to patients’ emotional and psychological health during treatment for infertility. Our data also suggest the need for education around the potential for decisional regret and the multiple paths to parenthood, any of which may prevent regret.

Figure 1. Distribution of Participants Decisional Regret Scores

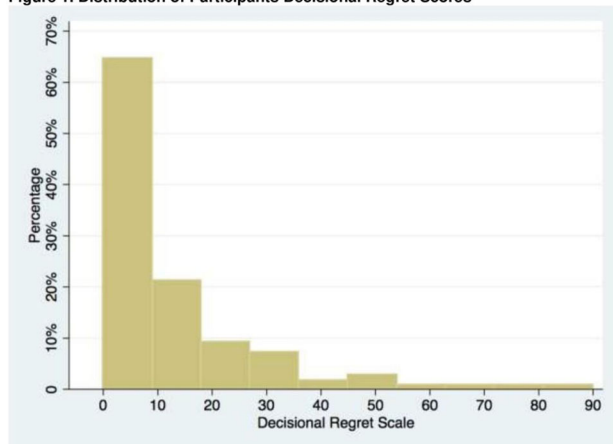


Table 1. Linear Regression of Decisional Regret and Path to Parenthood, adjusting for Gender

Variable	Coefficient	Standard Error	p-value
Path to Parenthood (ref = no child)			
Biological Child	-20.5	6.21	.000
Other Path (adoption, foster, stepchild)	-24.0	4.42	.000
Female (ref = male)	-1.31	3.11	.675

### (B201.10) Video-assisted thoracoscopic surgery (VATS) versus thoracotomy for locally advanced lung cancer: A retrospective longitudinal cohort study based on the Patient-Reported Outcomes

Xing Wei, Department of Thoracic Surgery, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, School of Medicine, University of Electronic Science and Technology of China, Chengdu 610041, China., Chengdu, China; Wei Dai, Department of Thoracic Surgery, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, School of Medicine, University of Electronic Science and Technology of China, Chengdu 610041, China., Chengdu, China; Hongfan Yu, School of public health and management, chongqing medical university, Chongqing, China; Qiang Li, Department of Thoracic Surgery, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, School of Medicine, University of Electronic Science and Technology of China, Chengdu 610041, China., Chengdu, China; Yongtao Han, Department of Thoracic Surgery, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, School of Medicine, University of Electronic Science and Technology of China, Chengdu 610041, China., Chengdu, China; Lin Peng, Department of Thoracic Surgery, Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, School of Medicine, University of Electronic Science and Technology of China, Chengdu 610041, China., Chengdu, China; Qiuling Shi, School of public health and management, chongqing medical university, Chongqing, China

Aims: Video-thoracoscopy as a minimally invasive technique has been widely used in the treatment of locally advanced lung cancers. The comparison of effects between this approach and the traditional thoracotomy approach are more based on traditional indicators in the existing studies. Using patient-reported outcomes, we aimed to define the patients’ experience of recovery after VATS surgery and open thoracotomy. Methods: Patients clinically diagnosed with lung cancer and scheduled for surgery were prospectively enrolled. Patient-Reported Outcomes (PROs) were collected using the MD Anderson Symptom Inventory for lung cancer (MDASI-LC) and single-item Quality of Life Scale (QOL). Longitudinal data about symptoms and functioning of the target patients were collected preoperatively, every post-operative day in hospital and weekly after discharge until 1 month after surgery or the beginning of postoperative adjuvant therapy. Descriptive statistics, Mann–Whitney  $U$  test, Chi Square Tests, Fisher exact probability test, and mixed effect model were used for data analysis. Results: From November 12, 2017 to November 18, 2019, we screened 512 patients, of whom 393 were inclusion. One hundred and seventeen primary lung cancer patients with stage II and III by pathological diagnosed were finally enrolled, with 63 patients in the VATS group and 54 in the thoracotomy group. Mixed effect model demonstrated that Compared with those under open thoracotomy, patients under VATS reported less fatigue (Estimate =  $-0.58$ , [95%CI  $-1.08$  to  $-0.08$ ],  $p = 0.0233$ ), disturbed sleep (Estimate =  $-1.02$ , [95%CI  $-1.08$  to  $-0.08$ ],  $p = 0.0021$ ), and impaired ability of walking (Estimate =  $-0.56$ , [95%CI  $-1.03$  to  $-0.10$ ],  $p = 0.0168$ ), activity (Estimate =  $-0.63$ , [95%CI  $-1.10$  to  $-0.15$ ],  $p = 0.0098$ ), and work (Estimate =  $-1.15$ , [95%CI  $-1.72$  to  $-0.58$ ],  $p = 0.0001$ ). Interactive effects of surgery type and time was found for disturbed sleep (Estimate =  $-1.02$ , [95%CI  $-1.08$  to  $-0.08$ ],  $p = 0.494$ ). Post-operative QOL was better in the VATS group compared to thoracotomy group (Estimate  $+0.83$ , [95%CI  $+0.26$  to  $+1.41$ ],  $p = 0.0049$ ). Conclusion: During the postoperative hospitalization, the VATS group demonstrated less symptom burden of symptoms and better daily functioning and quality of life. This results suggest that VATS could benefit patients with locally advanced lung cancer in terms of their perception of recovery.

**Table 1. Baseline Demographic and Patient Characteristics**

Variable	Thoracotomy (n=43)	VATS (n=83)	Thoracotomy (%)	VATS (%)	p-value
Age (mean, median (SD))	58.88 (58.24-64.55)	58.88 (58.24-64.55)	57.96 (58.49-64.90)	59.25 (58.49-64.90)	0.6275
BMF (kg/m <sup>2</sup> , median (IQR))	22.79 (21.45-24.48)	21.84 (20.26-24.74)	21.23 (20.38-22.53)	22.05 (21.45-24.48)	0.6865
Gender					0.2384
Male	36 (79.1%)	47 (57.4%)	4485 (40.8%)	4485 (40.8%)	
Female	28 (21.9%)	36 (42.6%)	1618 (15.2%)	1618 (15.2%)	
Education level					0.1514
Middle school graduate or below	61 (88.2%)	4266 (47%)	19 (27.3%)	19 (27.3%)	
Middle school graduate above	36 (28.2%)	21 (23.3%)	13 (27.3%)	13 (27.3%)	
Residence area					0.3111
Non-remote	47 (37.4%)	28 (31.7%)	1633 (15.0%)	1633 (15.0%)	
Remote	21 (88.2%)	31 (42.6%)	2408 (22.0%)	2408 (22.0%)	
Former remote	18 (13.7%)	6 (6.9%)	18 (18.7%)	18 (18.7%)	
Drinking history					0.6379
Yes	73 (56.7%)	4863 (55%)	21 (28.9%)	21 (28.9%)	
No	45 (28.2%)	23 (26.3%)	21 (48.7%)	21 (48.7%)	
ASA classification					0.3446
I-II	49 (49.5%)	28 (34.6%)	21 (28.9%)	21 (28.9%)	
III	48 (28.2%)	35 (42.6%)	31 (81.1%)	31 (81.1%)	
CCI score					0.3446
I-II	44 (54.7%)	37 (38.7%)	21 (28.9%)	21 (28.9%)	
III	17 (81.2%)	24 (28.2%)	27 (60.0%)	27 (60.0%)	
Lesion location					0.1135
Right Upper Lobe	25 (21.4%)	16 (22.4%)	6 (16.7%)	6 (16.7%)	
Right Middle Lobe	23 (18.8%)	9 (10.6%)	9 (24.7%)	9 (24.7%)	
Right Lower Lobe	24 (28.2%)	12 (13.9%)	11 (29.7%)	11 (29.7%)	
Right Middle and Lower Lobe	12 (9.8%)	8 (9.2%)	11 (29.7%)	11 (29.7%)	
Left Upper Lobe	14 (11.7%)	11 (12.6%)	11 (29.7%)	11 (29.7%)	
Left Lower Lobe	19 (16.2%)	9 (10.4%)	18 (48.4%)	18 (48.4%)	
Other	4 (3.3%)	2 (2.3%)	2 (5.3%)	2 (5.3%)	
Resection type					0.6865
Subtotal resection	7 (2.4%)	5 (6.0%)	6 (16.7%)	6 (16.7%)	
Lobectomy	38 (88.6%)	36 (43.6%)	36 (97.3%)	36 (97.3%)	
Extended resection	34 (29.7%)	18 (21.9%)	24 (64.9%)	24 (64.9%)	
Lymphadenectomy					0.6665
Systematic lymph node dissection	49 (39.8%)	32 (38.9%)	21 (28.9%)	21 (28.9%)	
Selective lymph node sampling	41 (28.2%)	41 (49.4%)	17 (45.9%)	17 (45.9%)	
Not performed	3 (2.4%)	1 (1.2%)	3 (8.1%)	3 (8.1%)	

**Table2. Symptoms compared by VATs vs Thoracotomy**

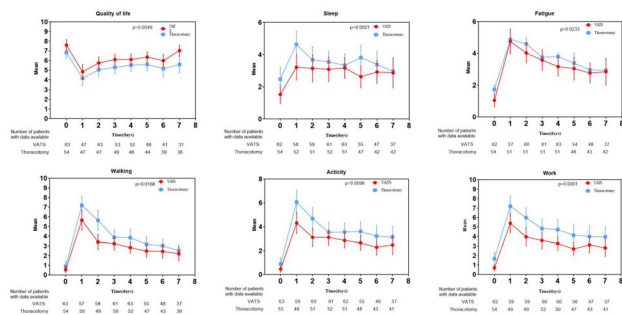
MDASI Symptoms	VATs vs Thoracotomy			Group-p	Time-p	group*time-p
	Estimate	SE	95%CI			
pain	-0.31	0.22	-0.75 to 0.13	0.1630	<0.0001	0.9196
fatigue	-0.58	0.25	-1.08 to -0.08	<b>0.0233</b>	<0.0001	0.4619
cough	-0.52	0.27	-1.05 to 0.02	0.0577	0.0019	0.0723
short breath	-0.11	0.31	-0.73 to 0.51	0.1842	<0.0001	<b>0.0338</b>
sleep	-1.02	0.32	-1.66 to -0.38	<b>0.0021</b>	<0.0001	<b>0.0494</b>
walking	-0.56	0.23	-1.03 to -0.10	<b>0.0168</b>	<0.0001	0.1384
activity	-0.63	0.24	-1.10 to -0.15	<b>0.0098</b>	<0.0001	0.3013
work	-1.15	0.29	-1.72 to -0.58	<b>0.0001</b>	<0.0001	0.7112
constipation	-0.22	0.25	-0.71 to 0.28	0.3904	<0.0001	0.3930
nausea	0.08	0.16	-0.23 to 0.40	0.6055	<0.0001	0.7210
vomiting	-0.06	0.13	-0.32 to 0.20	0.6513	0.0008	0.9202
drowsy	-0.56	0.31	-1.17 to 0.05	0.0735	<0.0001	0.1534
<b>QOL</b>	<b>0.83</b>	<b>0.29</b>	<b>0.26 to 1.41</b>	<b>0.0049</b>	<0.0001	0.5891

Methods: mixed effect model was used. The operation mode and time were fixed effects, and the patients were random effects. The least square method is used to estimate the relative structure as unstructured (compared with unstructured, simple and conformer).

Group \*time refers to the interaction between groups and the total time. If there is no interaction, the interaction is removed and the model is rematched.

In VATs vs Thoracotomy, estimate refer to the estimated mean increase or decrease in VATS compared with Thoracotomy, which is more reliable.

As shown in the above table, the distribution of fatigue, sleep, walking, activity, work and QOL in the whole time (a2-pod7) track was statistically significant.



**B202: Methods**

**(B202.2) The Quality of Life in Bipolar Disorder (QoLBD) Questionnaire a decade on—a systematic review of the measurement of condition-specific aspects of quality of life in bipolar-disorder**

Emma Morton, PhD, Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada, Vancouver, British Columbia,

Canada; Greg Murray, PhD, Faculty of Health, Arts and Design, Swinburne University, Melbourne, Australia, Victoria, Australia; Lakshmi Yatham, MBBS, DPM, FRCPC, MBA (Exec), Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada, Vancouver, British Columbia, Canada; Raymond Lam, MD PhD, Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada, Vancouver, British Columbia, Canada; Erin Michalak, PhD, Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada, Vancouver, British Columbia, Canada

**Aims:** Condition-specific quality of life (QoL) instruments assess domains specifically impacted by an illness, and are more representative of the priorities of people with lived experience. As such, the development of the first and only bipolar disorder (BD) specific measurement instrument, the Quality of Life in Bipolar Disorder (QoLBD) questionnaire, marked an important step forward for the literature. The present systematic review firstly aims to characterize the uptake of the QoLBD in the BD literature, including the geographical location and study designs in which it has been applied. Secondly, we aim to review findings on the psychometric properties of the QoLBD, as well as the impact of mood symptoms, psychosocial variables, and treatments on condition-specific aspects of QoL. **Methods:** Peer-reviewed papers citing the QoLBD were included for analysis if they reported original empirical data using the QoLBD in a BD population. No restrictions were placed on language or study type. **Results:** 110 articles citing the QoLBD were identified, and 35 publications were retained for analysis. Factor analytic methods and evaluation of psychometric properties provided support for five cross-cultural (Turkish, Persian, Spanish, and Chinese) and one alternate form (web-based) adaptations of the original scale. Fourteen clinical trials investigating QoLBD outcomes were identified, the majority of which (78.6%) described psychological interventions. Promising effect sizes were observed for recovery-focused cognitive behavior therapy, cognitive training, and online recovery-focused psychoeducation and mindfulness interventions. Fifteen studies examined clinical, functional, and psychological correlates of QoLBD scores. Depression was found to have a negative impact on condition-specific aspects of QoL, while mixed findings were reported regarding the influence of mania. **Conclusion:** A sizeable, international body of empirical evidence now exists regarding the measurement, correlation, and treatment of condition-specific aspects of QoL in BD. Clinical trials typically had small samples and were under-powered to detect significant treatment effects. The evidence base on potential predictors of QoLBD scores is limited in that reviewed studies predominantly used a cross-sectional design, and as such do not permit inferences about causal relationships. Further large scale clinical and prospective trials are needed to identify effective treatments and variables which predict improvements in QoL outcomes prioritized by patients.

**(B202.3) A PROMISING prospect of measuring pediatric general health: a comparison of the PROMIS® pediatric Global Health scale (PGH-7) and the Pediatric Quality of Life Inventory (PedsQL™)**

Michiel Luijten, MSc, Emma Children’s Hospital, Amsterdam UMC, Amsterdam, Netherlands; Lotte Haverman, PhD, Emma Children’s Hospital, Amsterdam UMC, Amsterdam, Netherlands; Raphaële van Litsenburg, PhD, Princess Maximá Center for Pediatric Oncology, Utrecht, Netherlands; Leo Roorda, PhD, Amsterdam Rehabilitation Research Center | Reade, Amsterdam, Netherlands; Martha Grootenhuys, Prof, Princess Maximá Center for Pediatric Oncology, Utrecht, Netherlands; Caroline Terwee, PhD, Amsterdam UMC, Utrecht, Netherlands

**Aims:** On February 18th 2020 the International Consortium for Health Outcomes Measurement (ICHOM) announced the release of the Standard Set for overall pediatric health. This outcome set contains the Patient-Reported Outcomes Measurement Information System (PROMIS) Pediatric Scale v1.0 Global Health (PGH-7 + 2) for measuring overall physical, mental and social health. Our aim was to assess the psychometric properties of the PGH-7 in the Dutch population and to compare the performance of the PGH-7 with the Pediatric Quality of Life Inventory (PedsQLTM). **Methods:** Children aged 8–18 years ( $n = 2654$ ), representative of the Dutch population on key demographics were asked to complete the PGH-7 (nitems = 7) and the PedsQL (nitems = 23). To assess structural validity of the PGH-7 a graded response model (GRM) was fitted to the data after assessing the following assumptions: Unidimensionality through CFA ( $CFI > .95$ ,  $TLI > .95$ ,  $RMSEA < .10$ ), local independence by residual correlations ( $r < .20$ ) and monotonicity by Mokken analysis ( $H > .50$ ,  $Hi > .30$ ). Item fit of the GRM model was inspected with S-X2, where  $p < .001$  indicates misfit. Additionally, convergent validity of the PGH-7 T-score with the PedsQL total score was assessed. A moderately strong correlation ( $> .50$ ) was expected, as both instruments measure physical, mental and social domains. Percentage of participants reliably measured was assessed using the standard error of measurement (SEM)  $< 0.32$  as a criterion (which equals a reliability of 0.90). Relative efficiency was calculated  $(1 - SEM^2)/nitems$  to compare how well both instruments perform relative to the amount of items administered. **Results:** In total 1082 (response rate = 40.8%) children completed both questionnaires. All GRM assumptions were met. PGH-7 displayed good structural (no misfit) and convergent ( $r = .65$ ) validity. Both questionnaires measured reliably (nPGH-7 = 74.5%, nPedsQL = 76.6%) at the mean and 2SD in clinically relevant direction. The relative efficiency of the PGH-7 was 2.6 in comparison to the PedsQL, indicating that, on average, the items in the PGH-7 are 2.6 times more informative than PedsQL items. **Conclusion:** The PGH-7 displays sufficient reliability and validity in the general Dutch pediatric population. The scale measures more efficiently than the most commonly used legacy instrument (PedsQL).

#### **(B202.4) Two-step screening for depressive symptoms in patients with end-stage kidney disease**

Sumaya Dano, MSc, Multi-Organ Transplant Program, University Health Network, Toronto, Ontario, Canada; Evan Tang, Multi-Organ Transplant Program, University Health Network, Toronto, Ontario, Canada; Faisal Jamil, Multi-Organ Transplant Program, University Health Network, Toronto, Ontario, Canada; Dean Christidis, Multi-Organ Transplant Program, University Health Network, Toronto, Ontario, Canada; Shahd Fulath Khan, Multi-Organ Transplant Program, University Health Network, Toronto, Ontario, Canada; Susan Barlett, PhD, Dept. of Clinical Epidemiology, Research Institute of the McGill University Health Centre, Toronto, Quebec, Canada; Doris Howell, RN, PhD, Princess Margaret Cancer Centre, Toronto, Ontario, Canada; Madeline Li, MD, PhD, Princess Margaret Cancer Centre, Toronto, Ontario, Canada; Marta Novak, MD, PhD, Centre for Mental Health, University Health Network, Toronto, Ontario, Canada; Istvan Mucsi, MD, PhD, Multi-Organ Transplant Program, University Health Network, Toronto, Ontario, Canada; MD, PhD, Multi-Organ Transplant Program, University Health Network, Toronto, Ontario, Canada

**Aims:** Routine screening for depressive symptoms can be resource intensive and burdensome for patients. Ultra-brief screening tools such as the single depression item from the Edmonton Symptom Assessment Survey-revised (ESASr-D) or the 2-item Patient Health Questionnaire (PHQ2) could quickly rule out depression. However, to

identify false positive cases from the first screening step, a more precise tool such as the PROMIS Depression Computer Adaptive Test (PROMIS Depression CAT) may be needed. We explore this approach to screen for depressive symptoms in patients with end-stage kidney disease (ESKD). **Methods:** A cross-sectional, convenience sample of adult kidney transplant recipients and patients on maintenance dialysis completed the ESASr, PROMIS Depression CAT and PHQ9. A PHQ9 score  $\geq 10$  identified moderate/severe depressive symptoms. ESASr-D and PHQ2 scores of  $\geq 1$  and  $\geq 2$  were evaluated for the first screening step. In the second step, a PROMIS Depression CAT T-score  $\geq 55$  was used to identify patients with potentially significant depressive symptoms. The number of true positive (TP), true negative (TN), false positive (FP) and false negatives (FN) cases, sensitivity, specificity, positive and negative predictive value (PPV and NPV) were calculated for the different scenarios. **Results:** Mean(SD) age of the 164 participants was 52(17), 68% were male, 62% Caucasian. Based on PHQ9, 16% ( $n=26$ ) had depression. ESASr-D  $\geq 1$  screened out 104 patients (Sensitivity:73%, Specificity:70%, PPV:0.32, NPV:0.93, FP: 41, FN:7). ESASr-D  $\geq 2$  screened out 123 patients (Sensitivity:69%, Specificity:83%, PPV:0.44, NPV:0.93, FP:23, FN:8). PHQ2  $\geq 1$  screened out 111 patients (Sensitivity:92%, Specificity:79%, PPV:0.45, NPV:0.98, FP:29, FN:2). PHQ2  $\geq 2$  screened out 121 patients (Sensitivity:92%, Specificity:86%, PPV:0.56, NPV:0.98, FP:19, FN:2). PROMIS Depression CAT  $\geq 55$  screened out 134 patients (Sensitivity:69%, Specificity:91%, PPV:0.60, NPV:0.94, FP:12, FN:8). PHQ2  $\geq 2$  combined with PROMIS Depression CAT provided the best two-step results (Sensitivity:65%, Specificity:94%, PPV:0.65, NPV:0.93; FP:9, FN:9). **Conclusion:** All the one-step screening options had high FP rates that may overburden the clinical system and may generate undue stress for patients falsely identified as having depression. The two-step screening had modest sensitivity but less FN. However, higher sensitivity may be preferable considering the negative implications of depression. Future studies should confirm the optimal screening combination using clinical assessment.

#### **(B202.5) Social media listening in PRO development: Can online conversations contribute to our conceptualization of a disease area?**

Sarah Knight, MSc, DRG (part of Clarivate), Bicester, United Kingdom; Chloe Johnson, DRG (part of Clarivate), Manchester, United Kingdom; Jaya Dulani, DRG, Guragon, India; Chris Marshall, DRG (part of Clarivate), Manchester, United Kingdom; Natalie Aldhouse, DRG (part of Clarivate), Manchester, United Kingdom; Jake Macey, DRG (part of Clarivate), Manchester, United Kingdom; Hannah Pegram, DRG (part of Clarivate), Manchester, United Kingdom; Maile Hunter, DRG (part of Clarivate), Nashville, Tennessee, United States; Tamara Al-Zubeidi, DRG (part of Clarivate), Bicester, United Kingdom; Stephanie McKee, DRG (part of Clarivate), Manchester, United Kingdom; Helen Kitchen, DRG (part of Clarivate), Manchester, United Kingdom; Helen Kitchen, DRG (part of Clarivate), Manchester, United Kingdom

**Aims:** To compare data obtained via social media listening (SML) and a traditional qualitative literature review to develop conceptual models in patient-reported outcome (PRO) development. **Methods:** Using a case study in mild Alzheimer's Disease (AD), the concepts identified from a qualitative literature review searching Ovid platforms (run by experts in PRO research) were compared to patient-reported data obtained via SML methods. Brandwatch was used to explore publicly available data from Twitter, Facebook, Instagram (hashtags), blogs/forums and YouTube (text descriptions). A comparison of results was conducted retrospectively. **Results:** Concepts from the qualitative literature review were categorized into symptoms

(memory loss, reduced cognition, apathy, reduced concentration and confusion) and impacts (on thought processing, daily activities, social life, emotional impact, and communication). The SML data were harder to categorize as the discussions were unconstrained and not guided by a research question. Data from patients and caregivers were available. Symptoms included memory loss, hallucinations, confusion, behavioral changes, sleep disorders, depression, delusion, brain fog, cognitive impairment, lack of empathy, agitation, and anosognosia. Alignment of patient-reported symptoms was identified between methods. Fewer and different impacts were described in the SML compared to the literature review; posts described limited information about the disease, fear of uncontrollable symptoms in the future, concern about side effects of experimental drugs, apprehension about future changes, patient and caregiver emotional burden (specifically worry), the daily struggle with disease, and future planning. Conclusion: SML can provide additional supplementary data in the preliminary stages of PRO development. SML can recover a vast quantity of data from different sources, and it can be a useful research method to implement when there is limited, published qualitative literature, for example, in a rare disease. However, SML does have limitations. The identity of the ‘poster’ cannot be confirmed; there is no evidence of clinician-confirmed/verified diagnosis, clinical or demographic information. Researchers are unable to query ambiguous or unclear concepts. A vast amount of data is also usually recovered, making analysis time intensive. For these reasons, it is recommended that SML should not be used as an alternative to traditional PRO development activities (e.g., qualitative patient interviews).

#### **(B202.6) Assessing HRQoL in chronic wounds across countries: the cross-cultural validity of the revised Wound-QoL questionnaire**

Catharina von Stülpnagel, German Center for Health Services Research in Dermatology (CVderm), Institute for Health Services Research in Dermatology and Nursing (IVDP), University Medical Center Hamburg-Eppendorf (UKE), Germany, Hamburg, Germany; Neuza da Silva, Dr., German Center for Health Services Research in Dermatology (CVderm), Institute for Health Services Research in Dermatology and Nursing (IVDP), University Medical Center Hamburg-Eppendorf (UKE), Germany, Hamburg, Germany; Matthias Augustin, Prof. Dr., German Center for Health Services Research in Dermatology (CVderm), Institute for Health Services Research in Dermatology and Nursing (IVDP), University Medical Center Hamburg-Eppendorf (UKE), Germany, Hamburg, Germany; Catherine van Montfrans, Dr., Department of Dermatology, Erasmus medical center, Rotterdam, The Netherlands, Rotterdam, Netherlands; Caroline Fife, MD, Baylor College of Medicine, Houston, Houston, Texas, United States; Ann-Mari Fagerdahl, PhD, Wound Centre, Södersjukhuset, Karolinska Institutet, Department of Clinical Science and Education, Stockholm, Sweden, Houston, Sweden; Alexander Gamus, PhD, Maccabi Health Services, Sackler Faculty of Medicine, Tel-Aviv University, Tel-Aviv, Israel, Tel-Aviv, Israel; Elena Conde Montero, PhD, Department of Dermatology, Hospital Universitario Infanta Leonor, Madrid, Spain, Madrid, Spain; Toni M. Klein, German Center for Health Services Research in Dermatology (CVderm), Institute for Health Services Research in Dermatology and Nursing (IVDP), University Medical Center Hamburg-Eppendorf (UKE), Germany, Hamburg, Germany; Christine Blome, PD Dr., German Center for Health Services Research in Dermatology (CVderm), Institute for Health Services Research in Dermatology and Nursing (IVDP), University Medical Center Hamburg-Eppendorf (UKE), Germany, Hamburg, Germany; Rachel Sommer, PD Dr., German Center for Health Services Research in Dermatology

(CVderm), Institute for Health Services Research in Dermatology and Nursing (IVDP), University Medical Center Hamburg-Eppendorf (UKE), Germany, Hamburg, Germany; Rachel Sommer, Dr., German Center for Health Services Research in Dermatology (CVderm), Institute for Health Services Research in Dermatology and Nursing (IVDP), University Medical Center Hamburg-Eppendorf (UKE), Germany, Hamburg, Germany

Aims: Chronic wounds often impair the life and wellbeing of people affected due to severe restrictions in all domains of health-related quality of life (HRQoL). The prevalence of chronic wounds is reported to be 1.04% of the German population and affects around 15% of patients over the age of 65 in the United States. The Wound-QoL is translated and validated for different countries. The aim of this study was to (1) test psychometric properties of the Wound-QoL across countries with the combination of classical test theory (CTT) and item response theory (IRT) methods and (2) revise the questionnaire accordingly. Methods: Cross-sectional datasets from six countries (US, Germany, the Netherlands, Sweden, Spain, and Israel) were combined, resulting in a total sample size of 1,185 patients. All patients were 18 years or older and diagnosed with a chronic wound. Sociodemographic and clinical variables were matched for age, sex, country, wound type and continent of origin. Results: Almost half of the patients were female (48.4%) and around 42% were diagnosed with leg ulcers. Metric invariance across countries was established for the original 17-items Wound-QoL ( $\Delta CFI = 0.012$ ,  $\Delta RMSEA = 0.001$ ). Nevertheless, IRT indicated several items with low item information and expert meetings discussed about content related issues. Problematic items were excluded of the Wound-QoL. The revised version consists of 14 items clustered in three dimensions with a good internal consistency in terms of the total Wound-QoL score ( $\alpha = 0.913$ ) and the “everyday life” ( $\alpha = 0.907$ ) dimension and acceptable for the dimensions “body” ( $\alpha = 0.709$ ) and “psyche” ( $\alpha = 0.877$ ). Furthermore, cross-cultural metric invariance was proved ( $\Delta CFI = 0.008$  and  $\Delta RMSEA = 0.001$ ), as well as strict invariance for other clinical and sociodemographic variables (e.g., age, sex, and wound type). Conclusion: The revised Wound-QoL is a reliable and valid instrument to measure the HRQoL of patients with chronic wounds across countries. This version may improve the health care of patients affected and is valid to assess the HRQoL in patients with chronic wounds in clinical practice and research. In future research the revised Wound-QoL should be analyzed for convergent validity with generic HRQoL questionnaires, as well as for sensitivity to clinical changes.

#### **(B202.7) Use of the Common Terminology Criteria for Adverse Events (CTCAE) framework to summarize symptomatic toxicities in the European Organisation for Research and Treatment of Cancer (EORTC) Item Library**

Claire Piccinin, MSc, EORTC, Brussels, Belgium; Andrew Bottomley, PhD, EORTC, Brussels, Belgium; Mogens Groenvold, MD, PhD, DSci, University of Copenhagen, Copenhagen, Denmark; Dagmara Kuliś, MA, EORTC, Brussels, Belgium; Galina Velikova, MD, PhD, University of Leeds, Leeds, United Kingdom; Alexandra Gilbert, MD, PhD, University of Leeds, Leeds, United Kingdom

Aims: The EORTC Item Library is an interactive platform comprised of 952 unique items from 67 different EORTC patient-reported outcome (PRO) questionnaires, covering a variety of symptomatic toxicities and types of functioning relevant to cancer patients. These PROs complement clinician-reporting using classifications like the CTCAE, the gold standard for adverse event (AE) reporting in oncology. In order to facilitate the identification of items and provide a common clinical framework, a mapping study was conducted,

linking all items from the Item Library to corresponding CTCAE symptomatic AEs, where relevant. **Methods:** Following a deductive coding methodology, items were searched for within the CTCAE version 5. Items were coded as linked if they were described within the title, description, or grading of an AE. Items not eligible for CTCAE coding were assigned a descriptive classification, using an inductive approach. Symptoms captured in EORTC items but not located in the CTCAE were also recorded. Two raters coded 249 items and agreement was calculated. The remaining 703 items were coded by one rater and verified by the second, with any discrepancies discussed between both until a consensus was reached. **Results:** Agreement for raters was 77.9% for at least one AE per item. Overall, 603 (63.3%) items were linked to 209 different AEs. The majority of linked items were associated with one (62.9%) or two (24.5%) AEs, with a smaller proportion associated with three or more (12.6%). Multiple linkage resulted from either multiple symptoms relating to the same diagnosis or one symptom relating to multiple diagnoses. Four symptoms covered by EORTC items but not found in the CTCAE were identified: bowel urgency, tenesmus, hair color change, and symptomatic skin fibrosis. Of the items not eligible for CTCAE linking, eight descriptive classifications emerged, with the majority of these items covering the emotional impact of cancer diagnosis/treatment (38.5%) and information provision/satisfaction with care (34.0%). **Conclusion:** Mapping symptomatic PRO items to the CTCAE clinical framework may facilitate PRO use in clinical trials and routine care as a systematic method of recording toxicity. In addition to symptomatic toxicities, important issues for cancer patients, including emotional concerns and satisfaction with care, are also represented.

**(B202.8) Developing and pre-testing a measure of quality of life (QoL) for people with primary sclerosing cholangitis (PSC) in the UK**

Elena Marcus, Marie Curie Palliative Care Research Department, Division of Psychiatry, UCL, London, United Kingdom; Douglas Thorburn, MB ChB, UCL Institute for Liver and Digestive Health, Division of Medicine, UCL & Royal Free London, NHS Foundation Trust, London, United Kingdom; Patrick Stone, MB ChB, Marie Curie Palliative Care Research Department, Division of Psychiatry, UCL, London, United Kingdom; Bella Vivat, PhD, Marie Curie Palliative Care Research Department, Division of Psychiatry, UCL, London, United Kingdom

**Aims:** Primary sclerosing cholangitis (PSC) is a rare incurable bile duct and liver disease which can considerably impact quality of life (QoL). We present the development, up to pre-testing, of a UK-developed measure of QoL for people with PSC (PwPSC): the UK-PSC-QoL. **Methods:** The study followed a two-stage mixed-methods design. Stage 1 extracted QoL issues from an earlier survey and a literature review of relevant QoL tools. Consensus on how to reliably stage PSC is lacking, so we initially hypothesized six categories of disease severity/comorbidities: PSC only, PSC with inflammatory bowel disease, awaiting liver transplant, post-transplant, recurrent PSC, PSC-related cancers. Issue relevance, importance, and phrasing were explored in individual and group discussions with UK PwPSC and clinicians. Decision rules, grounded in theoretical principles for tool development, guided issue selection. Retained issues were constructed as items for the provisional UK-PSC-QoL. Stage 2 pre-testing involved exploring item comprehension, acceptability, relevance, redundancy, and response distributions with UK PwPSC grouped in eight categories; subdividing each of the two least severe

of the previous six categories into mild or moderate-severe. **Results:** Stage 1 identified 396 QoL issues, explored with 28 PwPSC and 11 clinicians. Following issue reduction, 83 items were constructed in six domains: Overall QoL (one item), General Health Perceptions (one item), Symptoms (17 items), Functioning (38 items), Self-Management (19 items), and Experience of Care (EoC) (seven items), plus a separate six-item module for PwPSC with a stoma. Stage 2 pre-testing was conducted with 60 PwPSC: 35 completed the measure online/by post, 25 during an interview. The rarity of PSC presented recruitment challenges, and we under-recruited in the four more severe categories (e.g., awaiting liver transplant). Some findings were therefore inconclusive. Analysis, where data permitted, resulted in modifying 24 items to address problematic phrasing, deletion of five items, and addition of one new item. Due to problems with the timeframe and item ambiguity, the EoC domain was constructed as a distinct module. **Conclusion:** The revised UK-PSC-QoL comprises 67 items in five domains, plus separate six-item stoma module and eight-item EoC module. The measure remains provisional, requiring further testing with larger groups of PwPSC within and beyond the UK.

**(B202.9) A suggestion for best practice for missing data in diary collection: exploring the missingness first**

Konstantina Skaltsa, PhD, IQVIA, Barcelona, Spain; Pavol Kral, IQVIA, Bratislava, Slovakia; Matthew Reaney, IQVIA, Reading, United Kingdom; Michael O'Kelly, IQVIA, Dublin, Ireland

**Aims:** Patient-reported outcome (PRO) data are often collected daily, with weekly scores calculated as 7-day average daily scores. Missing data are common; both intermittent, when patients miss a day (or more) but contribute data later, and monotone, when patients stop contributing daily data before the end of the week or discontinue early. Both may prevent calculation of weekly scores but may be telling different stories. A critical step, often overlooked or rushed, is data exploration. We provide empirically informed recommendations that could help analysts justify the assumptions of their analysis. **Methods:** The following two steps could help understand and account for missingness: Explore patterns of missingness Count the number of patients with non-missing days Explore evidence for whether missingness is intermittent and seemingly random (e.g., Monday of week 1; Thursday of week 2), intermittent and seemingly non-random (e.g., Saturday of week 1 and week 2), or monotone (e.g., days 5–7 consistently missing). If treatment administration, or other recorded events, is expected to impact missingness, explore these days. Understand the potential impact of missingness Create spaghetti plots showing patients' trajectories over time Consider whether to provide supportive analyses excluding selected patients where evidence suggesting informative missingness exists, e.g., a sleep diary that is not completed each Saturday may be because the patient is sleeping through the completion period and ignoring this may bias the estimate. Similarly, if a patient completes a pain diary for 4 days and then stops, this may be due to significant worsening. Tabulate reasons and timing of missingness Plot the mean change from baseline by timepoint for different cohorts, e.g., drop-outs vs completers, cohorts defined by reason/time of discontinuation A Kaplan–Meier plot of time to discontinuation could highlight the difference in the proportion of discontinuations between arms **Results:** A thorough investigation of missing data at the day and week level provides a basis for making decisions on the appropriate analysis methods and the underlying assumptions. **Conclusion:** Exploration of missing PRO diary data should be a mandatory step in any analysis in which change in that data is the outcome of interest.

**B203: Clinical Care Applications****(B203.1) Patient-reported Symptoms Burden after CAR T-cell Therapy for Relapsed/Refractory Lymphoma**

Xin Shelley Wang, MD, MPH, MD Anderson Cancer Center, Houston, Texas, United States; Samer Srour, MD, MD Anderson Cancer Center, Houston, Texas, United States; Ishwaria Subbiah, MD, MD Anderson Cancer Center, Houston, Texas, United States; Meagan Whisenant, PhD, Cizik School of Nursing, Houston, Texas, United States; Loretta Williams, PhD, MD Anderson Cancer Center, Houston, Texas, United States; Tsun Hsuan Chen, PhD, MD Anderson Cancer Center, Houston, Texas, United States; Darcy Ponce, BS, MD Anderson Cancer Center, Houston, Texas, United States; Araceli Gonzalez, MD, PhD, MD Anderson Cancer Center, Houston, Texas, United States; Mona Kamal, MD, PhD, MD Anderson Cancer Center, Houston, Texas, United States; Sairah Ahmed, MD, MD Anderson Cancer Center, Houston, Texas, United States; Partow Kebriaei, MD, MD Anderson Cancer Center, Houston, Texas, United States; Partow Kebriaei, MD, MD Anderson Cancer Center, Houston, Texas, United States; Sattva Neelapu, MD, MD Anderson Cancer Center, Houston, Texas, United States; Elizabeth Shpall, MD, MD Anderson Cancer Center, Houston, Texas, United States

**Aims:** Although the benefits of cellular therapies in oncology are increasingly recognized, many of them, including chimeric antigen receptor (CAR)- T-cell therapy, may have both acute and late-onset severe symptomatic toxicities that require initial hospitalization followed by careful monitoring. This study examined patient-reported symptoms and functioning during the first year of CAR T-cell therapy to provide a better understanding of significant symptoms that may need to be monitored and triaged in real-world care. **Methods:** This cross-sectional pilot study surveyed relapsed/refractory lymphoma patients who were anytime within 12 months received standard of care CAR T-cell therapy at MD Anderson Cancer Center in 2019. PROs were assessed using The MD Anderson Symptom Inventory (MDASI) and the PROs Measurement Information System® 29 (PROMIS). Twenty-two relevant symptom items related to CAR T-cell therapy, identified by hematologists and symptom researchers based on clinical experience and currently available literature, were added to the original 13 MDASI core symptom items. All MDASI items are rated on a 0–10 scale to describe symptom severity during the past 24 h. **Results:** A total of 57 patients were included; 49 (86%) received Axicabtaneciclocleucel (Yescarta); 11 (19%) developed grade 3–4 neurotoxicity. Using MDASI, the severity of symptoms were significantly more severe in the first 30 days ( $n = 28$  patients) than during 30–180 days after therapy ( $n = 17$ ) and those > 180 days after therapy ( $n = 12$ ); these included fatigue, poor appetite, inability to eat and interference with general activity (all  $p < .05$ ). Within 30 days of CAR T-cell therapy, over 5 ~ 22% of patients rated the 20 symptoms on MDASI plus CAR-T items as severe (7–10/10), suggesting that intervention was needed. Physical function was the only domain on PROMIS 29 that showed significantly different scores at 3 time periods ( $p = .046$ ). The PRO completion rate was 95%. **Conclusion:** Patients with refractory/relapsed lymphoma who receive CAR T-cell therapy self-reported unique profiles of physical, psychological, and cognitive symptom burden during the first year. Additional qualitative interviewing and cognitive debriefing of patients to identify symptoms of acute and follow-up phases of CAR T-cell therapy is needed to optimize and implement routine symptom monitoring in CAR T-cell patient care.

**(B203.2) Health-related quality of life assessment for patients with advanced or metastatic renal cell carcinoma (mRCC) treated with a tyrosine kinase inhibitor (TKI) using electronic patient-reported outcomes in daily clinical practice: QUANARIE trial**

Guillaume Mouillet, MD, PhD, Department of Medical Oncology, University Hospital of Besançon, Besançon, France; Antoine Falcoz, Methodological and Quality of Life Unity, University Hospital Besançon, Besançon, France; Philippe Barthelemy, Department of Medical Oncology, Institut de Cancérologie Strasbourg Europe, Strasbourg, France; Lionnel GEOFFROIS, Department of Medical Oncology, Institut de Cancérologie de Lorraine, Vandoeuvre-lès-Nancy, France, France; Jean-Christophe Eymard, Department of Medical Oncology, Institut Jean Godinot, Reims, France; Sylvain LADDOIRE, Department of Medical Oncology, Centre Georges-François Leclerc, Reims, France; Jerome Plaza, Department of Medical Oncology, Hôpitaux Privés de Metz, Metz, France; Erika Viel, Department of Medical Oncology, Hopital Privé Sainte Marie, Chalon-sur-Saône, France; Tristan Maurina, Department of Medical Oncology, University Hospital Besançon, Besançon, France; Fabien Calcagno, Department of Medical Oncology, University Hospital Besançon, Besançon, France; Joelle Fritzsich, Department of Medical Oncology, University Hospital Besançon, Besançon, France; Joelle Fritzsich, Methodological and Quality of Life Unity, University Hospital Besançon, Besançon, France; Oumelkheir Djoumakh, Methodological and Quality of Life Unity, University Hospital Besançon, Besançon, France; Dewi Vernerey, Methodological and Quality of Life Unity, University Hospital Besançon, Besançon, France; Antoine Thiery-Vuillemin, Department of Medical Oncology, University Hospital Besançon, Besançon, France; Sophie Paget-Bailly, Department of Medical Oncology, University Hospital Besançon, Besançon, France

**Aims:** Routine Electronic Monitoring of Health-Related Quality of Life (HRQoL) (REMOQOL) in daily clinical care with real-time feedback to physicians could help to enhance patient-centered care. We evaluated the feasibility of REMOQOL for patients with metastatic renal cell carcinoma (mRCC) treated with TKI at multi-center scale in the French context. **Methods:** QUANARIE study (NCT03062410) is an interventional, prospective, multicenter trial involving 9 French oncological centers. Patients diagnosed with mRCC initiating TKI anti-VEGF treatment (Sunitinib or Pazopanib) were invited to complete the European Organisation for Research and Treatment of Cancer (EORTC) QLQ-C30 cancer-specific questionnaire before each visit with the physician on tablets and/or computers in the hospital or at home. During the visit, physicians had real-time access to visual summaries of HRQoL scores. The primary objective was to assess the proportion of patients having good compliance with REMOQOL during the first 12 months, which is defined as at least 66% of patients with filled out questionnaires during follow-up. We hypothesized that 80% of patients having good compliance with REMOQOL would be meaningful, and as defined by Fleming's one-stage design and defined parameters, we used data of the first 45 evaluable patients. **Results:** Between April 2017 and September 2018, 56 patients were included. Among them, 50 patients were evaluable: 25 treated with Sunitinib and 25 with Pazopanib. The 6 non-evaluable patients died ( $n = 4$ ), had disease progression ( $n = 1$ ) in the first 3 months or did not fill HRQoL questionnaire after the start of treatment ( $n = 1$ ). Mean age of evaluable patients was 65.5 years, and 72% ( $n = 36$ ) were male. In the first 12 months, the median number of HRQoL assessments was 11 (IQR 7–14), 60% were completed at home. At baseline, relatively high mean symptoms scores were observed for fatigue, sleep disturbance and pain. The proportion of patients with an adequate compliance rate was 97.8% ( $n = 44/45$ ).

Thirty patients (66.7%) had a compliance rate of 100%. Conclusion: The results of QUANARIE study suggest that REMOQOL is feasible in terms of patient's compliance in the French context at a multicenter scale.

### (B203.3) Patient and caregiver experiences informing an early palliative care pathway: a qualitative study

Sadia Ahmed, University of Calgary, Calgary, Alberta, Canada; Sveda Farwa Naqvi, University of Calgary, Calgary, Alberta, Canada; Aynharan Sinnarajah, Alberta Health Services, Calgary, Alberta, Canada; Gwen McGhan, University of Calgary, Calgary, Alberta, Canada; Jessica Simon, Alberta Health Services, Calgary, Alberta, Canada; Maria Jose Santana, University of Calgary, Calgary, Alberta, Canada

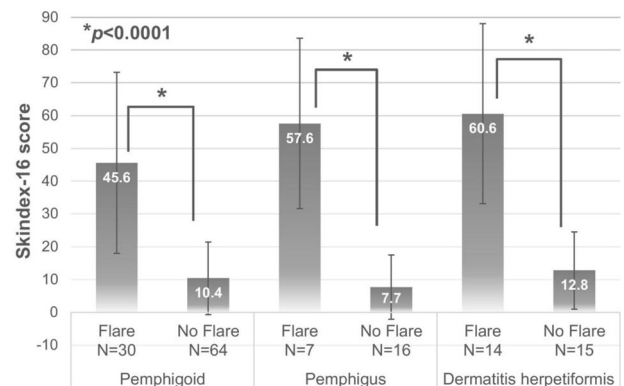
**Aims:** The World Health Organization defines palliative care as an approach that improves the quality of life of patients as they face life-threatening conditions by attending physical, psychosocial, and spiritual needs. The Palliative Care Early and Systematic (PaCES) program in Alberta implemented an early palliative care pathway for advanced colorectal cancer patients in January 2019, defined as a consultative visit from a specialist palliative care provider, palliative homecare service or hospice admission greater than or equal to 3 months before death. This study aims to understand the experience of patients and family caregivers receiving early palliative care supports, and compare those experiences with participants experiencing standard oncology care. **Methods:** This is a qualitative and patient-oriented study. Patient partners supported the development of the interview guide, along with healthcare providers on the team. Participants in Calgary were recruited with the support of a specialist palliative care nurse over the phone, and followed up by a researcher after consent to contact was given. Semi-structured telephone interviews with patients living with advanced colorectal cancer and family caregivers were conducted to explore their experiences with an early palliative approach to care. Interviews were audio-recorded. Interviews were transcribed, and the data thematically analyzed supported by the qualitative analysis software, NVivo. **Results:** A total of 12 participants (7 patients, 5 family caregivers) were interviewed over the phone after implementation of the care pathway. Participants expressed that visits from their early palliative care nurse was helpful, improved their understanding of palliative care, and improved their care. Three main themes shaped their experience of early palliative care: care coordination, coping with advanced cancer, and patient and family engagement. Main differences before and after implementation of the care pathway was in care coordination and communication with and among healthcare providers, understanding of palliative care, involvement of the family physician, and advance care planning discussions. **Conclusion:** Early palliative care delivered by a specialist nurse can improve advanced cancer care, including an improved understanding and acceptance of early palliative care. The PaCES program is currently underway in Calgary for advanced colorectal cancer however it can be expanded to other cancer conditions.

### (B203.4) Patient-reported outcomes can identify skin flaring in autoimmune blistering diseases

Aaron Secrest, MD, PhD, University of Utah Health, Salt Lake City, Utah, United States; Vanina L Taliercio, MD, MS, University of Utah Department of Dermatology, Salt Lake City, Utah, United States; Zachary H Hopkins, MD, Broward Health Department of Dermatology, Fort Lauderdale, Florida, United States

**Aims:** Patient-reported outcomes (PROs) can help to quantify the patient's voice during clinic visits to help shift the focus to patient-centered care. Despite the resources required for implementation, capturing real-time PRO data for clinical use is essential in understanding the true burden and impact of disease. Using Skindex-16, we explored the impact of autoimmune blistering diseases on patients' QOL, both during flares and in remission, in an unselected population seen at our institution. **Methods:** Since September 2016, the University of Utah Department of Dermatology has collected PROs electronically at clinic visits. The department has a weekly Autoimmune Skin Disease Clinic, staffed by five specialists. Patients either completed PROs online via a secure web link or on a tablet in the waiting room after check-in with PRO scores available during the clinic visit. Skindex-16 is a dermatology-specific QOL measure that assesses skin disease impact on symptoms, emotions, and functioning (range 0-100, higher scores = greater QOL impact). Demographic and clinical data were retrieved via manual chart review and linked to PRO scores for each clinic visit. Categorical variables were compared with Chi square and continuous variables with Student t-tests. SPSS v26 was used for analysis. **Results:** Between September 2016 and July 2019, 164 patients (89 female/75 male) with an autoimmune blistering disease completed Skindex-16 assessments at 249 visits: 160 visits for pemphigoid, 43 for pemphigus, and 46 for dermatitis herpetiformis. Patients reported their blistering disease was flaring at 31% of visits. Flaring was more common in female and older patients and was associated with higher PROMIS-Depression and lower PROMIS-Physical Function scores and poorer self-reported general health (Table 1). Overall Skindex-16 scores were four-to-seven-fold higher for autoimmune blistering patients when flaring than when in remission (Fig. 1). **Conclusion:** Autoimmune blistering diseases are severe skin conditions with high morbidity and mortality. Treatments can cause these blistering rashes to go into remission, but flares can happen and are associated with significant negative impacts on QOL. These results show the clinical potential for PROs to monitor patients with chronic relapsing and remitting skin conditions like autoimmune blistering diseases.

**Figure 1.** Mean ( $\pm$ SD) Overall Skindex-16 by autoimmune blistering disease and flaring status.



**Table 1.** Demographic and clinical characteristics and Skindex-16 scores of autoimmune blistering patients by flaring status (mean  $\pm$  SD,  $n=249$ ).

	Range	Flaring	Not Flaring	p-value
N (%)		77 (30.9%)	172 (69.1%)	---
Female (N (%))		51 (66.2%)	88 (51.2%)	<b>0.03</b>
Age		68.0 ( $\pm$ 14.2)	62.7 ( $\pm$ 14.4)	<b>0.007</b>
Age at diagnosis (mean $\pm$ SD)		56.1 ( $\pm$ 16.1)	59.8 ( $\pm$ 16.8)	0.15
Years since diagnosis (mean $\pm$ SD)		6.5 ( $\pm$ 9.0)	8.2 ( $\pm$ 8.0)	0.11
Disease Severity (self-reported)	1–10	6.1 ( $\pm$ 2.3)	1.9 ( $\pm$ 1.3)	<b>&lt;0.001</b>
General Health (self-reported)				<b>0.006</b>
Poor		2 (2.7%)	4 (2.3%)	---
Fair		24 (32.4%)	24 (14.0%)	---
Good		26 (35.1%)	55 (32.2%)	---
Very Good		17 (23.0%)	68 (39.8%)	---
Excellent		5 (6.8%)	20 (11.7%)	---
PROMIS-Depression	0–100	52.9 ( $\pm$ 9.2)	47.0 ( $\pm$ 8.1)	<b>&lt;0.001</b>
PROMIS-Physical Function	0–100	44.3 ( $\pm$ 9.8)	47.3 ( $\pm$ 9.1)	<b>0.02</b>
Skindex-16	0–100	51.4 ( $\pm$ 27.7)	10.3 ( $\pm$ 11.0)	<b>&lt;0.001</b>
S1 – Itch	0–6	3.13 ( $\pm$ 2.14)	1.30 ( $\pm$ 1.47)	<b>&lt;0.001</b>
S2 – Burn	0–6	2.65 ( $\pm$ 1.98)	0.62 ( $\pm$ 1.23)	<b>&lt;0.001</b>
S3 – Hurt	0–6	2.80 ( $\pm$ 2.13)	0.53 ( $\pm$ 1.08)	<b>&lt;0.001</b>
S4 – Irritated	0–6	4.02 ( $\pm$ 1.72)	1.13 ( $\pm$ 1.55)	<b>&lt;0.001</b>
Skindex-16-Symptoms	0–100	53.4 ( $\pm$ 28.2)	15.0 ( $\pm$ 16.7)	<b>&lt;0.001</b>
S5 – Recurrence Fear	0–6	4.15 ( $\pm$ 1.76)	1.18 ( $\pm$ 1.62)	<b>&lt;0.001</b>
S6 – Worry	0–6	3.68 ( $\pm$ 2.36)	1.02 ( $\pm$ 1.57)	<b>&lt;0.001</b>
S7 – Appearance	0–6	3.17 ( $\pm$ 2.45)	0.55 ( $\pm$ 1.13)	<b>&lt;0.001</b>
S8 – Frustrated	0–6	4.00 ( $\pm$ 2.16)	0.90 ( $\pm$ 1.38)	<b>&lt;0.001</b>
S9 – Embarrassed	0–6	2.64 ( $\pm$ 2.42)	0.29 ( $\pm$ 0.66)	<b>&lt;0.001</b>
S10 – Annoyed	0–6	3.83 ( $\pm$ 2.06)	0.96 ( $\pm$ 1.38)	<b>&lt;0.001</b>
S11 – Depressed	0–6	3.19 ( $\pm$ 2.30)	0.26 ( $\pm$ 0.69)	<b>&lt;0.001</b>
Skindex-16-Emotions	0–100	57.6 ( $\pm$ 29.7)	12.3 ( $\pm$ 14.9)	<b>&lt;0.001</b>
S12 – Interactions with Others	0–6	2.64 ( $\pm$ 2.34)	0.17 ( $\pm$ 0.51)	<b>&lt;0.001</b>
S13 – Being with Others	0–6	2.28 ( $\pm$ 2.38)	0.10 ( $\pm$ 0.37)	<b>&lt;0.001</b>
S14 – Showing Affection	0–6	1.89 ( $\pm$ 2.17)	0.10 ( $\pm$ 0.37)	<b>&lt;0.001</b>
S15 – Daily Activities	0–6	2.85 ( $\pm$ 2.17)	0.42 ( $\pm$ 1.01)	<b>&lt;0.001</b>
S16 – Hard to Work	0–6	2.68 ( $\pm$ 2.37)	0.30 ( $\pm$ 0.80)	<b>&lt;0.001</b>
Skindex-16-Functioning	0–100	41.2 ( $\pm$ 34.5)	3.9 ( $\pm$ 7.8)	<b>&lt;0.001</b>

### (B203.5) Assessing what really matters to individuals with upper limb amputation: the patients and clinicians' perspective

Diana Zidarov, University of Montreal, Montreal, Quebec, Canada; Catherine Robert, MSc, Institut universitaire sur la réadaptation en déficience physique de Montréal, Montreal, Quebec, Canada; Cyril Duclos, PhD, University of Montreal, Montreal, Quebec, Canada; Sara Ahmed, PhD, McGill University, Montreal, Quebec, Canada

**Aims:** Loss of an upper limb has an extensive impact on people's lives, including function, sensation, communication and appearance. Major upper limb amputation (ULA) requires comprehensive and multidisciplinary rehabilitation. Use of standardized measurement is essential for assessing the patient's needs, determining the type of prosthesis required, following progress, and assessing treatment effectiveness. Deciding on what to measure in clinical practice must include the voice of patients and clinicians to ensure outcomes reflect what is important to individuals with ULA. The objective of this study was to identify the most important domains of health-related quality of life (HRQoL) affected by ULA from the perspective of individuals with ULA and clinicians. **Methods:** A cross sectional study was conducted from 2016 to 2020 with 33 individuals with ULA admitted for functional rehabilitation. The Patient Generated Index, an individualized measure of quality of life, was used to assess the most important domains of quality of life affected by the amputation. An electronic survey was administered to clinicians ( $n = 40$ ) in two-rehabilitation centers providing multidisciplinary services to persons with ULA. The survey comprised a list of HRQoL domains from the Patient-Reported Outcomes Measurement Information System (PROMIS) framework to be ranked by importance. Areas affected by ULA identified in the PGI were mapped to The International Classification of Functioning, Disability and Health (ICF). **Results:** According to clinicians, the five most relevant HRQoL domains to be assessed in routine clinical care were upper extremity function, physical function, pain interference, ability to participate in social roles and activities and pain intensity. The five areas that were the most valued by individuals with ULA were recreation and leisure,

driving, remunerative employment, preparing meals, and unclassified domains such as physical appearance and being independent. In total, these represented 54% of all nominated areas. **Conclusion:** These domains provide the most valued and relevant domains to be addressed by multidisciplinary rehabilitation teams and highlight some discrepancies between patient and clinician perspectives. The results can be used in clinical care for joint decision-making and treatment planning and to identify appropriate outcome measures to assess the outcomes of multidisciplinary interventions for individuals with ULA.

### (B203.6) Developing a validated patient-reported outcome instrument to assess financial hardship among hematopoietic cell transplant recipients for sickle cell disease

Staci Arnold, MD, MBA, MPH, Afac Cancer and Blood Disorders Center at CHOA and Emory University, Atlanta, Georgia, United States; Diana Ross, Emory University, Atlanta, Georgia, United States; Cynthia Sinha, Emory University, Atlanta, Georgia, United States; Crystal Smith, Emory University, Atlanta, Georgia, United States; Lakshmanan Krishnamurti, MD, Afac Cancer and Blood Disorders Center at CHOA and Emory University, Atlanta, Georgia, United States

**Aims:** Allogeneic hematopoietic cell transplantation (HCT) is the only proven curative therapy for sickle cell disease (SCD). Due to the high risk of infection during the early post-transplant period, HCT requires weeks of inpatient hospitalization and physical isolation of both patient and caregiver. This translates into missed work and wages but improves survival rates. The resulting short- and long-term financial ramifications are not adequately reported. This study documents the development of a survey to query these factors among HCT recipients for SCD with the following aims: to develop and adapt an existing financial hardship measure and to validate the measure for use in this population over time **Methods:** A literature review identified less than 5 existing measures of financial hardship among SCD, HCT, and/or pediatric hematology/oncology. After, soliciting expert opinion and end-user input, the 43-item questionnaire assessing financial hardship, income, employment, and insurance status developed at the Dana Farber Cancer Institute was most amenable to use in our population. Consultation with the survey developer resulted in a modified 38 question patient-reported financial hardship assessment tool. To perform concept elicitation, test-retest reliability, and validate content, internal consistency, convergence, and divergence, we surveyed 5 study participants using standardized interviews. **Results:** Initial validation phases revealed lack of internal consistency between scaled measures and discrepancies around timing of survey responses. Survey participants were able to consistently define key concepts like household income, occupation, health, financial hardship, quality of life, and family leisure activities. Specifically, health was consistently defined as "overall well-being" and "ability to do things" you enjoy. Participants were unaware of resources including family medical leave, retirement, and hospital-based resources. Participants reported an ability to adequately convey the financial impact of HCT within the scope of the survey. **Conclusion:** Iterative survey validation will provide the first validated instrument assessing financial hardship among pediatric and adult HCT recipients for SCD. This instrument will allow providers to better inform and educate patients of HCT financial risks. It will also serve as an integral mechanism to develop patient resources and support. Finally, baseline and annual assessments will allow evaluation of these resources and identification of any ongoing financial hardships.



### (B203.7) Satisfaction of patients with chronic mechanical back pain with physiotherapy management in a low-resource setting

Adesola Odole, Ph.D, Department of Physiotherapy, College of Medicine, University of Ibadan, Ibadan, Nigeria; Olu Awosoga, Ph.D, Faculty of Health Sciences, University of Lethbridge, Lethbridge, Alberta, Canada; Blessing Adolo, B. Physiotherapy, Department of Physiotherapy, College of Medicine, University of Ibadan, Ibadan, Nigeria; Ayomikun Ayodeji, B.Physiotherapy, Department of Physiotherapy, College of Medicine, University of Ibadan, Ibadan, Nigeria; Henrietha Nwankwo, M.Sc., Emerging Researchers and Professionals on Ageing- African Network, Port Harcourt, Nigeria; Chidozie Mbada, Ph.D, Department of Medical Rehabilitation, College of Health Sciences, Obafemi Awolowo University, Port Harcourt, Nigeria

**Aims:** Neck pain is the third most common musculoskeletal pain and a leading cause of morbidity and disability in older adults. With healthcare shifting towards a more patient-centered approach and patient satisfaction emerging as a critical outcome of care, there is a need to consider the views and opinions of patients to further improve healthcare delivery. Thus, this study was aimed to determine the level of patients' satisfaction with physiotherapy in the management of chronic mechanical neck pain (CMNP). **Methods:** This study involved a convergent parallel mixed-method design of a cross-sectional survey (CSS) and a qualitative study. For the CSS, participants were selected purposively and data were collected using the MedRisk instrument for Measuring Patient Satisfaction with Physiotherapy. Data were analyzed using inferential statistics of Mann–Whitney  $U$  and Kruskal–Wallis tests at  $p \leq 0.05$ . Using a phenomenology qualitative approach, five purposively selected patients from two of the out-patient physiotherapy clinics participated in a Focus Group Discussion (FGD). Data were analyzed using content thematic analysis. **Results:** Participants (28 females; 23 males) for the CSS were aged  $54.24 \pm 14.08$  years. Almost half (49.0%) reported excellent satisfaction level with physiotherapy in the management of CMNP. Only a few (4) reported a fair satisfaction level. There was no significant difference in the level of satisfaction between female and male patients ( $U = 280.500$ ;  $p = 0.395$ ). There was no significant difference in the level of satisfaction on the basis of marital status ( $H = 3.603$ ;  $p = 0.165$ ). Participants for the qualitative study (4 females; 1 male) were aged  $62.8 \pm 6.85$  years. Four themes (patients' experience with physiotherapists; patient perception about physiotherapy services; patient satisfaction with physiotherapy services; patient satisfaction with other health care services) and eight sub-themes emerged from the discussion. The findings from the FGD further explained that participants were satisfied with physiotherapy management of their neck pain. **Conclusion:** Patients being managed for CMNP at selected outpatient physiotherapy facilities in Nigeria are satisfied with physiotherapy care for their CMNP. However, areas of improvement such as collaboration between physiotherapists and the records office in booking appointments to improve convenience and compliance with physiotherapy appointments were identified.

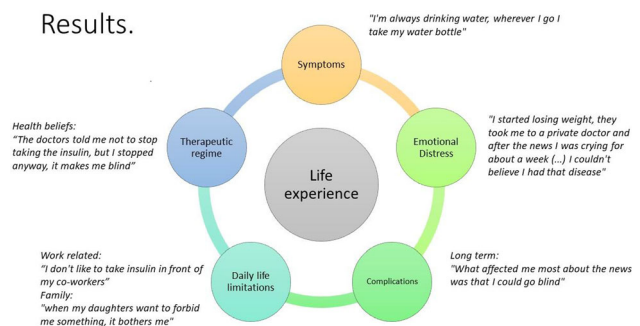
### (B203.8) Patients' experience of living with diabetes in Ecuador: preliminary results of focus groups

Jimmy Martin-Delgado, MD, Foundation for the Promotion of Health and Biomedical Research, Sant Joan, Spain; Jose Joaquin Mira, PhD, Health District Alicante-Sant Joan, Alicante, Spain; Mercedes Guilabert, PhD, Miguel Hernández University, Elche, Spain

**Aims:** Diabetes is considered a global health problem. In Ecuador it is the second cause of death, only after ischemic heart disease. The objective of this study was to explore the experience of "living with

diabetes" of the users of the national health system of Ecuador with the intention of developing the first PROM tool for the Ecuadorian environment. **Methods:** Qualitative research that included four focus groups and six semi-structured interviews with adults with type 2 diabetes treated in primary care. A purposive sampling strategy was used to recruit individuals who might be interested in discussing their life experience. All participants voluntarily agreed to participate and signed an informed consent, all sessions were recorded in audio and subsequently transcribed. In order to obtain culture, beliefs, demographic, diet, type of treatment and degree of engagement, participants from the highlands, coastal, indigenous population and urban or rural areas were included. The information was analyzed based on the following mutually exclusive categories: personal, social and occupational dimensions of the disease. Information capture was continued until data saturation was reached. **Results:** 10 men and 32 women between 30 and 75 years old participated. Of these, 19 participants belonged to rural areas and 23 participants to urban areas. Among the most prevalent symptoms, thirst was described as a persistent cause of discomfort, with fatigue added. A proportion of the participants accepted their pathology but not the treatment, this is motivated by a high prevalence of alternative treatments, the lack of information, the low level of health literacy and "fear" to insulin therapy. Therapeutic goals agreed with the patients were not set. Among the main fears are the long-term complications (diabetic nephropathy and retinopathy) since this would detract from the autonomy they maintain and limit them from leading a "normal life." **Conclusion:** Developing these tools respond to the objectives of achieving patient-centered care and, therefore, add value to health care by expanding the indicators that monitor the quality of assistance provided. This type of procedure allows patients to be involved in the care process, thus establishing a framework to achieve better clinical results and greater patient satisfaction with the system.

### Results.



### Participants.

- All participants were users of the ministerial health services.
- Coastal region:
  - Manabi (highest prevalence of diabetes)
  - Guayas (most populated province)
- Andes region:
  - Pichincha (includes the capital city)
  - Cotopaxi (highest prevalence of indigenous community)

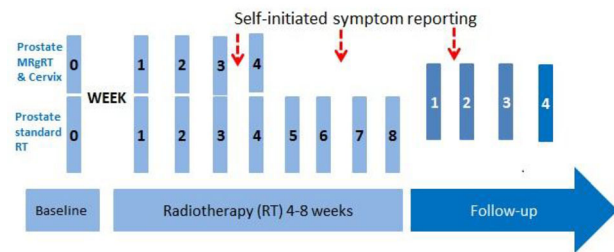


### (B203.9) Development and preliminary testing of a pelvic item set for acute patient-reported toxicity to MR-guided radiotherapy

Pia Krause Møller, MPH, PhD student, Odense University Hospital, Odense NV, Denmark; Pia Krause Møller, BN, MPH, PhD student, Research Unit of Oncology, Odense, Denmark; Helle Pappot, Professor, Department of Oncology, Rigshospitalet, University Hospital of Copenhagen., Copenhagen, Denmark; Tine Schytte, Chief physician, Associate professor, Ph.D., Department of Oncology, Odense University Hospital, Institute of Clinical Research, University of Southern Denmark, MANTRA., Odense, Denmark; Uffe Bernchou, Physicist, Associate Professor, Ph.D., Laboratory of Radiation Physics, MANTRA (New MAGnetic resonance Technology for Response Adapted radiotherapy) Frontline Research Center, Odense University Hospital and Department of Clinical Research, University of Southern Denmark., Odense, Denmark; Karin Brochstedt Dieperink, RN, Ph.D., Associate Professor, Department of Oncology, Odense University Hospital. Deputy Head at Department of Clinical Research, and Head of Research at Family focused healthcare research Centre (FaCe), University of Southern Denmark (SDU). Head of the Danish Family Nursing Association., Department of Oncology, Odense University Hospital, Academy of Geriatric Cancer Research (AgeCare) and Department of Clinical Research, University of Southern Denmark, Odense, Denmark

**Aims:** In cancer treatment, the MR-linac is a new technology providing magnetic resonance guided radiotherapy (MRgRT). This makes it possible to minimize the treatment volume potentially affecting the acute toxicity. Electronic patient-reported outcomes (ePROs) can be used to evaluate the acute toxicity of the patients with this new technology. To our knowledge, no PRO instrument is developed catching the acute toxicity during pelvic MRgRT. The objective of this study was to develop and test a pelvic item set for MRgRT. **Methods:** The study is a mixed-methods study in two phases. An initial item selection through 1) a literature review and 2) a journal audit of clinician-reported toxicity for pelvic patients at the MR-linac. The items in phase 1 were applied in a prospective cohort in phase 2 for weekly reporting during radiotherapy and 4 weeks after (Fig. 1). Self-initiated symptom reporting was also possible at any time. A cut-off of > 20% having reported the symptom in both groups was the criteria for the final model. Patients referred for primary pelvic radiotherapy (MR-guided or standard) were included. Here we present preliminary data from 5 months of data collection. **Results:** Until now, 33 patients (25 with prostate cancer and 7 with cervical cancer) were included. The initial item selection resulted in a pelvic item set with 18 symptomatic adverse events (AEs) from the PRO-CTCAE library (NCI) and EORTC item library being tested in the pilot study. All 18 acute AEs were reported in both patient groups. However, in the preliminary data two symptoms is below the cut-off (vomiting and blood in stools) suggesting a pelvic item set with 16 AEs. For cervical cancer patients there is a need for an add-on to the model with diagnose-specific symptoms like vomiting, tinnitus, head-ache and pain in the irradiated area. **Conclusion:** A pelvic item set was developed and the preliminary data from the test of the model in the pilot study points at a model with 16 acute symptomatic adverse events being relevant for both prostate and cervical cancer patients undergoing radiotherapy. There is, however a need for a diagnose-specific add-on for cervical cancer patients.

### ePRO during pelvic radiotherapy



### (B203.10) How to foster successful implementation of a PREM in the disability sector? An example of developing strategies in co-creation

Marjolein van Rooijen, Master, Maastricht University, Maastricht, Netherlands; Stephanie Lenzen, PhD, Zuyd University of applied sciences, Heerlen, Netherlands; Anneke van Dijk—de Vries, PhD, Maastricht University, Maastricht, Netherlands; Ruth Dalemans, PhD, Zuyd University of applied sciences, Heerlen, Netherlands; Albine Moser, PhD, Zuyd University of applied sciences, Heerlen, Netherlands; Anna Beurskens, PhD, Maastricht University, Heerlen, Netherlands

**Aims:** Successful implementation of a PREM within an organization depends on the involvement of stakeholders from different levels. This is challenging in the disability sector, in which not only managers, team leaders and care professionals need to be involved in the implementation process, but also care users who are communication vulnerable. This study aims to provide insight into supportive pre-conditions for valuably engaging communication vulnerable stakeholders, and the impact of engaging relevant stakeholders on the development of implementation strategies. **Methods:** Participatory Action Research (PAR) was used to develop strategies: (1) development of draft strategies, (2) testing of usability in context, (3) reflection and evaluation, and (4) development of final strategies. Two types of groups, a project group and two development groups, met on a regular basis. A project group initiated draft strategies based on a previously performed problem analysis and consisted of one care-user, two care professionals, four management level employees and four researchers. The concept strategies were iteratively tested between the project group and development groups, composed of communication vulnerable care users ( $n = 8$ ) and professionals ( $n = 12$ ). Data collection consisted of audio tapes, reports, and researchers' notes. We performed directed content analyses. **Results:** Supportive preconditions enabling communication vulnerable stakeholder engagement were co-creative methods, an equal number of care users and professionals per group, delimited sessions with a focus on one goal, physical concepts of draft strategies, to-the-point questions and visual session reports. The impact of the development group on strategies was mainly on the lay-out of micro level-oriented strategies in using bright colors, drawings, photos, pictos or smileys and using short sentences, using comprehensible words and having one focus per strategy. The project group impact was identified in all PAR steps, as they iteratively reflected upon development group outcomes, and had impact on content and lay-out decisions of the final implementation strategies. **Conclusion:** Engaging all relevant end-users in PREM implementation strategy development, including users who are communication vulnerable, is feasible. It requires adaptation in communication, delimited sessions, a safe group environment and to equip every participant to have impact on tailoring the implementation strategies to their needs, preferences and routines.

Screenshot of minutes (in Dutch) drawn from a development group session using communication supportive strategies.

**191218 Verslag 6e ontwikkelgroep Aan de Pas**

**Infographic**

- Met de **infographic** wordt het doel van 'Dit vind ik ervan!' duidelijker gemaakt.
- De infographic is een **poster** die over locaties van SGL verspreid kan worden.
- Er is versie voor **cliënten** en een versie voor **begeleiders**.

**Film**

- Met de **film** wordt het doel van 'Dit vind ik ervan!' duidelijker gemaakt.
- We hebben niet zelf een **film** gemaakt.
- Samen is **besloten** dat de **bestaande film** ingezet moet gaan worden op **SGL-locaties**.
- De **bestaande film** werd nog niet gebruikt bij SGL.
- Dit is de link om de **film** te bekijken: {HYPERLINK "<https://www.youtube.com/channel/UCxdqoKz0Ox7b8-ZngF2gmZw>"}

**Zakboekje**

- Het doel van het **zakboekje** is:
  - o **Helpen herinneren** voorafgaand aan en tijdens het '**Dit vind ik ervan!** gesprek
  - o **Toegankelijke informatie** voor **begeleider en cliënt**
- Met elkaar hebben we een **versie** voor **cliënten** gemaakt.
- Op het **zakboekje** was nog feedback
  - o De **websites** op pagina 2 en 3 staan **onvolledig**
  - o De tekst **ik bereid me voor** komt 'dwingend' over. Deze kan helemaal weg.





## B204: Methods II

### (B204.1) An Optimal Short-form Chinese Nonrestorative Sleep Scale

Daniel Fong, PhD, The University of Hong Kong, Hong Kong, Hong Kong; Sha Li, The University of Hong Kong, Hong Kong, Hong Kong; Janet Wong, The University of Hong Kong, Hong Kong, Hong Kong; Kate Wilkinson, University of Toronto, Toronto, Ontario, Canada; Colin Shapiro, University of Toronto, Toronto, Ontario, Canada; Edmond Choi, The University of Hong Kong, Toronto, Hong Kong; Bradley McPherson, The University of Hong Kong, Hong Kong, Hong Kong; Esther Lau, The Education University of Hong Kong, Hong Kong, Hong Kong; Cindy Lam, The University of Hong Kong, Hong Kong, Hong Kong; Lixi Huang, The University of Hong Kong, Hong Kong, Hong Kong; Mary Ip, The University of Hong Kong, Hong Kong, Hong Kong; Mary Ip, The University of Hong Kong, Hong Kong, Hong Kong

**Aims:** Nonrestorative sleep has gained increasing attention as a treatment target. It can be assessed by the 12-item Nonrestorative Sleep Scale (NRSS) which has been translated and tested in Hong Kong Chinese. However, whether the length of the instrument can be reduced without compromising its validity and reliability has not been explored. Hence, this study aimed to determine whether a shortened version of the Chinese NRSS could validly and reliably facilitate the assessment of NRS in research and clinical practice. **Methods:** We targeted community-dwelling adults in Hong Kong who were

recruited in two cross-sectional studies. They completed a standardized questionnaire that included the Chinese NRSS and Pittsburgh Sleep Quality Index (PSQI). We fitted a Graded Response Model (GRM) and used an iterative Wald test to assess differential item functioning (DIF) by gender. After excluding items showing DIF, optimal test assembly (OTA) was used to obtain a short-form NRSS that did not compromise the test information, concurrent validity, convergent validity with PSQI and internal reliability of the full version. Results: A total of 404 Chinese adults (60% female; mean age: 45 years [range 18–88]) completed the questionnaire. Exploratory factor analysis showed the adequacy of a single factor and local independence with residual correlation ranged from 0.21 to 0.22. No NRSS item showed gender DIF. OTA identified 9 items of the Chinese NRSS that kept at least 95% of the reliability, concurrent validity, and convergent validity as well as 92% of test information of the original 12-item scale. Under GRM, the 9-item shortened version had discrimination and difficulty parameters ranging from 0.92 to 2.68 and – 6.16 to 1.97, respectively, with a Cronbach's alpha of 0.83. Conclusion: The 9-item Chinese NRSS is reliable and a valid alternative to the original full version that can be utilized for more efficient assessment of NRS.

### (B204.2) Measurement properties of the Satisfaction With Life Scale in cardiac arrest survivors

Kristofer Årestedt, Professor, Linnaeus University, Faculty of Health and Life Sciences, Kalmar, Sweden; Johan Israelsson, PhD, Region Kalmar County, Kalmar, Sweden; Ina Marteinsdottir, Professor, Linnaeus University, Faculty of Health and Life Sciences, Kalmar, Sweden; António Macedo, PhD, Linnaeus University, Faculty of Health and Life Sciences, Kalmar, Sweden; Evalill Nilsson, PhD, Linnaeus University, Faculty of Health and Life Sciences, Kalmar, Sweden; Mikael Rask, Associate professor, Linnaeus University, Faculty of Health and Life Sciences, Kalmar, Sweden; Jalal Safipour, Associate professor, Linnaeus University, Faculty of Health and Life Sciences, Växjö, Sweden; Hanna Hanna Tuveson, Associate professor, Linnaeus University, Faculty of Health and Life Sciences, Växjö, Sweden; Amanda Hellström, PhD, Linnaeus University, Faculty of Health and Life Sciences, Kalmar, Sweden

**Aims:** Few patient-reported outcome measures (PROMs) have been developed or validated for use in cardiac arrest (CA) survivors. Life satisfaction is an important outcome and the aim was therefore to evaluate the factor structure and reliability of the Satisfaction With Life Scale (SWLS) in CA survivors. **Methods:** A postal questionnaire containing demographic questions and the SWLS was sent to 251 survivors six months after the CA. We used a Swedish version of SWLS that includes 5 items with a 5-point Likert response format. Data about the CA was taken from the Swedish Register of Cardiopulmonary Resuscitation. To evaluate the hypothesized one-factor structure, confirmatory factor analysis (CFA) for ordinal data were used (polychoric correlations and WLSMV estimation). The reliability was evaluated using ordinal alpha (ordinal version of Cronbach's alpha) and composite reliability coefficients. The analyses were conducted in R 4.0.0, including the Psych and Lavaan packages. Results: The final sample consisted of 212 CA survivors, 136 in-hospital and 76 out-of-hospital. The mean age was 66.6 (SD = 11.9) years and 23.6% were females. The polychoric correlations across items ranged between 0.64 and 0.88. The hypothesized one-factor model was overall supported after the residual variance of item 4 and 5 were allowed to correlate: RMSEA = 0.11, 95% CI 0.05/0.17, Pclose = 0.053, CLI = 0.99, TLI = 0.99, SRMR = 0.02. The standardized factor loadings were all significant ( $p < 0.001$ ) and ranged between 0.75 and 0.99. No Heywood cases were identified. The reliability was good: ordinal  $\alpha = 0.95$  and component

reliability = 0.95. Conclusion: Overall, the SWLS with a 5-point Likert response format showed sound measurement properties in the present sample and the instrument can therefore be used, pending further validations, to assess life satisfaction in CA survivors.

#### **(B204.3) A review of the content and psychometric properties of patient-reported outcome measures (PROMs) used to assess cancer-related fatigue in intervention trials**

Rachel Campbell, PhD, University of Sydney, Faculty of Science, School of Psychology, Sydney, NSW, Australia; Renée Bultjnick, MSc, Department of Radiation Oncology and Experimental Cancer Research, Ghent University, Ghent, Belgium; Jasmine Yee, PhD, University of Sydney, Faculty of Science, School of Psychology, Sydney, Australia; Chindhu Shunmuga Sundaram, MSc, University of Sydney, Faculty of Science, School of Psychology, Sydney, Australia; Haryana Dhillon, PhD, University of Sydney, Faculty of Science, School of Psychology, Sydney, Australia; Joanne Shaw, PhD, University of Sydney, Faculty of Science, School of Psychology, Sydney, Australia

**Aims:** Cancer-related fatigue (CRF) is a common symptom experienced by people with cancer, often caused by the disease and/or treatment. CRF places a significant burden on patients and survivors, highlighting the need for effective supportive care interventions to reduce fatigue. A recent systematic review of interventions targeting CRF identified large variability in PROMs used to assess CRF across interventions, limiting comparability of findings and robust conclusions about relative effectiveness. This review aims to evaluate the content and psychometric properties of PROMS used to assess CRF in interventions designed to alleviate fatigue, to inform the selection of robust measures in future studies. **Methods:** All included PROMs were identified from a previous systematic review of CRF intervention trials conducted by our group. General characteristics of each measure were extracted and content was assessed against domains specified by the National Comprehensive Cancer Network (NCCN) definition of CRF. Psychometric properties were evaluated using adapted criteria from the COSensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist. Recommendations for appropriate use were generated by the investigator team. **Results:** We identified 27 measures: 18 were fatigue-specific and 9 were a fatigue subscale or single item within a broader measure (e.g., QLQ-C30 or Edmonton Symptom Assessment System). Seventeen were unidimensional and 10 multidimensional. There was large variability between measures in content, length, number and type of domains covered. The FACIT-Fatigue and Piper fatigue scale were the most commonly used measures to assess CRF in intervention studies. **Conclusion:** A wide range of measures have been used to assess CRF in intervention studies, each varying in content and domains covered, making it difficult to interpret effects across studies. This may be, at least partly, due to lack of consensus on an appropriate conceptual framework and gold standard definition of CRF. The final evaluation of the content and psychometric properties of included measures will be presented, along with recommendations for the selection of CRF measures in future studies.

#### **(B204.4) Incorporating clinicians' CTCAE grades as auxiliary variables when imputing patients' PRO-CTCAE scores via multiple imputation in Alliance A091105**

Gina Mazza, Alliance Statistics and Data Center, Scottsdale, Arizona, United States; Molly M. Petersen, Mayo Clinic, Scottsdale, Arizona, United States; Lauren J. Rogak, Memorial Sloan Kettering Cancer

Center, New York City, New York, United States; Mrinal M. Gounder, Memorial Sloan Kettering Cancer Center, New York City, New York, United States; Michelle R. Mahoney, Alliance Statistics and Data Center, Rochester, Minnesota, United States; Gary K. Schwartz, Columbia University, Rochester, New York, United States; Gita Thanarajasingam, Mayo Clinic, Rochester, Minnesota, United States; Ethan Basch, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States; Amylou C. Dueck, Alliance Statistics and Data Center, Scottsdale, Arizona, United States

**Aims:** The Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) is an item library assessing 78 symptoms from the CTCAE. Patients complete a subset of items to report treatment-relevant symptoms. The National Cancer Institute recommends reporting the PRO-CTCAE in conjunction with the CTCAE. However, missing scores complicate PRO-CTCAE reporting because patients with and without missing scores may systematically differ. Because interest lies in whether CTCAE grades can inform missing PRO-CTCAE scores, we outline criteria for evaluating CTCAE grades' utility as auxiliary variables and compare strategies for imputing PRO-CTCAE scores using Alliance A091105, a phase III, randomized, placebo-controlled trial of sorafenib in patients with desmoid tumors. **Methods:** Patients completed PRO-CTCAE items assessing insomnia, constipation, pain, fatigue, nausea, vomiting, diarrhea, rash, hand-foot syndrome, decreased appetite, and mouth/throat sores before randomization and at each cycle for 8 cycles. Clinicians graded patients' fatigue, papulopustular rash, palmar-plantar erythrodysesthesia syndrome, diarrhea, anorexia, nausea, vomiting, abdominal pain, mucositis oral, hypertension, arthralgia, and myalgia at each cycle. We examined associations between PRO-CTCAE scores for the same symptom at different cycles and between CTCAE grades and PRO-CTCAE scores for the same symptom. We then performed multiple imputation with and without auxiliary variables to assess the added benefit of including CTCAE grades in the imputation model relative to only including PRO-CTCAE scores across all cycles. **Results:** The number of patients who completed the PRO-CTCAE ranged from 63 (baseline) to 30 (Cycle 8; total  $n = 64$ ). The strongest correlations occurred between PRO-CTCAE scores for the same symptom at different cycles. Correlations between CTCAE grades and PRO-CTCAE scores for the same symptom varied widely, though many exceeded .40 (e.g., fatigue, diarrhea). Multiple imputation with and without CTCAE grades as auxiliary variables yielded similar results, though complications arose due to characteristics of the CTCAE (e.g., data sparseness, including no clinician-reported grade 1 + vomiting at Cycle 8). **Conclusion:** Properly addressing missing PRO-CTCAE scores supports more accurate and generalizable causal inferences regarding treatment tolerability. CTCAE grades can inform missing PRO-CTCAE scores, though model simplicity and computational ease may support other strategies. **Support:** UG1CA189823; U01CA233046; Daniel J. Sargent, Ph.D., Career Development Award in Cancer Research; Bayer; <https://acknowledgments.alliancefour.org>.

#### **(B204.5) Practical guidance for the meaningful analysis of patient-reported data stored in clinical registries**

Jae-Yung Kwon, University of British Columbia, Vancouver, British Columbia, Canada; Richard Sawatzky, Trinity Western University, Langley, British Columbia, Canada; Jennifer Baumbusch, University of British Columbia, Vancouver, British Columbia, Canada; Sandra Lauck, Centre for Heart Valve Innovation, St. Paul's Hospital, Vancouver, British Columbia, Canada; Pamela A. Ratner, University of British Columbia, Vancouver, British Columbia, Canada

**Aims:** Although patient-reported outcomes (PRO) data collected for clinical registries have the potential to provide valuable information about the impact of treatment, there is little guidance about how to meaningfully analyze these data. Our aim is to explicate the process of conducting such analyses as a guide by answering the following questions: (1) What information can be extracted from linked clinical and administrative data sources? (2) How should we best accommodate the unavoidable missing data within these datasets? and (3) How do we determine an appropriate analysis strategy for longitudinal PRO data? **Methods:** The primary methods included observations and reflections gained from the longitudinal analysis of PRO data (Atrial Fibrillation Effect on Quality of Life Questionnaire) collected between 2008 and 2016 from a provincial cardiac registry in British Columbia (Canada) and linked with administrative health data. This presentation focuses on key methodological challenges and solutions while working with these data (rather than on the results of the statistical analyses). **Results:** A large part of the study was focused on how to represent change over time in PROs, and then establish a defensible process to represent the individual variation in the patients' trajectories. To extract information from linked data sources, substantial data preparation and cleaning were needed. Relevant information including patients' comorbidities and interventions received were extracted with validated algorithms and graphical techniques were employed to visualize relevant information (e.g., stacked bar plots and times series plots). To accommodate missing data, multilevel multiple imputation was used with all available auxiliary variables. Full information maximum likelihood was used to account for unequal number and duration of follow-ups. To model patient trajectories, several methods including multilevel and latent growth models were compared. However, an emerging statistical approach known as growth mixture modeling allowed for the identification of multiple subgroups of trajectories and nesting of time to account for individually varying times of observation. **Conclusion:** Analysis of PROs in clinical registries requires complex computer programming and knowledge of various analytical software packages. Meaningful analysis methods need to account for both the different subgroups of patient trajectories and variability in the frequency and timing of measurement occurrences.

#### **(B204.6) Defining meaningful change for the nasal polyps score (NPS) and nasal congestion score (NCS)**

Jessica Braid, Roche, Welwyn Garden City, United Kingdom; Cristian Gugu, PhD, Clinical Outcomes Solutions, Chicago, Illinois, United States; Lutaf Islam, DVM MSc, Roche, Welwyn Garden City, United Kingdom; Theodore Omachi, MD, Genentech, San Mateo, California, United States; Helen Doll, DPhil, Clinical Outcomes Solutions, Folkestone, United Kingdom

**Aims:** To determine meaningful change thresholds (MCTs) and minimal important differences (MIDs) for the co-primary endpoints NPS and NCS from two Phase 3, randomized, double-blind, placebo-controlled clinical trials (GA39668 and GA39855) of omalizumab (Xolair®) in patients with chronic rhinosinusitis with nasal polyps. **Methods:** Analyses focused on assessment of meaningful change (estimation of MCT and MID) using anchor-based methods (including ROC analysis, CDF and PDF curves), supported by distribution-based methods, by assessing within and between anchor group differences in NPS and NCS change scores. A nasal polyp-specific subscale of the Sino-nasal Outcome Test-22 (SNOT-22) was developed using exploratory factor and Rasch analyses. The identified MCTs and MIDs were then used in unblinded responder analyses to compare the proportions of patients in each treatment group achieving a meaningful improvement and group-level differences (mean treatment effect). **Results:** The SNOT-22 total and SNOT-22 Sino-nasal

Symptoms Subscale scores were sufficiently correlated ( $\geq 0.30$ ) with the NPS and NCS to be used as anchors. Based on anchor-based analyses, MCTs of  $-1.0$  for the NPS and  $-0.5$  for the NCS were estimated; MIDs were  $-0.5$  and  $-0.35$ , respectively. These estimates were consistent across the studies and different methodological approaches and were supported by slightly smaller distribution-based estimates. Approximately twice as many patients achieved a meaningful change in the NPS and NCS in the omalizumab vs the placebo group (NPS: 55.9% vs. 33.3% GA39668, 58.0% vs. 26.6% GA39855; NCS: 57.6% vs. 25.4% GA39668, 60.0% vs. 35.9% GA39855), with statistically significant differences between the treatment groups. Group differences in mean change were statistically significant and exceeded the estimated MIDs (NPS:  $-1.14$  GA39668,  $-0.59$  GA39855; NCS:  $-0.55$  GA39668,  $-0.50$  GA39855). **Conclusion:** MCTs and MIDs for the NPS and NCS were estimated from GA39668 and GA39855. Use of nasal polyp-specific measures (SNOT-22, SNOT-22 SNSS) supports the relevance of the estimated thresholds which, when applied in unblinded responder analysis, showed that in both studies around twice as many patients in the omalizumab arm as in the placebo arm achieved a change perceived as meaningful by patients. These thresholds could be used in future trials to assess NPS and NCS meaningful change.

#### **(B204.7) Goal quality, education, and cognition: preliminary analysis of self-management goals formulated by people living with HIV**

Maryam Mozafarina, MSc, McGill University, Montreal, Quebec, Canada; Fateme Rajabiyazdi, PhD, McGill University Health Centre, Montreal, Quebec, Canada; Amanda Austin-Keiller, MSc, McGill University, Montreal, Quebec, Canada; Marie-Josée Brouillette, MD, McGill University, Montreal, Quebec, Canada; Lesley Fellows, MD, PhD, McGill University, Montreal, Quebec, Canada; Nancy E. Mayo, PhD, McGill University, Montreal, Quebec, Canada

**Aims:** Goal setting is an important step towards achieving favorable health outcomes. In the context of chronic conditions, it is of great value to understand to what extent patients are capable of setting self-management goals independently. An ongoing goal setting trial is evaluating the effectiveness of a personalized health profile on quality of self-management goals (NCT04175795). A preliminary analysis of the quality of self-defined goals as well as the correlations between goal quality and participants' cognitive ability and education is presented. **Methods:** Participants enrolled in the Canadian HIV Brain Health Now (BHN) study were randomized to either (1) receive their personalized health profile created based on their first and last recorded visit (intervention) or (2) not receive the profile before setting goals (control). Both the intervention and control groups received instructions on goal setting. Self-defined goals in free text were collected through an online platform. The outcome, goal quality, was measured as the number of specific words that matched a goal setting lexicon applied through text mining. Polyserial correlations were calculated with education (5 levels) and cognitive ability, the latter of which was measured using a computerized test and classified as excellent, good, and fair. **Results:** The self-management goals of the first 50 participants (32 French and 18 English) were analyzed. A total 212 goals were formulated (132 in French/80 in English). Text mining identified 1914 usable words of which 437 were specific (236 English/201 French). Regardless of the language, goals were mainly missing "measurable" and "timebound" words; while, specific nouns and actionable verbs, which are representative of the specificity of the goal, were mostly present. Number of ambiguous nouns and neutral verbs was relatively small. The correlation between cognitive ability and goal specificity was strong ( $r = 0.63$ ; 95% CI 0.50–0.76) and with education it was weak ( $r = 0.34$ ; 95% CI 0.21–0.47). **Conclusion:** The

preliminary findings show encouraging signs of the usefulness of text mining techniques in measuring goal quality. People with higher cognitive ability, but not necessarily more education, formed stronger goals indicating that people with cognitive challenges would need more assistance in forming self-management goals.

**(B204.8) Cancer-specific and generic health utilities—The psychometric performance of the EORTC QLU-C10D in comparison to the EQ-5D in three cancer clinical trials**

Eva Maria Gamper, Department of Psychiatry, Psychotherapy and Psychosomatics, Innsbruck Medical University, Innsbruck, Austria, Innsbruck, Austria; Micha Johannes Pilz, Department of Psychiatry, Psychotherapy and Psychosomatics, Innsbruck Medical University, Innsbruck, Austria, Innsbruck, Austria; Femke Jansen, Department of Otolaryngology-Head and Neck Surgery, Amsterdam UMC, Vrije Universiteit Amsterdam, Cancer Center Amsterdam, Amsterdam, Netherlands; I.m. Verdonck-de Leeuw, Department of Clinical, Neuro- and Developmental Psychology, Faculty of Behavioral and Movement Sciences, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, Netherlands; Peter D. Siersema, Departments of Gastroenterology and Hepatology, Erasmus MC, University Medical Center Rotterdam, the Netherlands, and Department of Gastroenterology and Hepatology, University Medical Center Utrecht, the Netherlands, Utrecht, Netherlands; Bram D. Vermeulen, Department for Gastroenterology and Hepatology, Radboud Medical University Center, Nijmegen, The Netherlands, Utrecht, Netherlands; Georg Kemmler, Division of Psychiatry I, Department of Psychiatry, Psychotherapy and Psychosomatics, Medical University of Innsbruck, Innsbruck, Austria, Innsbruck, Austria; Bernhard Holzner, Department of Psychiatry, Psychotherapy and Psychosomatics, Innsbruck Medical University, Innsbruck, Austria, Innsbruck, Austria; Richard Norman, School of Public Health, Curtin University, Perth, Western Australia, Australia, Perth, Australia; Andrew Bottomley, European Organisation for Research and Treatment of Cancer, Brussels, Belgium, Brussels, Belgium; Madeleine King, European Organisation for Research and Treatment of Cancer, Brussels, Belgium, Brussels, Belgium; Madeleine King, Faculty of Science, School of Psychology, The University of Sydney, Sydney, NSW, Australia

**Aims:** The EORTC QLU-C10D (Quality of Life Utility Core 10 Dimensions) is a relatively new preference-based and cancer-specific measure designed for delivering utility weights for use in health economic evaluations. It was derived from the non-preference-based EORTC QLQ-C30 (Quality of Life Questionnaire Core 30) containing 10 of its health dimensions. We investigated the psychometric properties of the QLU-C10D by retrospectively calculating utilities from QLQ-C30 data of already published cancer clinical trials. We compared it with regard to sensitivity and responsiveness with a standard measure in economic evaluations, the EQ-5D-3L (3 level version of the Euroqol 5 Dimensions). **Methods:** We extracted data of three cancer randomized controlled clinical trials (RCTs) that had collected both the EQ-5D and the QLQ-C30, of which we calculated QLU-C10D utilities. We assessed the instruments' accordance (Bland–Altman plots and intraclass correlations (ICCs)) and relative efficiencies (REs) in detecting known group differences and changes over time separately for each trial and for respective outcome measurement time points. **Results:** The RCTs targeted different cancer populations and interventions including surgical, radiotherapeutic, and psychosocial interventions and included samples sizes between 156 and 209 patients. ICCs and Bland–Altman plots indicated an overall moderate to good agreement between the measures (ICCs between 0.523 and 0.678). Ceiling effects were higher for the EQ-5D (8.5–20.8%) than for the QLU-C10D (0.7–4.8%) throughout all trials

and assessment time points. There was agreement with regard to the ability to detect known groups/changes over time in the majority (67%) of performed tests. The measures' efficiencies varied across trials and across measurement time points. In 70% of cross-sectional known group comparisons, the EQ-5D showed a higher RE and in 67% of the responsiveness tests the QLU-C10D showed higher RE. **Conclusion:** Results indicate that the QLU-C10D and the EQ-5D in overall seem to capture a similar construct but with differing efficiencies, which will require further investigation including additional data in different cancer populations. Such analyses are currently being performed on additional trial data sets and prospectively collected data as part of an international research project. More results will be available by the time of the conference.

**(B204.9) A dictionary-based text mining approach for identifying patient-centric symptoms from patient interviews**

Xiaoran Li, Adelphi Values, Boston, Massachusetts, United States; Michael DeRosa, Adelphi Values, Boston, Massachusetts, United States; Brad Padilla, Adelphi Values, Boston, Massachusetts, United States; Benjamin Banderas, Adelphi Values, Boston, Massachusetts, United States; Linette Duluc, Adelphi Values, Boston, Massachusetts, United States

**Aims:** Concept elicitation (CE) is a critical first step in developing a patient-reported outcome (PRO) that is fit for purpose in a specific disease population. As part of the CE process, open-ended interviews are typically conducted with experts and/or patients. Following completion of interviews, qualitative data analysis is conducted using patient interview transcripts as source data via a qualitative analytical program (e.g., Atlas ti). While there are many robust methodologies to qualitative CE analyses, care must be taken to be consistent in the coding approach. Deviations can present inconsistent interpretation of output, gaps in capturing important concepts and is time consuming if re-analysis must be conducted. The authors sought to explore a complementary and confirmatory methodology utilizing text mining technique to confirm consistency and confidence in primary qualitative analysis. **Methods:** To identify the patient-centric symptoms of a dermatological disease from 23 interviews for patients with the condition, a dictionary-based text mining approach was applied. Traditionally, qualitative analysis involves a theoretical approach (e.g., grounded theory, ethnography, etc.), which may include the development of a coding framework and collaboration among team members (i.e., coders) who organize and catalogue text data from the transcript. The proposed text mining approach firstly creates an a priori dictionary containing 17 symptoms from a previously conducted literature review and expert interviews. Subsequently, an R program (R version 3.5.1) was employed to automatically read the interview transcripts, identify the patient-centric symptoms based on the a priori dictionary, and extract relevant patient quotes to cross-validate with the traditional method. **Results:** Fifteen patient-centric symptoms were identified by the proposed dictionary-based text mining method, whereas 16 were identified by the traditional qualitative coding method in Atlas.ti. The agreement rate in extracted patient quotes of the relevant symptoms was 90% between the two methods. The proposed text mining approach took less than 2 h with most of time spent on constructing the dictionary. **Conclusion:** There is preliminary evidence supporting the proposed dictionary-based text mining as a complementary approach to identifying patient-centric symptoms from the qualitative data.

### **(B204.10) Natural language processing and sentiment analysis on bladder cancer patients' goals before and after cancer surgery**

Yuelin Li, PhD, Memorial Sloan Kettering Cancer Center, New York, New York, United States; Thomas Atkinson, PhD, Memorial Sloan Kettering Cancer Center, New York, New York, United States; Bernard H. Bochner, MD, Memorial Sloan Kettering Cancer Center, New York, New York, United States; Bruce D. Rapkin, PhD, Albert Einstein College of Medicine/Montefiore Medical Center, Bronx, New York, United States

**Aims:** Sentiment Analysis (SA) is an increasingly popular Natural Language Processing technique to quantify affective states such as survey responses, online reviews, social media postings and QOL [https://doi.org/10.1016/j.ijmedinf.2019.02.008]. We applied SA to interview data collected as part of a prospective, longitudinal study of QOL in patients undergoing radical cystectomy and urinary diversion for bladder cancer (ClinicalTrials.gov NCT00745355). **Methods:** Patients were asked: "For you to have the most satisfying life possible, what are the main things you want to accomplish or prevent?" using Ideographic QOL interview probes at baseline (before surgery), 6, and 12 months post-surgery. R package tidytext was used to calculate the sentiment on a scale from  $-5$  (very negative) to  $+5$  (very positive) (1022 entries). Bayesian hierarchical linear modeling fitted sentiment changed over time and across surgical interventions. **Results:**  $n = 473$  bladder cancer patients were enrolled (217 and 256 underwent Neobladder and 256 Ileal Conduit (IC) reconstruction, respectively; age 65 years; 18% female; and 95% White). On "things to accomplish," the IC group at baseline had a statistically reliable positive sentiment of 1.53 (95% posterior HDR = (1.13,2.11), excluding zero). Neobladder group had a marginally lower sentiment ( $-0.33$ , HDR = ( $-0.90,0.35$ )). At 6-months, the IC group had a reduced positive sentiment by  $-0.33$  (HDR = ( $-0.80,0.22$ )) that improved by 0.30 at 12-months (HDR = ( $-0.76,0.29$ )). By contrast, the Neobladder group had an increased positive sentiment at 6-months (0.38, HDR = ( $-0.28,1.16$ )) and by 12-months returned to a comparable sentiment as IC. "Things to prevent" prompted a negative sentiment at baseline (IC =  $-1.14$ , HDR = ( $-1.40, -0.82$ ); Neobladder =  $-1.06$ , HDR = ( $-1.43, -0.62$ )). At 6 months, sentiment remained negative (IC =  $-0.97$ , ( $-1.27, -0.60$ ); Neobladder =  $-0.95$ , ( $-1.16, -0.74$ )) and stayed negative up to 12 months (IC =  $-1.22$ , ( $-1.46, -0.97$ ); Neobladder =  $-1.15$ , ( $-1.37, -0.92$ )). **Conclusion:** SA holds promise in QOL/PRO research, as shown in the Neobladder reconstruction showing a slightly greater positive sentiment than the IC at 6-months. This subtle increase faded by 12-months, when presumably both groups returned to a mildly positive outlook in their recovery. Overall, patients at baseline had a mildly positive sentiment on hopes for surgery success and recovery. Thinking on "things to prevent" elicited a statistically discernible negative sentiment for up to 12 months. SA captures patient experience without the need for conventional fixed-length assessment measures.

## **B205: Clinical research II**

### **(B205.1) Hearing aid use and quality of life in a diverse population of older adults**

C. Makenzie Danley, Doctorate in Audiology, The University of Kansas Medical Center, Kansas City, Kansas, United States; Frances M. Yang, PhD, The University of Kansas Medical Center, Kansas City, Kansas, United States

**Aims:** The purpose of this study is to determine the association between hearing aid use and quality of life in a diverse population of older adults, while controlling for deaf status, sociodemographic

characteristics, and common chronic health conditions among older Americans, aged 65 and older. **Methods:** Using the 2019 National Health and Aging Trend Study (NHATS,  $n = 4826$ ). The quality of life score is represented by 11 well-being items determined from a bifactor analysis that showed sufficient unidimensionality. Hierarchical multiple regression analysis was used to determine the significance, direction, and magnitude of the relationship between hearing aid use and quality of life while controlling for common chronic health conditions and sociodemographic characteristics. **Results:** The quality of life score on average was 32 points (Standard Deviation (SD) = 8.83) out of 41 points for those with hearing aids, while it was 32 points (SD = 9, Range 0–41) for those without a hearing aid. In Model 1, on average, older adults who wear a hearing aid report having 0.88 (Standard Error (SE) = 0.32,  $p < 0.01$ ) points higher on the quality of life score compared to those who do not wear a hearing aid. In Model 2, upon including deaf status, on average, older adults who wear a hearing aid continue to report having higher quality of life ( $b = 0.87$ , SE = 0.32,  $p < 0.01$ ) compared to those who do not wear hearing aids. However, those who are deaf report, on average, 16.5 (SE = 6.34,  $p < 0.01$ ) points lower on the NHATS well-being scale compared to those who are not deaf. After controlling for sociodemographic variables (Model 3), we see that the use of a hearing aid reaches a higher magnitude of 1.63 points (SE = 0.32,  $p < 0.001$ ) indicating a significant association between hearing aid use and sociodemographic characteristics. Model 4 controls for common chronic health conditions in older adults, which shows that hearing aid use continues to have a positive and significant relationship with quality of life ( $b = 0.87$ , SE = 0.28,  $p < 0.001$ ). **Conclusion:** The research concludes that hearing aid use is significantly associated with better quality of life in this national population of older adults.

### **(B205.2) The profiles of upper-limb function and associated risk factors for breast cancer patients after surgery: a cross-sectional study**

Changrong Yuan, PhD, Fudan University, Shanghai, China; Qingmei Huang, Fudan University, Shanghai, China; Fulei Wu, Fudan University, Shanghai, China; Tingting Cai, Fudan University, Shanghai, China

**Aims:** Upper extremity functional limitations are common as a result of breast cancer surgery among breast cancer patients. This study aimed to capture distinct upper-limb function profiles for women with breast cancer, and explore potential risk factors that could be used to predict patients with poor upper-limb function. **Methods:** A cross-sectional study was conducted in patients who had received breast cancer surgery in the past 3 months. The Patient-Reported Outcomes Information Measurement System (PROMIS)-Upper Extremity short form 7a was used to evaluate the upper-limb function of the participants. Latent profile analysis (LPA) was performed to categorize participants into latent subgroups with distinct upper-limb function profiles. Demographic information such as age, marriage status, education level, monthly income was collected. In addition, adverse symptoms including sleep impairment, pain, anxiety, depression, self-efficacy, and social support were assessed. On the other hand, multivariate logistic regression analysis was used to identify the risk factors for patients with poor upper-limb function. **Results:** A total of 187 eligible patients were included for data analysis, with median age of 50 years, ranging from 17 to 77 years. Two profiles were identified by LPA, with profile 1 showed high scores across all 7 items of PROMIS- Upper Extremity, while profile 2 demonstrated low scores across all the items. Additionally, patients in profile 1 were labeled as subgroup of "good upper-limb function." In the contrary, patients with profile 2 were classified as "poor upper-limb function" with a

group size of 38.5%. As for significant risk factors, patients with poor upper-limb function were low income (OR 2.64, 95% CI 1.32–5.26), with worse pain symptom (OR 1.12, 95% CI 1.04–1.21), more anxiety symptom (OR 1.07, 95% CI 1.02–1.12), and lower social support (OR 0.95, 95% CI 0.91–0.98). Conclusion: It was showed in the current study that more attention should be paid to breast cancer patients with poor upper-limb function after surgery. In addition, risk factors identified in this study could be used to help healthcare providers to provide early detection and targeted intervention for vulnerable patients.

### **(B205.3) The use of an electronic patient-reported outcome measure in the management of patients with advanced chronic kidney disease—The RePROM pilot trial**

Derek Kyte, Centre for Patient-Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Jon Bishop, Birmingham Clinical Trials Unit (BCTU), Institute of Applied Health Research, University of Birmingham, Birmingham, UK, Birmingham, United Kingdom; Elizabeth Brettell, Birmingham Clinical Trials Unit (BCTU), Institute of Applied Health Research, University of Birmingham, Birmingham, UK, Birmingham, United Kingdom; Melanie Calvert, Centre for Patient-Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Marie Chadburn, Birmingham Clinical Trials Unit (BCTU), Institute of Applied Health Research, University of Birmingham, Birmingham, UK, Birmingham, United Kingdom; Paul Cockwell, Department of Renal Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK, Birmingham, United Kingdom; Mary Dutton, Department of Renal Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK, Birmingham, United Kingdom; Helen Eddington, Department of Renal Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK, Birmingham, United Kingdom; Gabby Hadley, Department of Renal Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK, Birmingham, United Kingdom; Louise Jackson, Health Economics Unit, Institute of Applied Health Research, University of Birmingham, Birmingham, UK, Birmingham, United Kingdom; Natalie Rowland, Health Economics Unit, Institute of Applied Health Research, University of Birmingham, Birmingham, UK, Birmingham, United Kingdom; Natalie Rowland, Birmingham Clinical Trials Unit (BCTU), Institute of Applied Health Research, University of Birmingham, Birmingham, UK, Birmingham, United Kingdom; Stephanie Stringer, Department of Renal Medicine, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK, Birmingham, United Kingdom

**Aims:** Effective management of patients with chronic kidney disease (CKD) relies on timely detection of clinical deterioration towards end stage kidney failure (ESKF). We aimed to conduct a randomized pilot/feasibility trial of an electronic Patient-Reported Outcome Measure (ePROM) system, which would allow patients with advanced CKD (pre-dialysis) to: (i) remotely self-report their symptoms using a simple and secure online platform; and (ii) share data with the clinical team in real-time via the electronic patient record to help optimize care. **Methods:** We conducted an open-label, two-arm randomized controlled pilot trial of participants  $\geq 18$  years with advanced CKD undergoing outpatient follow-up at Queen Elizabeth Hospital Birmingham (QEHB). The primary outcome was feasibility.

Participants were randomized to receive either usual care, or usual care supplemented with an ePROM intervention accessed through the existing hospital patient portal 'myHealth@QEHB.' Participants within the intervention arm were asked to submit monthly self-reports of their health status using the ePROM system. Measures of study feasibility, participant quality of life, CKD severity and healthcare utilization were assessed at baseline, 3, 6, 9 and 12 months. Qualitative process evaluation was conducted. **Results:** During the 12 month recruitment period, 721 patients were assessed for eligibility within the low-clearance renal clinic at QEHB, 452 were deemed eligible, 166 approached and 52 randomized (intervention  $n = 24$ ; usual care  $n = 28$ ), representing a conversion rate of 31% (monthly recruitment rate = 4.3). Trial follow-up will end in April 2020 and qualitative evaluation is ongoing. To-date:  $n = 19$  patients are in-trial follow-up;  $n = 14$  have completed;  $n = 17$  have exited ( $n = 16$  due to ESKF,  $n = 1$  death); and average case report form return rates across all assessment timepoints  $> 80\%$ . Patients receiving the intervention have returned 63.74% of expected ePROM forms, reporting a total of 487 symptoms (59.55% mild; 36.55% moderate; 3.90% severe) and triggering 14 automated email notifications for severe and current pruritus (42.86%), fatigue (28.57%), shortness of breath (14.29%), difficulty sleeping (7.14%) and ankle swelling (7.14%). **Conclusion:** The full pilot trial results will be presented, including findings from qualitative debriefing interviews aimed at exploring the experiences of patients and healthcare staff involved in the study.

### **(B205.4) Measuring the well-being of cancer patients with generic and a disease specific instruments**

Gang Chen, PhD, Monash University, Melbourne, Australia; Norma Bulamu, Flinders University, Adelaide, Australia; Ellen McGrane, Monash University, Melbourne, Australia; Jeffrey Richardson, Monash University, Melbourne, Australia

**Aims:** It has been argued that generic instruments used to measure patient's quality of life (QoL) may be insensitive to disease-specific patient problems. However, disease-specific instruments may, in turn, be insensitive to more general domains of well-being, which are important for QoL. This paper aims to explore and compare the associations between QoL (as measured by cancer-specific instruments), generic health state utility (HSU) instruments, and by instruments measuring subjective well-being (SWB). Furthermore it examines the view that generic measures may be insensitive to disease-specific symptoms, but that disease-specific (DSU) measures may fail to address the broader domains of life satisfaction captured by SWB instruments. **Methods:** Data were drawn from the Multi-Instrument Comparison survey. The cancer-specific instrument (QLQ-C30) was also transferred onto the EORTC-8D and QLU-C10D utility scores. Linear regression was used to explore to what extent the cancer-specific instrument subscales explain HSU and SWB. The relative importance of seven key life domains in relation to overall life satisfaction was further studied using both a linear and a non-linear quantile regression model. **Results:** All correlations with the QLQ-C30 functional subscales were positive, while correlations with the symptom subscales were negative. The QLQ-C30 subscales explained the majority of the variance in EORTC-8D and QLU-C10D utility scores, followed by generic HSU instruments and SWB instruments. Among the seven life domains studied, achievement in life was the most important. Personal health and relationships were



also significantly more important for cancer patients than for the general public. Both disease-specific HSU instruments were sensitive to the cancer-specific symptoms, to a varied degree. On the other hand, the inclusion of a generic instrument improved the explanation of SWB, which the DSU measures failed to do. Conclusion: The cancer-specific measure (QLQ-C30) is more closely correlated with HSU than SWB measures, and the more sensitive utility measure was the QLU-C10D, which may be complemented by SWB. Consistent with suggestions in the literature, generic HSU instruments were less sensitive to cancer-specific symptoms, meanwhile at least one of these generic measures was more sensitive to broader life domains than the disease-specific instruments.

### (B205.5) Diabetes self-management and the impact of severe mental illness: a comparison of the DAWN-SMI and DAWN2 type 2 diabetes samples

Jennie Lister, Department of Health Sciences, University of York, York, United Kingdom; Jo Taylor, Department of Health Sciences, University of York, York, United Kingdom; Richard I. G. Holt, Faculty of Medicine, University of Southampton, Southampton, United Kingdom; Najma Siddiqi, Hull York Medical School & Department of Health Sciences, University of York, York, United Kingdom; Jan R. Boehnke, School of Health Sciences, University of Dundee, Dundee, United Kingdom

**Aims:** The life expectancy of people living with severe mental illness (SMI) is reduced by ~ 15 years compared to the general population. A higher prevalence of diabetes and its associated complications contribute to this health inequality. A variety of self-management activities have been identified to support health and quality of life in people living with diabetes. We investigate factors affecting diabetes self-management in people living with type 2 diabetes, with and without SMI, and whether SMI is an independent contributor to engaging in self-management activities. **Methods:** We recruited people living with diabetes and SMI (mean age  $\pm$  SD 55.6  $\pm$  10.5 years; 57% male; see table 1) and combined our data with UK general population participants from the DAWN2 study (age 60.4  $\pm$  11.0 years; 59% male). Both surveys covered participants' demographics, physical and mental health, health-related quality of life, type of care received, social support, and diabetes-specific constructs. Both surveys collected data on diabetes self-management using the Summary of Diabetes Self-Care Activities which explores the frequency of engagement in six self-management activities (table 1). We used zero-inflated Poisson and negative binomial regressions to firstly investigate which factors contribute to whether participants engaged in these self-management activities at all; and secondly, which factors are associated with the degree of engagement (i.e., how often participants engaged in an activity). **Results:** In the regression analyses, the only variable consistently associated with whether participants engaged at all, was diabetes empowerment (as measured with the DES-DSF). And with the exception of physical activity, participants living with SMI were less likely to engage in self-management activities. Regarding the degree of engagement, fewer variables (Table 1) were associated with how often participants engaged in an activity and none of the variables showed a consistent association across the six activities. **Conclusion:** The factors contributing to diabetes self-management are diverse. The identified correlates encompass different areas of people's lives and additionally reflect severity of the condition as well as treatment recommendations. The regression analyses showed that participants living with SMI are more likely not to engage in these activities at all; and this association remained when controlling for a range of demographic and condition-related factors.

Table 1. Descriptive statistics of the Summary of Diabetes Self-Care Activities (SDSCA-6) items for people living with diabetes (PWD) and with diabetes and SMI (PWD+SMI); and an aggregate summary of the regression analyses using 15 demographic and psychosocial variables from the DAWN2 protocol.

Self-management activity	PWD+SMI			PWD			R <sup>2</sup>	Regressions	
	N	Mean	SD	N	Mean	SD		N <sub>pred</sub> Engaging	N <sub>pred</sub> Degree
Eating healthily	244	4.07	2.63	419	5.16	2.08	0.43	3	2
30 minutes activity	243	2.61	2.66	419	2.81	2.54	0.47	7	6
Checking feet	242	2.45	2.79	419	3.63	2.74	0.39	6	2
Testing blood sugar	244	2.41	3.14	419	3.28	2.98	0.10	5	1
Testing blood sugar (as recommended)	235	2.77	3.26	419	3.53	3.08	0.54	5	1
Taking medication*	208	6.46	1.55	320	6.49	1.34	0.33	4	2

Note. \*Analysis based only on participants who took medication for their diabetes. R<sup>2</sup> calculated from correlation between observed and predicted values. N<sub>pred</sub> Number of predictors. "Engaging" describes the selection part of a regression model, i.e. whether this activity is performed at all. "Degree" describes the individual differences in how often an activity is performed.

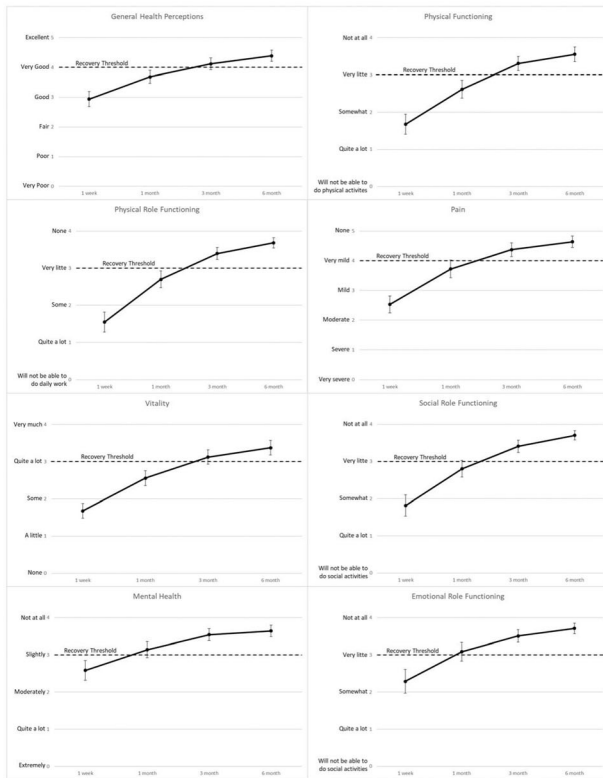
### (B205.6) Preoperative expectations of quality of life outcomes among patients undergoing cancer surgery

Nikhil Panda, MD, Massachusetts General Hospital, Boston, Massachusetts, United States; Ian Solsky, MD MPH, Montefiore Medical Center, Albert Einstein College of Medicine, Bronx, New York, United States; Brandon Neal, MPH, Ariadne Labs, Boston, Massachusetts, United States; Stuart Lipsitz, SCD, Ariadne Labs, Boston, Massachusetts, United States; Christopher Gibbons, PhD, MD Anderson Cancer Center, Houston, Texas, United States; Carrie Lubitz, MD MPH, Massachusetts General Hospital, Houston, Massachusetts, United States; Christy Cauley, MD MPH, Massachusetts General Hospital, Boston, Massachusetts, United States; Alex Haynes, MD MPH, Dell Medical School at The University of Texas at Austin, Austin, Texas, United States

**Aims:** Patient expectations of the impact of surgery on postoperative quality of life (QoL) inform treatment decisions. Patients are provided counseling regarding recovery from surgery, but perceptions of this information remain unknown. We sought to determine expectations of QoL outcomes among patients undergoing cancer surgery. **Methods:** Adults undergoing cancer surgery from 2017 to 2019 at a single-institution were eligible for inclusion. After scheduling surgery, patients received a remote electronic survey assessing postoperative recovery expectations. Patients reported expected QoL at 1 week and 1, 3, and 6 months postoperatively based on the eight Short-Form 36 (SF36 version 1) domains (physical functioning, role physical, bodily pain, general health, vitality, social functioning, role emotional, and mental health) reflecting a total of 32 Likert-type items (score range 0 [lowest QoL] to 4/5 [highest QoL]). Expected recovery was defined as the first time period at which population-level mean domain scores and 95% confidence intervals were above the 2nd best QoL response. Comparisons in expected QoL between 1 week and 6 months postoperatively were determined in each domain. **Results:** Among the 99 patients who received the survey, 73 provided complete or partial responses (74% completion rate). Respondents' mean age was 53 years (standard deviation [SD] 14), predominantly female (66%), with a mean American Society of Anesthesiology score of 2.0 (SD 0.6). Similar number of patients were scheduled to undergo cancer surgery for breast (34%), abdominal (30%), and skin/soft-tissue (36%) tumors. Patients expected consistent improvement in QoL throughout recovery, with significant changes from 1 week to 6 months in all domains (e.g., vitality: difference 1.69 95%CI [1.46–1.93],  $p < 0.0001$ , Table 1). Patients expected to achieve recovery in mental health, role emotional, social functioning, bodily pain physical functioning, and role physical by 3 months, and general health and vitality by 6 months (Fig. 1). **Conclusion:** Prior to treatment, patients facing cancer surgery reported differential expectations for QoL outcomes and timing of recovery by health domain. Future studies are needed to evaluate the association of patient and provider expectations for recovery, as well as how preoperative expectations align with measured postoperative QoL and behavioral outcomes.

**Table 1:** Difference in 1 week and 6-month expected QoL outcomes among patients undergoing cancer surgery. Patients expected significantly improved QoL outcomes in all health domains ( $p < 0.0001$ ). CI: confidence interval.

Health Domain	Scale Range	n	Difference in Mean Domain Score	Lower 95% CI	Upper 95% CI
General Health	0-5	67	1.45	1.17	1.72
Physical Functioning	0-4	65	1.88	1.58	2.18
Role Physical	0-4	66	2.14	1.87	2.41
Bodily Pain	0-5	72	2.11	1.85	2.37
Vitality	0-4	72	1.69	1.46	1.93
Social Functioning	0-4	71	1.89	1.59	2.18
Mental Health	0-4	70	1.06	0.80	1.32
Role Emotional	0-4	69	1.42	1.11	1.73



**Figure 1:** Expected QoL Outcomes after Cancer Surgery. Trend in expected population-level mean QoL scores with 95% confidence intervals at 1 week and 1, 3, and 6 months postoperatively are shown. Expected recovery was defined as the first time period at which population-level mean scores and 95% confidence intervals were above the 2<sup>nd</sup> best QoL response for the individual domain.

**(B205.7) Applying multidimensional computerized adaptive testing to the MSQOL-54**

Andrea Giordano, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy; Silvia Testa, Associate Professor, Department of Human and Social Sciences, University of Aosta Valley, Aosta, Italy; Marta Bassi, Associate Professor, Department of Biomedical and Clinical Sciences L. Sacco, Università di Milano, Milano, Italy; Sabina Cilia, UOC Servizio di Psicologia ASP di Catania, Catania, Italy; Antonio Bertolotto, MD, Neurology Unit & Regional Referral Multiple Sclerosis Centre (CRESM), University Hospital San Luigi Gonzaga, Orbassano, Italy; Maria Esmeralda Quartuccio, Department of Neuroscience, San Camillo-Forlanini Hospital, Orbassano, Italy; Erika Pietrolongo, Department of Neurosciences, Imaging and Clinical Sciences, University G. d’Annunzio, Chieti, Italy; Monica Falautano, Servizio di Psicologia e Neuropsicologia, UO di Neurologia e Riabilitazione Specialistica, San Raffaele Hospital, Milano, Italy; Monica Grobberio, Laboratory of Clinical Neuropsychology, Psychology Unit, ASST Lariana, Como, Italy;

Claudia Nicolai, IRCCS Don Gnocchi Foundation, Florence, Italy; Beatrice Allegri, IRCCS Don Gnocchi Foundation, Florence, Italy; Beatrice Allegri, Multiple Sclerosis Center, Neurology Unit, Hospital of Vaio, Fidenza, Italy; Rosa Gemma Viterbo, Department of Basic Medical Sciences, Neurosciences and Sense Organs, University of Bari, Bari, Italy; Paolo Confalonieri, Multiple Sclerosis Center, Unit of Neuroimmunology and Neuromuscular Diseases, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy; Ambra Mara Giovannetti, Unit of Neuroepidemiology, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy; Eleonora Cocco, Department of Medical Science and Public Health, University of Cagliari, Cagliari, Italy; Maria Grazia Grasso, Multiple Sclerosis Unit, IRCCS S. Lucia Foundation, Rome, Italy; Alessandra Lugaresi, Dipartimento di Scienze Biomediche e Neuromotorie, Università di Bologna, Bologna, Italy; Elisa Ferriani, UOC Psicologia Ospedaliera, AUSL di Bologna, Bologna, Italy; Ugo Nocentini, Department of Clinical Sciences and Translational Medicine, University of Rome “Tor Vergata,” Rome, Italy; Mauro Zaffaroni, Multiple Sclerosis Centre, ASST Valle Olona, Gallarate, Italy; Alysha De Livera, Neuroepidemiology Unit, Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Italy; Alysha De Livera, Neuroepidemiology Unit, Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Italy; George Jelinek, Neuroepidemiology Unit, Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Italy; Alessandra Solari, Unit of Neuroepidemiology, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy; Rosalba Rosato, Department of Psychology, University of Turin, Turin, Italy

**Aims:** The Multiple Sclerosis Quality of Life-54 (MSQOL-54) is one of the most used MS-specific health-related quality of life inventories. Availability of an adaptive short version that immediately processes and scores the items may improve instrument usability and validity. Just like the SF-36, two composite scores (Physical Health Composite and Mental Health Composite) are derived by combining scores of the relevant subscales. However, multidimensional computerized adaptive testing (MCAT) has not previously applied to MSQOL-54 items. The aim was to develop an MCAT version of the MSQOL-54. **Methods:** Responses from a large international sample of MS patients were assessed. First, multidimensional item response theory (MIRT) analysis was conducted, using a bifactor model. Second, MCAT simulations were implemented with different estimators, item selection methods, and measurement precision criteria. CAT latent trait estimates were evaluated in terms of bias, root mean square deviation (RMSD), and correlation. **Results:** Our dataset included 3669 MS patients (mean age 43 years [range 18-87], 74% women, 54% with a mild level of disability). The bifactor model outperformed the unidimensional model in all the statistics used for comparison. The explained variance was 74% and 43%, respectively. All items loaded satisfactorily on the general factor (range 0.60–0.84). Loadings on the specific factors were all > 0.50, except for six items (range 0.28–0.46). The MCAT MSQOL-54 was almost 70% shorter (average number of items: 15) than the fixed-length MSQOL-54 and had satisfactory accuracy (correlations > 0.90, bias < 0.03, and RMSD < 0.41). The reliability of the general factor estimated by CAT administration was 0.8, by setting the standard error of measurement to 0.4. **Conclusion:** The bifactor model is a useful approach for modeling the second order structure of the MSQOL-54, because it allows evaluation of the contribution of the general factor and the extent to which items load onto their specific (group) factors, when their relationship with the main factor is accounted for. The CAT administration proved to be parsimonious, saving more than 2/3 of

items. In conclusion, the MCAT MSQOL-54 can be used in MS clinical practice and research.

**(B205.8) Age-based differences in cardiopulmonary symptoms among patients with atrial fibrillation**

Brian Zenger, University of Utah, Salt Lake City, Utah, United States; Jeffrey L. Turner, DO, University of Utah, SLC, Utah, United States; T. Jared Bunch, MD, University of Utah, SLC, Utah, United States; Rachel Hess, MD, University of Utah, Slc, Utah, United States; Benjamin A. Steinberg, MD, University of Utah, SLC, Utah, United States

**Aims:** Age-based differences in patient-reported outcomes with atrial fibrillation (AF) and other cardiac arrhythmias have been well documented. However, the differences in symptoms have not been corroborated with real-time ambulatory rhythm monitoring. We sought to characterize the age differences in symptom-rhythm correlation among patients with AF. **Methods:** Clinically ordered ambulatory rhythm monitoring studies among patients with a history of AF were analyzed (duration 7–30 days). Patients were instructed to trigger and document symptoms (including shortness of breath [SOB], chest pain, dizziness, palpitations, or tiredness). Heart rhythm was simultaneously recorded and annotated. Patient age was dichotomized at 65 years, based on sample distribution. **Results:** 236 patients under 65 and 438 patients over 65 underwent ambulatory monitoring. Patients reported a total of 2885 symptomatic events ( $n = 1741$  age < 65, 60.3%). There were baseline differences between older and younger patients with respect to hypertension (85.4% vs. 62.7%  $p < 0.001$ ), hypothyroid (28.6 vs. 16.5,  $p = 0.001$ ), renal disease (23.6% vs. 16.5%,  $p = 0.042$ ), and stroke (38.0% vs. 22.5%,  $p < 0.001$ ), but not for myocardial infarction (33.4% vs. 26.7%), heart failure (35.7% vs. 28.4%,  $p = 0.066$ ), pulmonary disease (37.1% vs. 31.4%,  $p = 0.162$ ), or depression (34.3% vs. 36.9%,  $p = 0.566$ ). Younger patients' symptoms correlated with a documented arrhythmic event in 48% of tracings compared with 71% in older patients (Figure,  $p < 0.001$ ). Younger patients were more likely to note chest pain (14.5%) and less likely to report SOB (21.3% vs. 15.6%,  $p < 0.001$ ) and tiredness (21.4% vs. 10.7%,  $p < 0.001$ ) compared with older patients. There were no differences in reported symptoms of palpitations and dizziness between younger and older patients. **Conclusion:** Younger patients' symptomatic events were less likely correlated with cardiac arrhythmia compared with older patients. Younger patients reported more clinical symptoms of chest pain and less likely to report SOB and tiredness compared to older patients. These data indicate that reported symptoms and identified cardiac arrhythmias may have significant differences between age groups and should be considered in clinical practice.

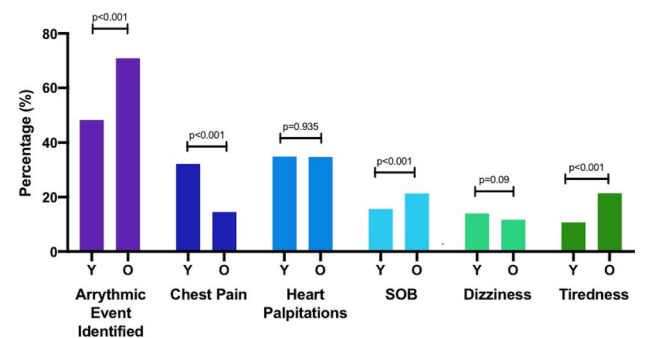
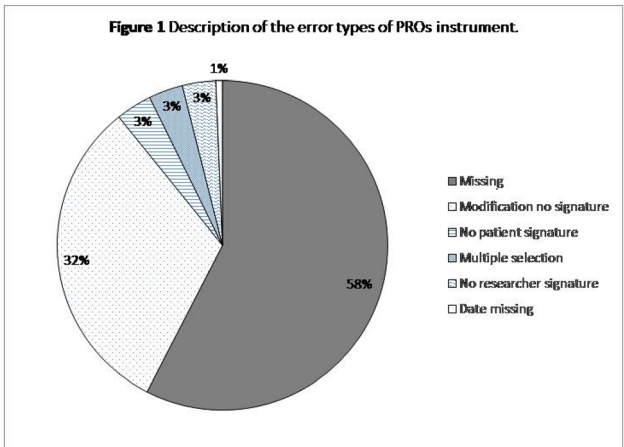


Table 1 Count of cases and error records with ePRO and paper method in PROs instrument for each error type, as well as the corresponding percentage for incompleteness.

Error type	Instrument	Incomplete of cases (n(%))		Errors (Count)	
		ePRO (n=84)	Paper (n=470)	ePRO (n=84)	Paper (n=470)
Item multiple selection	MDASI_LC	0	41(8.72)	0	48
	QOL	0	3(0.64)	0	4
	Total	0	43(9.15)	0	52
Item missing	MDASI_LC	21(25.00)	242(51.49)	85	733
	QOL	1(1.19)	55(11.70)	2	76
	Total	21(25.00)	262(55.74)	87	809
Item modification no signature	MDASI_LC	0	147(31.28)	0	443
	QOL	0	35(7.45)	0	50
	Total	0	164(34.89)	0	493
No patient signature	MDASI_LC	2(2.38)	9(1.91)	5	23
	QOL	2(2.38)	9(1.91)	5	22
	Total	2(2.38)	12(2.55)	10	45
No researcher signature	MDASI_LC	0	13(2.77)	0	30
	QOL	0	10(2.13)	0	21
	Total	0	13(2.77)	0	51
Form date missing	MDASI_LC	0	4(0.85)	0	6
	QOL	0	2(0.43)	0	4
	Total	0	5(1.06)	0	10



**(B205.10) Are important health dimensions similar across cultures? A qualitative study of four Asian populations**

Rachel Lee-Yin Tan, National University of Singapore, Singapore, Singapore; Ataru Igarashi, Yokohama City University School of Medicine, Yokohama, Japan; Zhihao Yang, Department of Health Services Management, Guizhou Medical University, Guizhou, China; Min-Woo Jo, University of Ulsan College of Medicine, Seoul, South Korea; Nan Luo, Saw Swee Hock School of Public Health, National University of Singapore, Seoul, Singapore

**Aims:** Though widely used in Asia to assess HRQoL, evidence for content validity of EQ-5D is scarce. Moreover, there is a lack of studies investigating the concept of health among general population in Asia. This study aimed to i) identify health dimensions the general public from China, Japan, South Korea, and Singapore use to conceptualize health, and ii) investigate the content validity of the EQ-

5D-5L in the four countries. Methods: Members of the general public were recruited from all four countries with quotas for age, disease experience, etc. One to one, semi-structured interviews were conducted in the participants' preferred language. In the interviews, open-ended questions (e.g., What is good/poor health to you?) were first used to elicit relevant health concepts. Then, participants were asked to complete EQ-5D before being asked to discuss its' adequacy. All interviews were transcribed verbatim and analyzed either in the original language or after being translated into English. Framework analysis was done separately for each country before pooling for comparison. Multiple analysts were used and codebooks were developed for each country to ensure consistency in coding. Discrepancies were resolved via consensus meetings. Results: 164 participants (60 Singaporeans, 60 Chinese, 24 Japanese, 20 Korean) were interviewed. A total of 81 health concepts (18 themes and 63 subthemes) were identified and organized into a framework covering the 5 domains of Physical Health, Mental Well-being, Social Relationship, Medical Condition & Treatment, and Health Promotion Knowledge & Behavior. Content pertaining to each of the five domains was spontaneously generated in all four countries. Frequently recurring themes in all four countries were closely related to the EQ-5D dimensions—('Mobility,' 'Self-Care,' 'Usual Activities,' 'Pain/Discomfort,' and 'Anxiety/Depression'). Though EQ-5D was considered an adequate questionnaire to measure health by participants, participants commonly suggested to include themes such as 'appetite' and themes under the 'Social Relationship' domain to improve the questionnaire. Conclusion: The results suggest that EQ-5D shows satisfactory content validity in these 4 countries, though additional dimensions or 'bolt-on's could potentially improve its cultural relevance. Further research in this direction is warranted.

### Outstanding Poster Award Finalists

#### (1001) Patient-centricity in action: using patient feedback to adapt the PRO-CTCAE instrument in a real-world study

Samuel Llewellyn, Vitaccess Ltd, Oxford, United Kingdom; Katie Hall, MPH, Vitaccess Ltd, Oxford, United Kingdom; Lucinda Camidge, MPH, Vitaccess Ltd, Oxford, United Kingdom; Catherine Bottomley, PhD, Vitaccess Ltd, Oxford, United Kingdom; Annabel Nixon, PhD, Chilli Consultancy, Salisbury, United Kingdom; Mark Larkin, PhD, Vitaccess Ltd, Salisbury, United Kingdom

Aims: The Melanoma UK study, conducted in collaboration with Melanoma UK, explores the real-world patient-reported impact of melanoma by collecting data via a bespoke 'bring your own device' mobile app. A previous version of the app contained 11 items from the PRO-CTCAE instrument item bank, selected by oncologists and results of a literature review. The objective of this project was to explore the relevance of the 78 AEs in the PRO-CTCAE instrument item bank with patients in relation to their melanoma treatment, to ensure that the AEs of greatest relevance to patients were included in the app. Methods: Melanoma UK study participants with any type or stage of melanoma were invited to participate in an online survey to rate whether each AE in the PRO-CTCAE item bank was 'very relevant,' 'a little bit relevant,' or 'not relevant.' Focus groups were then conducted with a subgroup of study participants to explore and understand in more depth the patient experience of the AEs reported as being most relevant. Results: The most relevant AEs identified by online survey respondents ( $n = 92$ ; 82% female) were (in order of decreasing relevance) fatigue, anxiety, aching joints, skin dryness, feeling sad, rash, itchy skin, swelling, insomnia, feeling discouraged and aching muscles. Eight of these were not included in the app already. Focus group participants ( $n = 9$ , 89% female) ratified these results, with a majority (63–100%) confirming that they had

experienced each of the 11 AEs, attributing the majority (e.g., skin dryness, rash, insomnia) to specific treatments and others (e.g., fatigue, anxiety) as both AEs and general impacts of living with melanoma. The eight additional identified AEs were compared against the summary of product characteristics for the reported melanoma treatments in the Melanoma UK study so far to ensure relevance. Conclusion: The majority of AEs identified as most relevant by melanoma patients had not previously been identified through literature searches and consultations with oncologists. This demonstrates the importance of a patient-centered approach in the design of patient-reported studies, ensuring that the most relevant data is captured. The Melanoma UK app will be updated to reflect the most relevant AEs as reported by patients.

#### (1002) Psychosocial illness impacts of Taiwanese cancer survivors

Grace Yao, PhD, National Taiwan University, Taipei City, Taiwan; Nai-Ying Ko, PhD, National Cheng Kung University, Tainan City, Taiwan; Jung-Der Wang, Sc.D., National Cheng Kung University, Tainan City, Taiwan; Hoi-San Lou, MS, National Taiwan University, Taipei City, Taiwan; Yi-Lin Wu, MS, National Cheng Kung University, Tainan City, Taiwan; Chia-Ling Tsai, BS, National Cheng Kung University, Tainan City, Taiwan; Nai-Jung Chiang, PhD, National Health Research Institutes, Tainan City, Taiwan

Aims: Past research has focused on cancer's negative consequences. Recent research has paid increasing attention to the more positive side of psychosocial adjustment to illness. The Patient-Reported Outcomes Measurement Information System (PROMIS) Illness Impact Scale, developed by the researchers at Northwestern University, contains 46 positive and 40 negative items that are classified into four subdomains: Self-Concept (SC), Social-Impact (SI), Stress-Response (SR), and Spirituality (SP). The purpose of this study was to investigate positive and negative illness impacts on Taiwanese cancer survivors. Methods: Cancer survivors ( $n = 251$ ; age:  $54.79 \pm 10.0$ ) completed the PROMIS Illness Impact Scale. We calculated change scores (i.e., impact) as the discrepancy between the participants' ratings of recalled experiences before cancer diagnosis and their ratings of post-cancer experiences. Descriptive statistics were calculated on change scores. Effect sizes (ES) provided standardized change scores. Results: Of the 86 items, more negative than positive impacts were found (59% vs. 41%). Specifically, more positive than negative impacts were noted in SP and SI subdomains, but more negative than positive impacts were noted in SR and SC subdomains. Participants showed more negative values on positive SC items but less on negative SC items. They felt stressed and upset, have more worry and fear, and felt they were a burden to family and others. However, they have no problem asking for help, to express their emotions, aware of people's love and support, and to show their appreciation. Conclusion: Taiwanese cancer survivors reported both positive and negative impacts after diagnosis. However, more negative than positive impacts were found. The psychosocial illness impacts can be classified into four categories: social relationship, appreciation of life, self-perception, and personal strength. The perceived illness impact of cancer survivors may have implications for healthcare services.

#### (1003) Rewarding participants through charitable donations to Melanoma UK motivates recruitment and engagement: a bring your own device (BYOD) study example

Katie Hall, Vitaccess Ltd, Oxford, United Kingdom; Chengyu Ouyang, Vitaccess Ltd, Oxford, United Kingdom; Anna Chilongo, Vitaccess Ltd, Oxford, United Kingdom; Samuel Llewellyn, Vitaccess Ltd, Oxford, United Kingdom; Casey Quinn, Vitaccess Ltd,

Oxford, United Kingdom; Mark Larkin, Vitaccess Ltd, Oxford, United Kingdom

**Aims:** The Melanoma UK study, developed by Vitaccess in collaboration with the patient advocacy organization Melanoma UK and the Royal Marsden NHS Foundation Trust, explored the real-world impact of melanoma via a bespoke ‘bring your own device’ mobile app. Through the app, participants were asked questions about symptom burden and quality of life via validated patient-reported outcome measures (PROMs). In May 2019, Vitaccess launched a monthly scheme pairing a charitable reward with the completion of PROMs to improve recruitment and engagement. **Methods:** For each of the first three PROMs participants completed each month (EQ-5D-5L, PRO-CTCAE or EORTC QLQ-C30), Vitaccess donated £1 to Melanoma UK, i.e., a maximum of £3 per-person per-month. Participants were notified each month of the total raised the previous month via an in-app notification. **Results:** Since the study launch, the total number of app users increased from 259 to 599 (March 2018–February 2020). After incentive scheme implementation, the mean monthly new users significantly increased from a mean (standard deviation) of 13 (6.5) to 19 (5.8) ( $p < 0.05$ ). Similarly, a chow test confirmed the total users increased at a significantly faster rate after May 2019 ( $p < 0.05$ ). Following incentive introduction, the monthly PROM completions per user significantly increased from a mean (standard deviation) of 2.79 (0.35) to 3.57 (0.51) ( $p < 0.05$ ). The mean monthly completion rate of the EQ-5D-5L and PRO-CTCAE also significantly increased after May 2019 from 0.046 completions (where 0.033 denotes completion of the PROM at least once per month) to 0.057, and 0.044 to 0.048, respectively ( $p < 0.05$ ). Contrastingly, the mean percentage of users completing at least one PROM monthly did not change, nor did mean completion rate of the EORTC QLQ-C30. Due to the incentive’s framework, £1,342 has been donated to Melanoma UK to date. **Conclusion:** A key aspect of the value of app-based digital real-world evidence studies is the ability to generate longitudinal data, which requires patient engagement over time. Implementation of a charitable reward scheme in the Melanoma UK study improved recruitment and user engagement in completing specific PROMs. This is in line with literature indicating higher motivation by small charitable donations for completing actions compared to direct monetary rewards.

#### **(1004) Guidance for the use of PROMIS Pediatric measures in ambulatory clinics**

Elizabeth Cox, MD PhD, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, United States; Sarah Dobrozi, MD MS, Medical College of Wisconsin, Milwaukee, Wisconsin, United States; Christopher Forrest, MD PhD, Children’s Hospital of Philadelphia, Philadelphia, Pennsylvania, United States; Wendy E. Gerhardt, MSN RN-BC, Cincinnati Children’s Hospital, Cincinnati, Ohio, United States; Harald Kliems, MS, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, United States; Bryce B. Reeve, PhD, Duke University School of Medicine, Madison, North Carolina, United States; Nan Rothrock, PhD, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Jin-Shei Lai, PhD OTR, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Jacob Svenson, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, United States; Lindsay Thompson, MD MS, University of Florida College of Medicine, Gainesville, Florida, United States; Thuy Dan Tran, MD MS, University of Florida College of Medicine, Gainesville, Florida, United States; Thuy Dan Tran, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, United States;

Carole A. Tucker, PhD, Temple University College of Public Health, Philadelphia, Pennsylvania, United States

**Aims:** Patient-Reported Outcomes Measurement Information System (PROMIS) Pediatric measures quantify health from the patient perspective, which is crucial to patient-centered care. To realize the benefits of PROMIS measures in pediatric settings, healthcare system and clinician leaders must attend to unique challenges to how the measures are implemented and used. **Methods:** To identify and address challenges to PROMIS use in the US ambulatory pediatric setting, 18 semi-structured telephone interviews of health system leaders, measurement implementers, and ambulatory pediatric clinicians were conducted. Five coders used thematic content analysis to iteratively identify and refine themes and subthemes in the interview data. These identified themes became the topics to which content experts responded, providing guidance and recommendations. **Results:** Analysis of the interviews yielded six themes: (1) selection of PROMIS measures, (2) method of administration, (3) use of PROMIS Parent-Proxy measures, (4) privacy and confidentiality of PROMIS responses, (5) interpretation of PROMIS scores, and (6) using PROMIS scores clinically. A total of 29 recommendations were made. For example, experts recommended engaging a wide array of stakeholders to choose measures that are meaningful, actionable, and best assessed by patient-report; having the child respond themselves whenever possible; and considering options for protecting children’s privacy early in the process. Training for clinicians was recommended to support score interpretation and timely discussion of both normal and concerning scores with patients. **Conclusion:** Based on the challenges encountered by pediatric clinicians and health system leaders, this work provides guidance for the integration of PROMIS measures in pediatric clinics. In some instances, data on which to make recommendations were lacking, highlighting opportunities for future research.

#### **(1005) Resilience among depressed individuals: cognitive appraisal processes as mediators**

Carolyn Schwartz, Sc.D., DeltaQuest Foundation, Inc., Concord, Massachusetts, United States; Roland Stark, M.Ed., DeltaQuest Foundation, Inc., Concord, Massachusetts, United States; Brian Stucky, Ph.D., DeltaQuest Foundation, Inc., Los Angeles, California, United States; Wesley Michael, M.B.A., Rare Patient Voice, LLC, Towson, Maryland, United States; Bruce Rapkin, Ph.D., Albert Einstein College of Medicine, Bronx, New York, United States

**Aims:** Depression can severely impair one’s ability to carry out major life activities. Cognitive-behavioral therapy is an effective non-pharmacologic intervention by which individuals are taught to identify and modify disempowering self-talk. There may, however, be other types of self-talk that can mitigate depression’s impact. Research has revealed that cognitive-appraisal processes (i.e., the way patients think about health and quality of life (QOL)) influence treatment burden, adherence, and treatment outcomes. We investigated how depressive symptoms influenced subsequent resilience, and whether cognitive-appraisal processes mediated this relationship differently in people with or without depression. **Methods:** This secondary analysis examined longitudinal data from 1481 chronically ill patients and caregivers. A depression-symptom scale was created with items from PROMIS-10; NeuroQOL Applied Cognition, Positive Affect & Well-Being; and Ryff Environmental Mastery. Resilience was operationalized as better-than-expected functioning despite physical or mental-health symptoms using the Centers for Disease Control Healthy Days Module. Appraisal was measured using the QOL Appraisal Profile v2. Receiver Operating Characteristic analysis identified a cut-point for depression as a comorbidity (at both

timepoints vs. at fewer; Concordance = 0.77). A multi-group, time-lagged path analysis was implemented for depressed and non-depressed individuals testing the hypothesis that the relationship between baseline depressive symptoms and resilience at follow-up was mediated by baseline appraisal, after demographic-covariate adjustment, and to evaluate the extent to which the mediation effect varied between depressed and non-depressed individuals. Results: The total relationship between depressive symptoms and resilience was stronger for depressed than non-depressed individuals (coef. = -0.38 vs. -0.20). Appraisal mediated the relationship, but only for depressed individuals (total indirect effect = -0.10 vs 0.01). Specifically, depressed individuals who at baseline focused more on wellness and less on health worries, evinced higher levels of resilience. The path model explained 27% vs. 15% of the variance for the depressed and non-depressed groups, respectively. Conclusion: There was a stronger relationship between baseline depressive symptoms and subsequent resilience for depressed than non-depressed individuals. When depressed individuals emphasized certain positive aspects of their experience, they were able to lessen the impact of this illness on their daily function. Cognitive-behavioral interventions might expand the target of the self-talk to embrace such health-specific appraisal processes.

#### (1006) Preliminary results of perioperative patient-reported outcomes in patients with lung cancer: a multicenter observational cohort study

Wei Dai, Department of Thoracic Surgery, Sichuan Cancer Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, China; Xing Wei, Department of Thoracic Surgery, Sichuan Cancer Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, China; Yaqin Wang, Department of Thoracic Surgery, Sichuan Cancer Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, China; Jia Liao, Department of Thoracic Surgery, Sichuan Cancer Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, China; Fang Liu, Traditional Chinese Medicine Department of Oncology, Sichuan Cancer Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, China; Yaqian Feng, Department of Thoracic Surgery, Sichuan Cancer Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, China; Wei Xu, School of Public Health and Management, Chongqing Medical University, Chongqing, China; Hongfan Yu, School of Public Health and Management, Chongqing Medical University, Chongqing, China; Qingsong Yu, School of Public Health and Management, Chongqing Medical University, Chongqing, China; Yang Pu, School of Public Health and Management, Chongqing Medical University, Chongqing, China; Xiaojun Yang, School of Public Health and Management, Chongqing Medical University, Chongqing, China; Xiaojun Yang, Department of Thoracic Surgery, Sichuan Cancer Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, China; Qiang Li, Department of Thoracic Surgery, Sichuan Cancer Hospital, School of Medicine, University of Electronic Science and Technology of China, Chengdu, China; Qiuling Shi, School of Public Health and Management, Chongqing Medical University, Chongqing, China

**Aims:** Lung cancer surgery can lead to severe perioperative symptom burden for patients. We aimed to explore the feasibility of longitudinal patient-reported outcomes data collection and profile the trajectories of symptoms in perioperative patients with lung cancer. **Methods:** Patients with initial diagnosis of lung cancer and planned surgery were recruited from 6 hospitals. The MD Anderson Symptom Inventory-lung cancer module (MDASI-LC) and single-item quality of life (QOL) scale (0–10) were administered before surgery, daily postoperatively, and weekly after discharge up to 4 weeks or the start of postoperative cancer treatment. MDASI-LC was administered via paper-and-pencil or WeChat-based electronic questionnaire. **Results:** Among 512 patients enrolled, 424 patients (82.8%) were lung cancer from November 2017 to January 2020. The MDASI-LC data completion rates were 100% before surgery, 94.2% during hospitalization after surgery, and 65.9% after discharge, respectively. Most of the patients were males (53.1%), received video-assisted thoracoscopic surgery (81.4%), underwent lobectomy (63.4%), and had an early stage tumor (63.2%). The median postoperative length of stay was 7 day. The 5 most common symptoms were pain, fatigue, shortness of breath, coughing, and disturbed sleep. Patients had the most severe symptom burden and worst QOL in the first 3 days after surgery. **Conclusion:** This real-world cohort study showed that longitudinal perioperative patient-reported outcomes data collection is feasible in patients with lung cancer. The results preliminarily provided perioperative symptom trajectories for patients undergoing lung cancer surgery. Further analysis of patient-reported symptoms and QOL will be made to find more meaningful comparisons of different surgical technique and perioperative care.

**Table 1.** Selected demographic and clinical characteristics of 424 lung cancer patients

Variables	Mean (SD)	Median (min-max)
Age	56.1 (10.4)	56 (26-80)
BMI	22.9 (2.8)	22.8 (16.6-32.4)
Postoperative LOS	7.7 (4.4)	7 (3-41)
LOS	14.2 (6.3)	13 (6-49)
Variables	Number	Percent
Male	225	53.1
VATS	345	81.4
Lobectomy	269	63.4
Adenocarcinoma	344	81.1
Early stage (Stage 0-I)	268	63.2
≥ Grade 2 perioperative complication	91	21.5

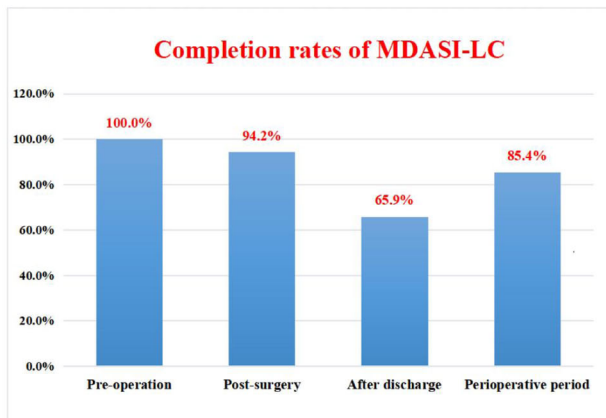
BMI, body mass index; LOS, length of stay; SD, standard deviation; VATS, video-assisted thoracoscopic surgery

**Table 2.** The median scores of top 5 symptoms and quality of life from pre-operation to 4 weeks after discharge

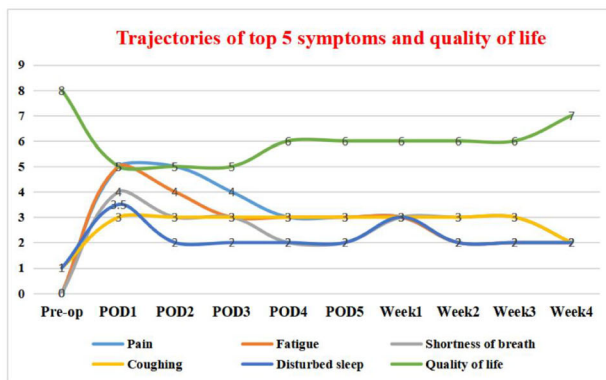
Items	Pre-op	POD1	POD2	POD3	POD4	POD5	Week1	Week2	Week3	Week4
Pain	0	5	5	4	3	3	3	2	2	2
Fatigue	0	5	4	3	3	3	3	2	2	2
Shortness of breath	0	4	3	3	2	2	3	3	3	2
Coughing	1	3	3	3	3	3	3	3	3	2
Disturbed sleep	1	3.5	2	2	2	2	3	2	2	2
Quality of life	8	5	5	5	6	6	6	6	6	7

POD, postoperative day; Pre-op, pre-operation

**Figure 1.** Completion rates of MDASI-LC data



**Figure 2.** Trajectories of top 5 symptoms and quality of life from pre-operation to 4 weeks after discharge



**New Investigator & Student Poster Award Finalists**

**New Investigator Poster Award Finalists**

**(2001) Association between untreated Dental Caries, low BMI and Oral Health-related Quality of life among children in Lagos, Nigeria**

Afolabi Oyapero, Lagos State University College of Medicine, Ikeja, Nigeria; Aderinsola Adenaike, BDS, FWACS., Lagos State University College of Medicine, Lagos, Nigeria

**Aims:** Dental caries is the most prevalent childhood disease, and a significant public health problem especially in socially disadvantaged populations. We explored the associations between Untreated Dental Caries, anthropometric measurements of children and OHRQOL at a Tertiary Paediatric Dental Clinic in Lagos, Nigeria. **Methods:** 278 child/parent on first visit appointment were randomly recruited from the pediatric dental clinic and were screened with a validated, interviewer administered questionnaire to obtain socio-demographic data, anthropometric measurements, OHRQOL and dental examinations. Caries was scored according to WHO criteria (DMFT) while untreated dental caries were measured by the Pulpal exposure, Ulceration, Fistula and Abscess (PUFA/pufa) index. The weight for age (WAZ), height for age (HAZ), and weight for height (WHZ) parameters were utilized to evaluate each child’s nutritional status. OHRQoL was measured with the Early Childhood Oral Health Impact Scale (ECOHS). Chi square test, ANOVA and Binary regression analysis were used to evaluate the effect of s-ECC on anthropometric measurements and OHRQoL. Nominal significance was set at  $p < 0.05$ . **Results:** 220 (79.1%) children had past or current caries experience (DMFT + dmft  $\geq 1$ ) of which 194 (70.0%) were previously untreated. 172 (61.8%) children had odontogenic infections (PUFA + pufa  $\geq 1$ ) due to untreated carious teeth [mean:  $2.42 \pm 1.95$ ]. 74 (26.6%) children were stunted; 12 (4.3%) were underweight while 30 (10.8%) were wasted. Pulpal exposure and odontogenic infection (PUFA + pufa  $\geq 1$ ) was associated with increased risk of wasting (OR 2.45; 95% CI 1.16–4.88). Children with a PUFA + pufa score  $\geq 1$  had higher mean ECOHS scores ( $17.21 \pm 5.86$ ) compared to children with pufa score = 0 ( $8.32 \pm 3.85$ )  $p < 0.001$ . The domains with the highest impact scores were pain, difficulty in chewing, disturbed sleep and absenteeism from school. Pain and difficulty in chewing were associated with lower body mass index-for-age z-scores ( $p < 0.01$ ). **Conclusion:** Untreated dental caries resulting in odontogenic infections was significantly associated with low weight for height as well as negative impact on the OHRQOL. Decision and policy-makers should be made aware of the additional burden that oral neglect can have on the anthropometric indices and QOL of children.

**(2002) Methodological challenges of using dissimilar multi-attribute utility instruments for people with MS: a direct comparison of commonly used instruments**

Julie A. Campbell, PhD, University of Tasmania, Menzies Institute for Medical Research, Hobart, Australia; Bruce V. Taylor, MBBS, PhD, University of Tasmania, Menzies Institute for Medical Research, Hobart, Australia; Ingrid van der Mei, PhD, University of Tasmania, Menzies Institute for Medical Research, Hobart, Australia; Suzi Claffin, PhD, University of Tasmania, Menzies Institute for Medical Research, Hobart, Australia; Mohammad Akhtar Hussain, PhD, University of Tasmania, Menzies Institute for Medical Research, Hobart, Australia; Andrew J. Palmer, MBBS, BMedSci, University of Tasmania, Menzies Institute for Medical Research, Hobart, Australia

**Aims:** Background/Aims: In Australia in 2017, there were almost 26,000 people with multiple sclerosis (MS) at a cost of \$1.75 billion AUD. MS is a chronic neurological disease where MS-onset generally

occurs between the ages of 20–40 years when people are establishing careers and families. Patient-centeredness, precision-medicine and patient-reported outcome measures (PROMs) are integral to the clinical assessment of people with MS (PwMS). Generic multi-attribute utility instruments (MAUI) are PROMs that assess health state utilities (HSU) as a measure of health-related quality of life for cost-utility (cost-effectiveness/economic evaluation) and clinical assessment. Different MAUIs yield dissimilar HSUs and cost-effectiveness calculations. This has important policy implications for MS management/resourcing decisions; if a new treatment or intervention fails to demonstrate cost-effectiveness, it may severely limit access to new and worthy treatments. We aim to identify the MAUI that preferentially assesses the complex physical/psychosocial health needs of PwMS. Methods: We will conduct the world's first comprehensive head-to-head comparison of key MAUIs for PwMS. The study's sample will be sourced from the Australian MS Longitudinal Study, with over 3,000 participants representative of Australia's MS population. Study design includes the AQoL-8D, EQ-5D-5L, MSQoL-54 (SF-6D) and PROMIS-29 + 2 MAUIs. MAUIs will be anchored to a disease-specific instrument, the remaining 18 MSQoL-54 questions, and last acute MS-episode for people with relapsing–remitting MS. As an alternative anchor point, subjective quality of life will be assessed with the Personal Wellbeing Index. Covid19-related questions will be included. Direct comparisons of MAUIs will include an array of comparative methodologies including examination of individual and summary HSUs and scores. Dimensional responses will be investigated. Participant representativeness and proportions of completion will be assessed. Discriminatory sensitivity will be compared with MS disease-severity. Bland–Altman analysis will be conducted. Population norms will be included. Results: The comprehensive study will be in the field, however, preliminary results regarding initial response, participant characteristics, and HSUs/scores will be presented. Conclusion: A fundamental question in MS health economics will be addressed: which MAUI is most appropriate to capture changes associated with progression of MS and impacts of interventions? These findings will lead to more robust economic evaluations for MS-specific preventions/interventions leading to improved quality of life and resourcing decisions for PwMS.

### (2003) The unique symptom burden of receiving CAR T-cell therapy

Meagan Whisenant, PhD, APRN, The University of Texas Health Science Center at Houston, Houston, Texas, United States; Samer Srour, MB ChB, MS, The University of Texas MD Anderson Cancer Center, Department of Stem Cell Transplantation and Cellular Therapy, Houston, Texas, United States; Ishwariah Subbiah, MD, MS, The University of Texas MD Anderson Cancer Center, Department of Palliative, Rehabilitation and Integrative Medicine, Houston, Texas, United States; Donna Griffin, The University of Texas MD Anderson Cancer Center, Department of Symptom Research, Houston, Texas, United States; Loretta A. Williams, PhD, APRN, The University of Texas MD Anderson Cancer Center, Department of Symptom Research, Houston, Texas, United States; Xin Shelley Wang, PhD, MD, The University of Texas MD Anderson Cancer Center, Department of Symptom Research, Houston, Texas, United States

Aims: There has been little research on the patient experience of symptom burden from CAR T-cell therapy and no validated measure specific to the symptoms of chimeric antigen receptor T-cell therapy currently exists. The purpose of this study was to identify symptoms experienced by patients receiving CAR T-cell therapy. Methods: This qualitative study aims to determine the content domain for a patient-reported outcome (PRO) CAR T-cell therapy symptom burden

measure. Twenty-one patients who had received CAR T-cell therapy as standard care described their experience in single qualitative interviews. Content analysis was used to define the symptom burden content domain for a PRO measure. Results: Mean patient age was 61.4 years (standard deviation = 13.1); 76.2% were male; 100% were White; 85.7% were diagnosed with diffuse large B cell lymphoma. 61.9% were within 3 months of their CAR T-cell infusion; 81.0% experienced clinician-documented cytokine release syndrome and 28.6% experienced neurotoxicity. Content analysis found 34 symptoms related to both disease and treatment. The most common disease symptom identified by patients as occurring at diagnosis or disease recurrence in the qualitative interviews was pain (42.9%). The most common symptoms identified by  $\geq 20\%$  of patients as related to CAR T-cell therapy in the qualitative interviews included fatigue (tiredness) (61.9%), lack of appetite (28.6%), headache (28.6%), chills/feeling cold (23.8%), and feeling confused (23.8%). The qualitative analysis also confirmed that symptoms interfere with daily activities, work, walking, relationships with others, and mood. Conclusion: Patients who receive standard CAR T-cell therapy experience numerous symptoms related to disease and treatment, including many symptoms related to the T-cell infusion. Symptoms may result in interference with daily activities, relationships, life plans, treatment adherence, and mood. Well-designed PROs are essential for accurate symptom assessment and management to maintain patient functioning. The content domain for a PRO symptom-burden measure of receiving CAR T-cell therapy encompasses the severity and activity interference of common symptoms.

### Student Poster Award Finalists

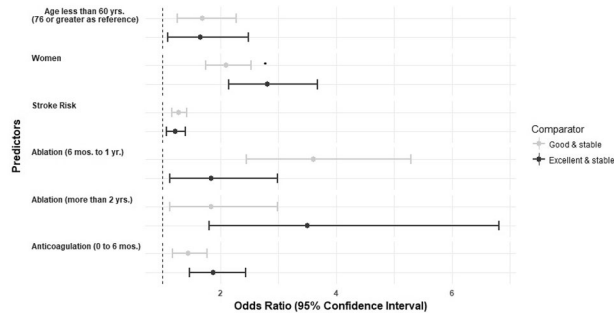
#### (2004) Health status trajectories of outpatients with atrial fibrillation

Jae-Yung Kwon, University of British Columbia, Vancouver, British Columbia, Canada; Richard Sawatzky, Trinity Western University, Langley, British Columbia, Canada; Jennifer Baumbusch, University of British Columbia, Vancouver, British Columbia, Canada; Sandra Lauck, Centre for Heart Valve Innovation, St. Paul's Hospital, Vancouver, British Columbia, Canada; Pamela A. Ratner, University of British Columbia, Vancouver, British Columbia, Canada

Aims: Health and quality of life are important treatment goals in patients with atrial fibrillation; however, patterns in patients' health trajectories over time are not well described. The purpose of this study was to identify health trajectories in outpatient setting of patients with atrial fibrillation and assess whether variation in their trajectories could be explained by differences in their age, gender, travel distance to the clinic, stroke risk, and when ablation or anticoagulation therapy was received. Methods: This was a retrospective analysis of data collected between 2008 and 2016 from a provincial cardiac registry in British Columbia (Canada) and linked with administrative health data. The Atrial Fibrillation Effect on Quality of Life Questionnaire was used to measure health status on repeated visits (up to 10 visits over 5 years). Longitudinal latent class models were used to identify latent classes of different health status trajectories while a three-step multinomial logistic regression approach was used to identify predictors. Results: The patients ( $n = 7439$ ) were primarily men (61.1%) over 60 years of age (72.9%). Three health status trajectories were identified. The "poor and improving health" class (63.6%) had relatively low baseline scores and gradually improved at each follow-up visit. The "good and stable health" (27.9%) class had higher baseline scores, which remained relatively stable over time. The "excellent and stable health" class (8.5%) had very high baseline scores with little change over time. Relative to the "good and stable health" and "excellent and stable health" classes,



people in the “poor and improving health” class were more likely to (a) be less than 60 years in age (versus 76 or older); (b) be women; (c) have higher scores for Atrial Fibrillation Stroke Risk; (d) have had ablation therapy within 6 months to 1 year or more than 2 years after the initial consultation; and (e) have had anticoagulation therapy within 6 months of the initial consultation visit (see corresponding figure). Conclusion: The change in the health trajectories in outpatient with atrial fibrillation may be modest at best. Age, gender, stroke risk score, and ablation and anticoagulation therapy at specified follow-up predicted membership in the lowest health trajectory.



### (2005) Health-related quality of life is strongly associated with self-efficacy, self-esteem, loneliness, and stress in 14 to 15-year-old adolescents

Hilde T. Mikkelsen, University of Agder, KRISTIANSAND S, Norway; Kristin Haraldstad, University of Agder, Kristiansand, Norway; Sølvi Helseth, Oslo Metropolitan University, Oslo, Norway; Siv Skarstein, Oslo Metropolitan University, Oslo, Norway; Milada C. Småstuen, Oslo Metropolitan University, Oslo, Norway; Gudrun Rohde, University of Agder, Oslo, Norway

**Aims:** To enhance and better understand health-related quality of life (HRQOL) in adolescents, it is important to study factors associated with HRQOL. The present study aimed to assess possible associations between selected sociodemographic variables, self-efficacy, self-esteem, pain, sleep, loneliness, stress and HRQOL in 14 to 15-year-old adolescents. All selected variables were theoretically known and clinically relevant variables reported in previous HRQOL research. **Methods:** A cross-sectional study was performed among 696 adolescents (14–15 years) in a school-based setting. Sociodemographic variables, self-efficacy, self-esteem, pain, sleep, loneliness, and stress were analyzed. The variables were all assessed with well-validated instruments. HRQOL was analyzed using KIDSCREEN 27. Analyses involved Chi square, independent t-tests, Mann–Whitney U tests, linear regression analyses and hierarchical regression analyses. The results from linear regression models were expressed as standardized beta. **Results:** The adolescents generally reported good HRQOL. However, girls scored significantly worse on HRQOL, self-efficacy, self-esteem, pain, sleep, loneliness, and stress compared to boys. Using hierarchical regression analyses we found that Self-efficacy (beta = 0.11–0.24), Self-esteem: (beta = 0.12–0.21), Loneliness: (beta = – 0.24 to – 0.45) and Stress: (beta = – 0.26 to – 0.34) revealed the strongest associations with the HRQOL dimensions. Sociodemographic-, pain- and sleep related covariates were all significantly associated with some of the KIDSCREEN subscales, however their effect on the outcome was smaller than for the psychosocial variables listed above. Being a girl, not living with both parents, not having both parents working, being absent from school more than 4 days, having pain and having lack of enough sleep were all independently negatively associated with HRQOL. **Conclusion:**

HRQOL is strongly associated with self-efficacy, self-esteem, loneliness, and stress in 14- to 15-year-old adolescents. Our findings indicate that positive psychosocial factors such as self-efficacy and self-esteem might play a buffer role for negative psychosocial factors (e.g., stress) in adolescents. Further, our results show that girls score significantly worse on factors that are associated to HRQOL compared to boys. Thus, in order to increase HRQOL in school-based populations of adolescents, we suggest that future interventions should aim to strengthen self-efficacy and self-esteem. We recommend gender specific interventions.

### (2006) Why invest in patient-reported outcome measures (PROMs)? An exploratory qualitative study to understand why hospital leaders established PROM programs at major healthcare organizations in the United States (US)

Danny Mou, MD, Brigham Women’s Hospital, Boston, Massachusetts, United States; Claire Sokas, MD, Center for Surgery and Public Health, Boston, Massachusetts, United States; Azan Virji, BA, Harvard Medical School, Boston, Massachusetts, United States; Marilyn Heng, MD, MPH, Massachusetts General Hospital, Boston, Massachusetts, United States; Andrea Pusic, MD, Brigham and Women’s Hospital, Boston, Massachusetts, United States; Christer Mjaset, MD, Harvard School of Public Health, Boston, Massachusetts, United States

**Aims:** There is increasing pressure to demonstrate value of medical care. PROMs assess how patients feel and function, and thus have been considered a critical component of high-value care. However, PROMs implementation requires significant resources that can pose a barrier to adoption. Despite this, several groups of hospital leaders across the US have invested significant resources to develop institution-wide PROMs programs. We interviewed these key stakeholders to understand the impetus for their investment. **Methods:** A semi-structured interview guide was developed from literature review and expert input. We conducted semi-structured interviews per snowball sampling with 23 hospital executives and PROM program directors across 4 major healthcare organizations in the United States. Mean interview time was 44 min. Interviews were recorded, transcribed, and coded. Interviews continued until thematic saturation. Data were analyzed using thematic analysis. **Results:** Preliminary results revealed that key proponents of the PROMs program often had expertise in patient-centered care, the biopsychosocial model, and value-based healthcare. Furthermore, the interviewed executives often believed that their hospitals delivered superior care relative to their competitors. They thought that PROMs may be a compelling way to showcase this. Potential financial incentives from payors were noted to encourage establishing institution-wide PROMs. However, key stakeholders claimed that beyond financial incentives, systematically collecting PROMs was the right thing to do. In fact, none of the interviewed hospital leaders received any external financial support to start their PROMs program. Furthermore, several leaders expressed concern about the uncertain return on investment (ROI) of PROMs. Finally, several key stakeholders posited that by being early adopters of PROMs, they may have the credibility to dictate future PROMS-related reimbursement terms. **Conclusion:** PROM programs vary significantly across institutions in the United States both in size and scope. Understanding why key stakeholders of major healthcare institutions invest in PROMs will be a critical to understanding the role PROMs is expected to play in US healthcare system. Importantly, these insights may serve to influence and direct the adoption of PROMs more broadly across healthcare organizations.

## Applied Clinical Research I

### (3001) Self-reported physical functioning as a predictor of unplanned hospitalization among adult patients with congenital heart disease

Misa Takegami, PhD, MPH, Department of Preventive Medicine and Epidemiology, National Cerebral and Cardiovascular Center, Suita, Japan; Hideo Ohuchi, MD, PhD, Department of Pediatric Cardiology, National Cerebral and Cardiovascular Center, Suita, Japan; Kenji Yasuda, MD, PhD, Department of Pediatrics, Shimane University Faculty of Medicine, Izumo, Japan; Yoshihiro Miyamoto, MD, PhD, Center for Cerebral and Cardiovascular Disease Information, Open Innovation Center, National Cerebral and Cardiovascular Center, Suita, Japan; Kunihiro Nishimura, MD, PhD, Department of Preventive Medicine and Epidemiology, National Cerebral and Cardiovascular Center, Suita, Japan; Kenichi Kurosaki, MD, PhD, Department of Pediatric Cardiology, National Cerebral and Cardiovascular Center, Suita, Japan; Satoshi Yasuda, MD, PhD, Department of Cardiovascular Medicine, National Cerebral and Cardiovascular Center, Suita, Japan

**Aims:** Increasing rates of survival has raised requisition of life-long care for persistent and late complications of adult patients with congenital heart disease. The clinical determinants of acquired complication remain unclear. We examined the longitudinal association between self-reported physical functioning and incidence of unplanned hospitalization among adult patients with congenital heart disease. **Methods:** The prospective cohort study included 240 adult patients with congenital heart disease (41.7% male, mean age = 29.4). We performed a questionnaire survey consisting of the 36-Item Short-Form Health Survey (SF-36). Self-reported physical functioning level was evaluated by SF-36 physical functioning subscale. The main outcome was the incidence of unplanned hospitalization. The Cox proportional hazard model was used to estimate hazard ratios (HRs) of physical function subscale categories grouped by quartile for incident unplanned hospitalization, adjusted by age, sex, anatomic complexity of congenital heart disease and the New York Heart Association (NYHA) classification of functional status. **Results:** Of the 240 patients, 49 (16.7%) had at least one unplanned hospitalization during a mean follow-up of 4.8 years. SF-36 physical functioning subscale was significantly related to anatomic complexity of congenital heart disease and the NYHA classification ( $p$  for trend,  $p < 0.001$ ,  $p < 0.001$ ). Compared with the lowest physical function group, the univariate and multivariate-adjusted hazard ratios (95% confidential interval) for unplanned hospitalization in the highest physical function group were 3.4 (2.0–5.6) and 1.8 (1.0–3.3), respectively. **Conclusion:** Self-reported physical functioning level was associated with increased risks of hospitalization for adult patients with congenital heart disease. Clinicians should carefully assess adult patients whose subjective perception of their physical functioning capacity is lower than those in similar age groups.

### (3002) Salutogenic factors of healthy eating habits in Japanese elderly people: a structural equation modeling approach

Junko Sakano, ph.D., Okayama Prefectural University, Okayama, Japan; Atsushi Ogata, Graduate School of Okayama Prefectural University, Okayama, Japan; Yoichi Sawada, Ph.D., Okayama Prefectural University, Okayama, Japan; Yuki Yajima, Ph.D., Niimi University, Okayama, Japan

**Aims:** The aim of this research was to examine whether a set of individual and social-environmental factors can predict healthy eating habits in Japanese elderly people, based on Antonovsky's salutogenic

framework. **Methods:** A questionnaire survey was conducted for 359 Japanese community-dwelling adults aged  $\geq 60$  years. The questionnaire included questions on healthy eating habits (weight control, food variety and balance, and social dining), Sense of Coherence (SOC), healthy eating literacy (HEL), healthy food environment, and cooking skills as well as respondent's socio-demographic characteristics. A structural equation modeling (SEM) approach was used to identify significant individual and socio-environmental factors of healthy eating habits. **Results:** The SEM supported a final model in which a latent variable of healthy eating habits was regressed on SOC, HEL, cooking skills, and healthy food environment as well as socio-demographic variables (Chi square = 65.29,  $df = 20$ ,  $p < 0.001$ , CFI = 0.933, RMSEA = 0.080). Significant factors that can promote healthy eating habits were SOC (std. beta = 0.19), HEL (std. beta = 0.28), Healthy food environment (std. beta = 0.13), Living with someone (std. beta = 0.13), Gender (std. beta = -0.13), and BMI (std. beta = -0.20). **Conclusion:** This study was the first to apply a salutogenic approach to the identification of individual and socio-environmental factors of health eating behaviors in Japanese elderly people. This study provided a valuable contribution to designing more effective salutogenic approaches to improve healthy eating behaviors of community-dwelling older adults.

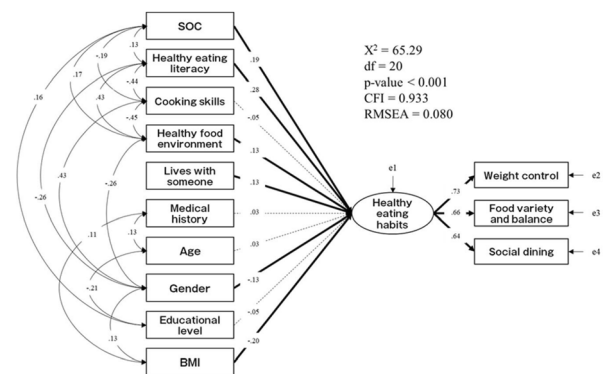


Fig. Salutogenic factors of healthy eating habits in Japanese elderly people: SEM approach

### (3003) Development of a novel item bank and multidimensional computerized adaptive test to measure patient-reported outcomes and improve management of kidney diseases: systematic review

Ben Fletcher, DPhil, Centre for Patient-Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Sarah Damery, PhD, NIHR Applied Research Collaboration (ARC) West Midlands, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Nicola Anderson, RN Dip HE, MSc, Centre for Patient-Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Olalekan (Lee) Aiyegbusi, MB BCH, PhD, Centre for Patient-Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Melanie Calvert, PhD, Centre for Patient-Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Paul Cockwell, MB BCH, PhD, University Hospitals Birmingham NHS Foundation Trust, Birmingham, United Kingdom

Kingdom; James Ferguson, MB ChB, MD, University Hospitals Birmingham NHS Foundation Trust, Birmingham, United Kingdom; Mike Horton, PhD, Psychometric Laboratory of Health Sciences, University of Leeds, Leeds, United Kingdom; Muirne Paap, PhD, Faculty of Behavioural and Social Sciences, University of Groningen, Groningen, Netherlands; Chris Sidey-Gibbons, PhD, MD Anderson Centre for INSPIRED Cancer Care, University of Texas, Houston, Texas, United States; Anita Slade, PhD, MD Anderson Centre for INSPIRED Cancer Care, University of Texas, Houston, Texas, United States; Anita Slade, PhD, Centre for Patient-Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Neil Turner, BM BCh, PhD, College of Medicine and Veterinary Medicine, University of Edinburgh, Edinburgh, United Kingdom; Derek Kyte, PhD, Centre for Patient-Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom

**Aims:** To inform the development of a new renal item bank and computerized adaptive test (CAT) for patients with chronic kidney disease (CKD), across all stages of the disease, by identifying: (i) symptoms and health-related quality of life (HRQOL) domains important in CKD, and where available, the prevalence of symptoms; (ii) existing validated patient-reported outcomes (PROs) used in CKD. **Methods:** Bibliographic databases (Ovid MEDLINE, Ovid PsycINFO and EBSCO CINAHL) were searched from January 2000 until present, for quantitative/qualitative articles focusing on symptom burden/HRQOL/PROs in adults with chronic kidney disease (stage 1–5, on dialysis, or in receipt of a renal transplant). A forward citation search and a hand search of the reference lists was conducted. Editorials, conference abstracts, case reports, systematic reviews and non-English papers were excluded. Data were extracted on: study characteristics, population, symptoms/symptom clusters, HRQOL domains and PRO characteristics/items. Two independent investigators conducted screening and data extraction, with a third on hand to mediate where required. Reported symptoms/HRQOL domains/PRO items were synthesized according to the categories presented by Lockwood et al. (2019) and van der Willik et al. (2019) and prevalence figures combined using meta-analysis where heterogeneity was acceptable. **Results:** The results of the review will be presented in full. To-date, searches have identified 541 studies reporting data on symptoms/HRQOL in > 150,000 patients with CKD; 75 qualitative studies where patients with CKD or HCPs were interviewed about symptoms and HRQOL; and 20 studies reporting development and/or validation of measures for use in CKD. Symptoms have been reported in 146 studies using 55 different outcome measures, the most common being the: Beck Depression Inventory ( $n = 39$ ), Center for Epidemiologic Studies Depression Scale ( $n = 11$ ), and Hospital Anxiety and Depression Scale ( $n = 13$ ). HRQOL has been reported in 371 studies using 10 outcome measures, the most common being the Medical Outcomes Study Short Form 36 ( $n = 126$ ) and Kidney Disease Quality of Life Short Form ( $n = 51$ ). **Conclusion:** The final review synthesis will, for the first time, include a comprehensive summary of symptom and HRQOL trajectories across all stages of CKD; and will inform the development of a novel renal item-bank and CAT.

#### **(3004) Health-related quality of life in adults born extremely preterm—a matched controlled cohort study**

Merete Benestad, Master Degree in Nursing, Haukeland University Hospital, Bergen, Norway; Jorunn Drageset, RD, PhD professor, Department of Clinical Science, Faculty of Medicine and Dentistry, University of Bergen, Bergen, Norway, Bergen, Norway; Geir Eide, PhD, Professor, Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway Centre for Clinical

Research, Haukeland, University Hospital, Bergen, Norway, Bergen, Norway; Bente Johanne Vederhus, PhD, Department of Pediatrics, Haukeland University Hospital, Bergen, Bergen, Norway

**Aims:** Increased numbers of extremely preterm-born (EPB) children survive into adulthood. Several studies have reported that health-related quality of life (HRQoL) is inferior in EPB individuals compared to term born (TB) peers in adolescence and in young adulthood. However, we have sparse knowledge about HRQoL in EPB into adult life. This study examines HRQoL and subjective health at 34 years of age in EPB and TB subjects. **Methods:** A population-based cohort of subjects born at gestational age  $\leq 28$  weeks or with birth weight  $\leq 1000$  g and individually matched term-born controls completed the Short Form Health Survey-36 (SF 36/RAND-36) and Health Behavior in School-aged Children–Symptom Checklist (HBSC-SCL) at 34 years of age. Higher score in RAND 36 indicates better HRQoL and higher score in HBSC-SCL indicates more health complaints. **Results:** Of the 48 eligible EPB subjects, 35 (73%) participated. The mean ( $\pm$  standard error) gestational age was 27.5 ( $\pm 1.5$ ) weeks, birth weight was 1010.0 ( $\pm 190.4$ ) g, and the mean age was 34.3 ( $\pm 1.4$ ) years. In total 35 (76%) matched TB subjects participated. Among the EPB participants, eight had education at college or University level and 27 subjects were working. Among the EPB subjects, six had disability and 29 were without disability. For the RAND36 subdomain Mental Health, the mean score difference (95% confidence interval) was  $-17.6$  ( $-35.1, -0.2$ ) in EPB with disability vs. TB subjects. EPB without disability reported mean scores differences in Vitality  $-19.2$  (35.0, 3.3) and Role Emotional  $-22.8$  ( $-45.6, -0.1$ ) vs. TB subjects. For HBSC-SCL there were no significant differences between the EPB adults with disability and TB subjects in self-rated somatic and psychological health complaints. EPB adults without disability reported 3.4 (0.6, 6.3) higher mean score in self-rated psychological health complaints, and total complaints mean score 5.6 (0.6, 10.7) vs. TB subjects. **Conclusion:** This study demonstrates that EPB adults with and without disability reported lower HRQoL especially in the psychosocial domains vs TB subjects. The subjects without disability reported significantly more subjective health complaints vs TB subjects.

#### **(3005) Interference Score: a tool to identify barriers to adhere to Antiretrovirals therapy**

Kedar K.V. Mate, Mayo Clinic/McGill University Health Centre, Montreal, Quebec, Canada; Kim Engler, PhD, Center for Outcomes Research and Evaluation, McGill University Health Centre Research Institute, Montreal, Quebec, Canada; David Lessard, PhD, Center for Outcomes Research and Evaluation, McGill University Health Centre Research Institute, Montreal, Quebec, Canada; Serge Vicente, McGill University, Montreal, Quebec, Canada; François Raffi, Université de Nantes, Nantes, France; Darrell Tan, University of Toronto, Nantes, Quebec, Canada; Christine Jacomet, CHU Clermont-Ferrand, Clermont, France; Sara Ahmed, PT, PhD, McGill University, Montreal, Quebec, Canada; Alexandra de Pokomandy, MD, Family Medicine, McGill University, Montreal, Quebec, Canada; Joseph Cox, MD, Family Medicine, McGill University, Montreal, Quebec, Canada; Bertrand Lebouhe, MD, Family Medicine, McGill University, Montreal, Quebec, Canada; Bertrand Lebouhe, MD, Family Medicine, McGill University, Montreal, Quebec, Canada

**Aims:** Adherence to antiretroviral therapy (ART) can be challenging for some people living with HIV (PLHIV). Routine screening for barriers to ART adherence could help make HIV care more patient-centered and prevent virologic failure. The objective of this project was to identify barriers to ART by PLHIV and healthcare providers (HCP) to develop a digital HIV-specific patient-reported outcome

measure of barriers to adherence, the Interference Score (I-Score), to be administered through a digital application prior to a routine clinical consultation. Methods: A multi-site two-step Delphi survey targeting PLHIV and HCP in Canada and France was shared from March–December 2019. The survey consisted of 100 items on barriers to ART identified from the literature and a pool of PLHIV and HCP. The participants rated each item on three qualifiers: importance as an adherence barrier, relevance for HIV care and clarity of items on a 4-point ordinal scale. Each item was given a total score as the product of importance and relevance qualifiers. Items were hierarchically arranged based on the number of responses (total score of  $\geq 9$ ; at least a rating of 3 for importance and relevance qualifiers). These high-ranking items will be reworded based on the ‘clarity’ qualifier by a pool of experts for the second round of the survey. Results: The survey was sent out to 28 PLHIV and 38 HCP in Canada and 22 PLHIV and 35 HCP in France with a response rate of 79% and 73% in Canada; and 84% and 69% in France respectively, whose characteristics are shown in Table 1. Table 2 presents the top 20 barriers and proportion of PLHIV and HCP endorsing these barriers on ‘importance’ and ‘relevant’ qualifiers. Over two-thirds (13/20) of items were concordant among the two groups. However, a higher proportion of HCP compared to PLHIV valued these barriers as ‘important’ and ‘relevant for HIV care.’ PLHIV endorsed barriers related to stigma, comorbidities, isolation, appearance, and distress to a greater extent than HCP (discordant). Conclusion: The Interference score will target all the different patterns endorsing barriers to adherence defined by PLHIV and not only those-focused on treatment- selected by HCP in order to improve care.

Table 1: Sociodemographic characteristics of the PLHIV and HCP

Profile variables	Mean (SD) or N [%]
<b>PLHIV (n=40)</b>	
Country of residence	
Canada	24 [60]
France	16 [40]
Language	
English	16 [40]
French	24 [60]
Education	
Primary	4 [10]
Secondary	9 [23]
College/CEGEP/Technical degree	13 [34]
University	11 [28]
Other	2 [5]
Sexual orientation	
Heterosexual	23 [58]
Homosexual	13 [33]
Bisexual	3 [8]
History of drug use	
Yes	10 [25]
No	30 [75]
<b>HCP (n=56)</b>	
Country of practice:	
Canada	32 [57]
France	24 [43]
Language	
English	28 [50]
French	28 [50]
Gender	
Women	38 [68]
Men	18 [32]
Duration of service delivery to people with HIV	
5-9 years	28 [51]
10-14 years	7 [14]
15-19 years	8 [15]
20 years or more	12 [20]
Location of practice	
Hospital	35 [65]
Private clinic	7 [13]
Both	12 [22]
Percent practice that involve people with HIV?	6.2 (2.7)

Table 2: Top barriers and proportion of endorsement by the PLHIV and HCP as ‘important’ and ‘relevant for HIV care’ (concordant items are shown in color)

Sr. no.	People living with HIV(n=40)		Healthcare providers*(n=56)	
	Items	%	Items	%
1.	Having another health condition to deal with (for example, depression, diabetes or heart disease)	70	Being tired of taking my medication every day	95
2.	Being tired of taking my medication every day	65	Having side effects from my medication	95
3.	Not feeling motivated to take my medication	65	Not having a stable or suitable place to live	93
4.	Not having a stable or suitable place to live	63	Having side effects that interfere with my daily activities	91
5.	Having side effects that interfere with my daily activities	63	Forgetting	88
6.	Being concerned about stigma or discrimination related to HIV	60	Not feeling motivated to take my medication	88
7.	Feeling sad or depressed	60	Feeling my medication is toxic or harmful	88
8.	Being too sick or ill	60	Not wanting others to notice that I take this medication	88
9.	Forgetting	58	Being concerned about stigma or discrimination related to HIV	88
10.	Having trouble getting food or the right kind of food	58	Not having insurance to cover my medication costs or not having enough coverage	88
11.	Fearing rejection because of HIV	23	Having trouble fitting my medication into my daily life	48
12.	Getting discouraging test results (viral load or CD4 cell count)	23	Feeling sad or depressed	48
13.	Having side effects from my medication	22	Not being given enough information by my primary provider about my medication or how to take it	47
14.	Feeling my medication is toxic or harmful	22	Having trouble getting food or the right kind of food	46
15.	An irregular or unpredictable schedule	22	My primary provider having an unsupportive or negative attitude	46
16.	Feeling isolated or alone	22	Feeling medication is only for when you feel sick	45
17.	Not getting the support I need from others	22	Doubting that I need my medication	45
18.	Worrying about my medication's effects on my physical appearance	22	An irregular or unpredictable schedule	44
19.	Not being informed enough about my medication	22	My medication schedule conflicting with my daily activities	44
20.	My primary provider having an unsupportive or negative attitude	21	Doubting my medication's effects on HIV	44

\* Healthcare providers included physicians, nurse practitioners, pharmacists, social workers, psychologists, and community workers.

**(3006) Women discharged from inpatient cardiology units report a worse experience: results from 4 years of survey data**

Kyle Kemp, University of Calgary, Calgary, Alberta, Canada; Hude Quan, University of Calgary, Calgary, Alberta, Canada; Colleen Norris, University of Alberta, Edmonton, Alberta, Canada; Maria Santana, University of Calgary, Calgary, Alberta, Canada

Aims: When compared with their male counterparts, women with cardiovascular disease experience disparities in the diagnosis, treatment, and ongoing management of their condition. It is unknown, however, whether sex differences exist with respect to one’s in-hospital patient experience (i.e., patient-reported feedback about care processes and perceived quality of care). Using validated surveys, our objective was to examine to the experience of patients discharged from cardiology units in two urban hospitals in Alberta, according to sex. Methods: Patients completed the Canadian Patient Experiences Survey—Inpatient Care (CPES-IC) via telephone within 6 weeks of discharge. Surveys and administrative records were linked to determine study eligibility. We examined 32 individual questions and 4 rating scales. Individual questions pertained to specific care processes (e.g., how often nurses explained things in a way you could understand, how often your pain was well controlled), while rating scales asked respondents about their overall care, and their willingness to recommend to hospital to friends/family. Survey responses were reported as percent in “top box,” as represented by the most positive answer choice to each question. Results: From April 2014 to March

2018, 2776 surveys (942 females, 1,834 males) were completed across the five inpatient cardiology units studied. Female respondents had lower “top box” response percentages on all 36 questions studied. The largest differences observed were for questions that pertained to being told about possible side effects of new medications (36.8% of females vs. 51.6% of males responded “always”), cleanliness of the hospital room/bathroom (47.3% vs. 61.0% responded “always”), and getting enough information about your condition and treatment while in the emergency department (46.2% vs. 59.2% responded “completely”). Women reported a lower overall hospital experience (mean score of 8.7 vs. 9.0 [out of 10]) and were also less likely to “definitely” recommend the hospital to friends and family (75.3% vs. 84.1%). Conclusion: Reports have shown that nearly 80% of women miss warning signs of heart events because of inequities and biases in the healthcare system. In stratifying inpatient experiences by sex, this study provides evidence to suggest that these inequities and biases may also be reflected in inpatient care.

### (3007) Health-related quality of life in Tetralogy of Fallot patients with surgery: systematic review and meta-analysis

Zaiba Shafik Dawood, Medical Student, Aga Khan University Hospital, Karachi, Pakistan; Dr.mahin Janjua, MD, The Aga Khan University Hospital, Karachi, Pakistan; Sardar Shahmir Chauhan, Medical Student, The Aga Khan University, Medical College, Karachi, Pakistan; Dr.mahim Malik, MD, The Aga Khan University Hospital, Karachi, Pakistan; Dr.laila Ladak, PhD,MscN,BscN,RN, The Aga Khan University, Karachi, Pakistan

**Aims:** With the improvement in diagnostics and clinical/surgical management, many patients with Tetralogy of Fallot (ToF) are surviving till adulthood. Due to increased survival, the impact on their quality of life following surgery becomes very important. The objective of this study is to conduct a systematic review and Meta-analysis of the Health-Related Quality of life (HRQOL) in patients who underwent ToF repair. The systematic review protocol is registered with PROSPERO—registration number CRD42019145140  
**Methods:** A systematic search was conducted using the PubMed, Web of Science and CINAHL databases. Studies focusing on HRQOL in ToF surgical patients, conducted after 1990, with patients’ mean age  $\geq 18$  years and that had comparison with healthy controls were included. The analysis was done using Revman V5.3 using random-effects model. Results: The 17 studies (16 using Short Form 36—SF-36) included in the meta-analysis encompassed 1818 patients and 32554 healthy controls. There was a higher proportion of males (59% males). The mean age of assessment was 30.3 years and the mean age at surgery was 5.37 years (range: 2.6–6.9 years). All the studies reviewed were from high income countries. This review reports that ToF repair patients had lower scores in the physical domain (SMD = -0.29 CI -0.55, -0.04), General health (SMD = -0.30 CI -0.64, 0.03), Vitality (SMD = -0.08 CI -0.28, 0.11), Mental Health (SMD = -0.01 CI -0.19, 0.17) and Mental component Summary (SMD = -0.62 CI -1.39, 0.14). However the HRQOL was either equal like in Limitation in Physical role (SMD = 0.06 CI -0.27, 0.38), Limitation in Emotional role (SMD = 0.00 CI -0.14, 0.13) or even better in bodily pain (SMD = 0.36 CI 0.09, 0.63). This review highlights a deficit in the literature from low middle income countries (LMIC), given higher incidence of congenital heart disease in these countries with resource scarcity. Conclusion: ToF surgical patients have similar or even better HRQOL as compared to healthy controls in almost all domains other than the physical domain. This implies that early ToF repair surgery has a good effect on HRQOL in the long term. Literature is needed from LMIC to get an understanding of the HRQOL in ToF from resource constrained countries.

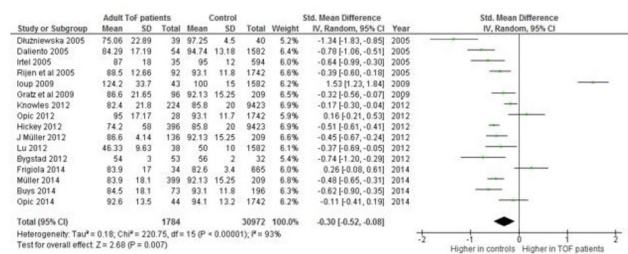


Figure 1: Forest Plot for physical functioning with a medium effect size.

### (3008) Factors associated with quality of sexual life among HIV positive men

Fatima Etemadi, Master, EA7334, Patient-Centered Outcomes Research, Paris, France; Pascal Bessonneau, PhD, Patient-Centered Outcomes Research, Paris, France; Ben Huntingdon, MD, University of Sydney, School of Psychology, Sydney, Australia; Aichata Fofana Dara, MD, University Paris-Diderot, EA7334, Patient-Centered Outcomes Research, Paris, France; Ilona Juraskova, PhD, University of Sydney, School of Psychology, Sydney, Australia, Sydney, Australia; Mariliza Henrique Da Silva, MD, Center of treatment HIV/AIDS, Sao Paulo, Brazil, Sydney, Brazil; James William Griffith, PhD, Northwestern University, Chicago, Illinois, United States; Fredrique Thonon, PhD, AP-HP, Hotel-Dieu Hospital, Health Economics Clinical Trial Center, Paris, France; Réjean Thomas, MD, CLINIQUE MÉDICALE L'ACTUEL, Montreal, Quebec, Canada; Sara Rodriguez, Master, hAP-HP, Hotel-Dieu Hospital, Health Economics Clinical Trial Center, Paris, France, Paris, France; Oliver Chassany, Master, hAP-HP, Hotel-Dieu Hospital, Health Economics Clinical Trial Center, Paris, France, Paris, France; Oliver Chassany, PhD, hAP-HP, Hotel-Dieu Hospital, Health Economics Clinical Trial Center, Paris, France, Paris, France; Martin Duracinsky, PhD, AP-HP, Bicetre Hospital, Internal Medicine and Clinical Immunology Department, Kremlin-Bicêtre, France, Paris, France

**Aims:** The high prevalence of sexual difficulties among people living with HIV implies the importance of evaluating the different dimensions of quality of sexual life. The objective of this study was to determine the factors associated with quality of sexual life among HIV positive men. **Methods:** We conducted a cross-sectional study in HIV centers in five countries (France, Australia, Brazil, USA and Canada) among HIV-positive men. Data related to physical and mental health status and HIV parameters including, were self-reported. We used the three main dimensions of PROQOL-SexLife instrument for measuring the outcome: positive sexual perception, sexual dysfunction and stigma/fear, PROQOL HIV Sexlife dimensions are scored from 0 (best) to 100 (impaired). **Results:** 107 Heterosexual men and 474 Men who have Sex with men (MSM) were included (Australia: 109; Brazil: 139; Canada: 79, France: 192; USA: 64). Among MSM, mean ( $\pm$  SD) score of the first, second and third dimensions were respectively 38(22.4), 42(23.9), 37 (25.4), and among Heterosexual men they were respectively 41(64), 36(24.3), 42(27.3). In linear mixed regression model, marriage and being in couple was associated with positive sexual perception, lower experiences of stigma and less sexual difficulties among MSM with regression coefficient of [-5.4, 95% CI (-9, -5),  $P < 0.05$ ] and [-22, 95% CI (-35, -10),  $p < 0.001$ ] meaning that being in couple improve positive sexual perception by 22 points. Lower mental health status was associated with higher sexual dysfunction, experience of stigma and more negative sexual perception ( $p < 0.05$ ). Having diabetes and being under lipid-lowering therapy was found to be in association with sexual dysfunction among MSM ( $p < 0.001$ ). Among heterosexual men, African ethnicity and unemployment was respectively associated with lower and higher values of the first dimension

[− 18.8, 95% CI (− 31, − 9),  $p < 0.05$ ] & [16.6, 95%CI (0.02, 34),  $p < 0.05$ ]. Conclusion: Stigma, fear and sexual dysfunction are important aspects for evaluating quality of sexual life among HIV positive men, which are in close association with poor mental health and comorbidity. Mental health support could be considered as an effective factor for improving sexual life among HIV positive men.

### (3009) The quality of informed consent: a Portuguese experience

Pedro Ferreira, PhD, CEISUC/FEUC, Coimbra, Portugal; Ana Barradas, MSc, CEISUC, Coimbra, Portugal

**Aims:** An informed consent is an ethical and legal mandatory document to be read and signed by a patient before a clinical trial. It expresses his/her decision to voluntarily participate in that trial and demonstrates that s/he has the necessary skills to perform it. It also specifies the participant's rights and the procedures s/he will be submitted. However, some studies reveal that participants do not always full understand the informed consent, one of the reasons for dropouts. The aim of this study was to evaluate the ability of understanding the informed consent and what were the possible conditions that lead to a better or worse understanding of it by the participants of a clinical trial. **Methods:** After being translated and validated into Portuguese, the Quality of Informed Consent (QuIC) questionnaire was implemented in 100 cardiac phase III clinical trials participants. We collected gender, age, family and professional situation and education level, as well as self-assessment of each participant about the study and his/her general health. **Results:** 85% of participants were male, their average age was 67.3, 70% were retired and 49% had only the primary school. All patients evaluated positively their participation and their own health, knew the main purpose of clinical trials, and 97% understood their role in helping future patients. 97% realized that by signing the informed consent they would be participating in a clinical trial. However, none of them knew that their experimental treatment was not proven to be the best alternative for their condition. At last, 70.8% mentioned that, when signed the consent, they understood what was the purpose of the trial. **Conclusion:** It was possible to conclude that the level of education and their social condition may not directly affect the understanding of consent, but the belief that the new treatment will be the only cure for their disease. It is necessary a greater awareness about the importance of reading the informed consent, so that participants could understand, as much as possible, the protocol of the clinical trial, especially the risks that come from an experimental treatment, and the clarification of their rights as a participant in the study.

### (3010) Preliminary validation of the Dutch Wound-QoL and findings on differences between self- and read-out-completion

Toni Maria Klein, M.Sc., University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Stella Amesz, University of Groningen, Groningen, Netherlands; Audrey Meulendijks, HU University of Applied Science, Utrecht, Netherlands; Tuong-Vi Nguyen, Erasmus University Medical Center, Rotterdam, Netherlands; Christine Blome, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Petrie Roodbol, University Medical Center Groningen, Hamburg, Netherlands; Catherine van Montfrans, Erasmus University Medical Center, Rotterdam, Netherlands

**Aims:** Chronic wounds have major impacts on patients' health-related quality of life (HRQoL). Validated and easy-to-use HRQoL questionnaires are essential for treatment. Therefore, this study aimed to translate and validate the German wound-specific HRQoL questionnaire Wound-QoL for Dutch patients. **Methods:** The original

questionnaire was translated using a forward-backward process. Participants answered the Wound-QoL, the EQ-5D-3L and a visual analogue scale (VAS) on pain at baseline and six weeks later. For patients who were not able to self-complete the questionnaires, nurses read out the questions (read-out group). Furthermore, sociodemographic data, medical record data, and wound size was obtained. Statistical analyses included calculation of floor and ceiling effects, internal consistency, item selectivity, convergent validity, and responsiveness. **Results:** Data of 120 participants showed few missing values, except for one item (about climbing stairs). Though, sensitivity analysis showed that this item had no impact on the validity results. Only minor ceiling effects were detected. Larger floor effects were seen, especially in the read-out group. Only in the self-completion group, global and subscale scores decreased significantly (i.e., HRQoL improved) over time. Item selectivity and internal consistency were similar in both patient groups. Cronbach's alpha was good regarding the global, everyday life, and psyche scales ( $\alpha = 0.794$  to  $\alpha = 0.925$ ), but reduced regarding the body scale ( $\alpha = 0.673$  to  $\alpha = 0.687$ ). Analyses on convergent validity and responsiveness showed significant associations between EQ-5D-3L and the Wound-QoL but inconsistent results for associations with pain VAS and wound size. **Conclusion:** The results are similar to those in other Wound-QoL validation studies, suggesting that the Wound-QoL is a valid and easy-to-use instrument. Strong correlations between generic and wound-specific HRQoL questionnaires but inconsistent results regarding pain and wound size suggest that other burdens experienced by patients (e.g., odor, exudate) might have even stronger impacts on HRQoL. It cannot be determined whether deviations between self-completing and read-out groups derived from differences occurring due to the method of data collection or from differences in age and wound duration between both groups. Incapability of older and more severely impaired patients to complete the current Wound-QoL version promotes the idea to develop a more visual, low-threshold version of this questionnaire.

### Applied Clinical Research II

#### (3011) A brief prototype weight-related quality of life index performs as well as or better than weight-specific or generic measures in bariatric surgery sample

Ana Maria Moga, Faculty of Medicine, School of Physical & Occupational Therapy, McGill University, Montreal, Quebec, Canada; Nancy E. Mayo, James McGill Professor, Faculty of Medicine, School of Physical & Occupational Therapy, McGill University, Montreal, Quebec, Canada; Laurie Twells, Associate Professor, Departments of Medicine and Pharmacy, Memorial University of Newfoundland and Labrador, St. John's, Newfoundland and Labrador, Canada

**Aims:** The increasing obesity prevalence, the number of weight-loss interventions, and the high demand for health care resources make evidence for comparative effectiveness, a matter of individual, clinical, public health importance. Preference-based health-related quality of life indices are fit for this purpose, but there are none for obesity. This study aimed to estimate the extent to which a prototype for a short multi-dimensional preference-based index of weight-related quality of life (PB-WRQL) distinguishes between known groups of individuals with severe obesity compared to concurrent measures at baseline and six-months post-surgery. **Methods:** The study data source was a Canadian longitudinal bariatric surgery cohort. Forty-eight items from the Impact of Weight on Quality of Life (IWQL), Euro-QoL-5D (EQ-5D), and the Short Form-12 (SF-12) were mapped to obesity-relevant domains. Rasch analysis identified one best performing item to form the prototype dimensions. Individuals' health

ratings were regressed on each response option of each prototype dimension, and the regression coefficients were used as weights in an additive model. Generalized estimation equations were used to compare measure parameters across groups and levels of converging constructs. Results: Table 1 presents the results from the 201 individuals (BMI:  $48.8 \pm 6.7$  kg/m<sup>2</sup>; Age:  $43 \pm 9.0$  years; 82% women) with data at baseline and those participants ( $n = 125$ ; 62%) with 6-month follow-up. The seven dimensions of the prototype PB-WRQL were: Physical Function, Mood, Participation, Pain, Vitality, Dyspnoea, and Ankle Oedema. There were substantial improvements from baseline to 6 months post-surgery all study measures with both weight-specific measures showing a greater change with bariatric surgery than the EQ-5D. Compared to the IWQL-Lite and EQ-5D, the prototype PB-WRQL showed a stronger relationship both with BMI ( $t = -3.68$ ) and self-reported health ( $t = 9.42$ ) at baseline. The prototype PB-WRQL was more sensitive to change in BMI ( $t = -3.42$ ) than the other two measures and equally sensitive to change in self-rated health ( $t = 2.27$ ). Conclusion: The current study shows that a brief prototype measure comprising seven dimensions weighted by health impact performed as well as the 31 items IWQL-Lite and better than the generic EQ-5D. These findings demonstrate the potential value of the brief PB-WRQL index and support its further development using preference weights.

Table 1. Effect of BMI and Health Rating on HRQL measures at baseline and six-months post-bariatric surgery

	BMI	VAS	HRQL Measure (0-100)		
			PB-WRQL	IWQOL-Lite	EQ-5D-3L
Mean baseline (SD)	48.8 (6.7)	59.7 (18.6)	43.0 (24.8)	43.2 (18.7)	77.9 (17.0)
Mean change (SD)	-11.1(3.0)	19.7 (18.5)	38.9 (24.7)	40.5 (19.6)	11.5 (15.0)
<b>HRQL baseline on BMI baseline (n=201)</b>					
Adjusted (age and sex)					
Estimate (SE)			-0.95* (0.25)	-0.43 (0.19)	-0.36 (0.18)
t-statistic			-3.68*	-2.18	-2.01
<b>HRQL change on BMI change six months post-surgery (n=125)</b>					
Adjusted (age and sex)					
Estimate (SE)			-2.63* (0.77)	-1.75 (0.57)	-0.95 (0.44)
t-statistic			-3.42*	-3.05	-2.16
<b>HRQL baseline on VAS baseline (n=201)</b>					
Adjusted (age and sex)					
Estimate (SE)			0.76* (0.08)	0.31 (0.07)	0.36 (0.06)
t-statistic			9.42*	4.37	6.02
<b>HRQL change on VAS change 6 months post-surgery (n=125)</b>					
Adjusted (age and sex)					
Estimate (SE)			0.30 (0.13)	0.30(0.09)	0.23 (0.07)
t-statistic			2.27	3.06	3.12

BMI: Body Mass Index (kg/m<sup>2</sup>)

VAS = Visual analogue scale for self-rated health on a 0 (worst) to 100 (best) scale.

\*Effect of variable is greater on prototype PB-WRQL than for IWQOL-Lite and EQ-5D (significant interaction between measure and variable in GEE model);

### (3012) Psychometric evaluation of the Indolent Systemic Mastocytosis Symptom Assessment Form (ISM-SAF) in patients with indolent systemic mastocytosis

Fiona Taylor, MBiochem, Adelphi Values, Boston, Massachusetts, United States; Cem Akin, MD, PhD, University of Michigan, Ann Arbor, Michigan, United States; Alan Shields, PhD, Adelphi Values, Boston, Massachusetts, United States; Shirley Li, PhD, Adelphi Values, Boston, Massachusetts, United States; Brenton Mar, MD, PhD, Blueprint Medicines Corporation, Cambridge, Massachusetts, United States; Ronny Oren, MS, Blueprint Medicines Corporation, Cambridge, Massachusetts, United States; Jeffrey McDonald, MS, Adelphi Values, Boston, Massachusetts, United States; Brad Padilla, MPH, Adelphi Values, Boston, Massachusetts, United States; Hui-Min Lin, PhD, Blueprint Medicines Corporation, Cambridge, Massachusetts, United States; Tanya Green, MS, Blueprint Medicines Corporation, Cambridge, Massachusetts, United States; Philina Lee, MS, Blueprint Medicines Corporation, Cambridge, Massachusetts, United States; Philina Lee, PhD, Blueprint Medicines Corporation, Cambridge, Massachusetts, United States; Frank Siebenhaar, PD Dr. med., Dermatological Allergy, Department of Dermatology and

Allergy, Charité -Universitätsmedizin Berlin; Corporate Member of Freie Universität Berlin, Humboldt - Universität zu Berlin; and Berlin Institute of Health, Berlin, Germany

**Aims:** To psychometrically evaluate the Indolent Systemic Mastocytosis Symptom Assessment Form (ISM-SAF©2018 Blueprint Medicines Corporation) among patients with indolent systemic mastocytosis (ISM), the most common subset of SM, a rare, clonal mast cell neoplasm associated with debilitating symptom burden. **Methods:** Administered as a daily diary in BLU-285-2203 (PIONEER), a Phase 2 trial to evaluate the safety and efficacy of avapritinib, a selective kinase inhibitor, in subjects with ISM, the ISM-SAF has 11 items assessing symptom severity on an 11-point numerical rating scale and one item assessing diarrhea frequency. The ISM-SAF severity items compose a Total Symptom Score (TSS), Gastrointestinal Symptom Score (GSS), and Skin Symptom Score (SSS). Psychometric evaluation of the ISM-SAF is supported by other BLU-285-2203 assessments: 12-item Short Form Health Survey (SF-12), Mastocytosis Quality of Life Questionnaire (MC-QoL), Five-level EuroQol Five Dimensions (EQ-5D-5L), Patients' Global Impression of Change (PGIC), and Patient Global Impression of Symptom Severity (PGIS). **Results:** Thirty-eight patients contributed to the analyses (78.9% female; mean age = 49 [ $\pm$  13]). At Baseline, internal consistency reliability ( $\alpha$ ) for the bi-weekly TSS, GSS, and SSS was 0.86, 0.83, and 0.82, respectively. Test-retest reliability among patients who exhibited no change in PGIS between Baseline and Day 15 was assessed using the intra-class correlation coefficient (ICC); ICCs exceeded 0.74 for all ISM-SAF scores. Construct validity and known-groups analysis showed ISM-SAF scores were moderately to strongly correlated ( $r = 0.382$ – $0.881$ ) to PGIS, MC-QoL symptom and skin scores and were able to distinguish among clinically unique groups (by PGIS, EQ-5D-5L Visual Analogue Scale, MC-QoL symptom, and SF-12 Physical Component Summary score). Correlations of ISM-SAF change score and other assessment change scores reflect evidence of score sensitivity. For the TSS, candidate clinically important between-group differences based on distribution-based methods ranged from 7 to 10; the clinically important response using a PGIS anchor was 19.0 (29.4% individual decrease from Baseline). **Conclusion:** The ISM-SAF produced reliable, construct-valid, and sensitive scores when administered in the target patient population. These results, along with the ISM-SAF's strong development history and evidence of content validity, support its use as the first fit-for-purpose daily symptom measure to evaluate clinical benefit of treatment interventions in individuals with ISM.

### (3013) Cardiopulmonary symptoms in atrial fibrillation: looking beyond the arrhythmia

Jeffrey Turner, DO, University of Utah, Salt Lake City, Utah, United States; Brian Zenger, University of Utah, Salt Lake City, Utah, United States; Jared Bunch, University of Utah, Salt Lake City, Utah, United States; Rachel Hess, University of Utah, Salt Lake City, Utah, United States; Benjamin Steinberg, University of Utah, Salt Lake City, Utah, United States

**Aims:** Patients with atrial fibrillation (AF) may have cardiopulmonary symptoms beyond those periods during which they are experiencing paroxysms of AF while in sinus rhythm (SR). It is unknown what symptoms are most commonly reported during SR in patients with AF and how this compares with patients without AF. **Methods:** We retrospectively evaluated ambulatory cardiac monitors to assess if patients with AF documented during the monitor reported different symptoms during SR compared with patients without any documented AF. **Results:** Among 1,075 patients (mean age 53, 64% female), 208 patients (19%) had a study with at least 1 AF episode (81% without

any AF). There were a total of 11,071 non-AF symptomatic events recorded in our dataset; 2103 (19%) in patients that also had paroxysmal AF on their monitors and 8,968 (81%) in patients without AF. Patients with documented AF were more likely to report symptoms of palpitations during SR than those without documented AF (40.8% versus 26.6%;  $p < 0.001$ ) and palpitations were the most commonly reported SR symptom in the AF group (Fig. 1). Those with AF had higher frequency of premature atrial contractions (PAC) (21.8% versus 6.3%;  $p < 0.001$ ). There was no difference in premature ventricular contractions between groups. Chest pain was more commonly reported in those without AF (33.0% versus 26.1%;  $p < 0.001$ ) and was the most common reported symptom in the non-AF group. Conclusion: Patients with documented AF on ambulatory monitoring are more likely to report palpitations when not in AF, compared with patients without AF. This may be related to cardiac preconditioning related to an increase in frequency of PACs, AF-related neurohormonal compensatory mechanisms still present during SR, or other noncardiac mechanisms. These findings provide valuable insight into symptom assessments in patients with AF and may inform treatment selection.

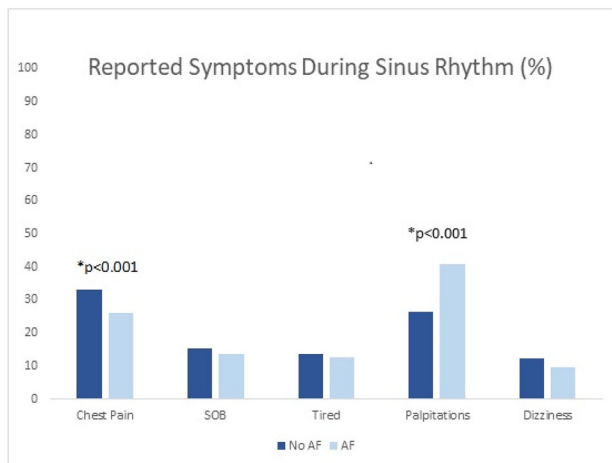


Figure 1: Symptoms reported during sinus rhythm in the group without AF present (No AF) and in those with AF documented on their monitor (AF)

#### (3014) The psychometric properties of the WHOQOL-BREF—a systematic review from 1998 to 2020

Li\_chung Lin, master, National Taiwan University, Taipei, Taiwan; Grace Yao, professor, National Taiwan University, Taipei, Taiwan

**Aims:** The short version of the World Health Organization Quality of Life (WHOQOL-BREF) is widely used in medical and public health fields. The psychometric properties of the WHOQOL-BREF have been examined in multiple countries and cultures. To more comprehensively understand its reliability and validity, a study for summarizing and synthesizing these results from the previous studies is necessary. The present study conducts a systematic review of the psychometric properties of the WHOQOL-BREF. **Methods:** We used the keywords (including WHOQOL-BREF, psychometric properties, construct validity, reliability, exploratory factor analysis, and confirmatory) to find more than 3,000 articles published from 1998 to April 2020 from MEDLINE and PsycINFO. After excluding the articles that were incompatible with the purpose of this study, 52 articles were

left. **Results:** The result showed that the physical, psychological, social relationship, and environmental domains of the WHOQOL-BREF have acceptable Cronbach'  $\alpha$  (with mean  $> .75$  for each domain, except for the social relationship domain with mean = 0.68) and test–retest reliability (with mean  $> .75$  for each domain). All these four domains showed good discriminant and criterion-related validity, representing a significant difference between the healthy and unhealthy groups, and exhibiting significantly positive or negative correlations with the questionnaires measuring health or syndrome of illness, respectively. However, the WHOQOL-BREF performed poorly in construct validity. Four- to eight-factor models were found from different studies. In other words, the four-factor model originally proposed by the WHO was not fully supported by these studies. After investigating these 52 articles, we found that inappropriate statistical analysis approaches may be a reason for poor construct validity. **Conclusion:** The WHOQOL-BREF has acceptable internal-consistency and test–retest reliabilities, and discriminant and criterion-related validities. However, its construct validity is inconsistent. Based on psychometric theory and simulation studies, we recommend the following procedure for evaluating the construct validity of the WHOQOL-BREF: (1) conduct exploratory factor analysis using iterative principal axis factoring, Promax rotation, and multiple methods to determine the number of factors, and then (2) conduct confirmatory factor analysis with stricter cut-off scores on goodness-of-fit indices and cautiously adopt modification index for evaluating the psychometric properties of the WHOQOL-BREF.

#### (3015) Quality of life in patients with primary hyperparathyroidism after parathyroidectomy

Inna Gladkova, University Hospital, Saint-Petersburg State University, Saint-Petersburg, Russia; Roman Chernikov, University Hospital, Saint-Petersburg State University, Saint-Petersburg, Russia; Sergey Efremov, University Hospital, Saint-Petersburg State University, Saint-Petersburg, Russia; Tatyana Ionova, University Hospital, Saint-Petersburg State University, Saint-Petersburg, Russia; Tatiana Nikitina, University Hospital, Saint-Petersburg State University, Saint-Petersburg, Russia; Vladimir Rusakov, University Hospital, Saint-Petersburg State University, Saint-Petersburg, Russia; Natalia Shablovskaya, University Hospital, Saint-Petersburg State University, Saint-Petersburg, Russia

**Aims:** Primary hyperparathyroidism (PHPT) is a complex endocrinopathy involving calcium metabolism and a potent hormone made by the parathyroid glands, which may lead to significant quality of life (QoL) impairment. We aimed to evaluate changes in QoL of PHPT patients after parathyroidectomy (PTX) and to test applicability of the Russian version of the disease-specific questionnaire PHPQoL in this patient cohort. **Methods:** Patients with PHPT who were referred to PTX filled out PHPQoL and SF-36 tools before and 3 mos after PTX. PHPQoL is a disease specific tool which includes 16 items; the total PHPQoL score is standardized from 0 (worst QoL) to 100 (best QoL). The Russian version of PHPQoL was developed in accordance with international guidelines. General linear models (GLM) and Wilcoxon test were used for statistical analysis. **Results:** 47 PHPT patients were enrolled in the study—mean age (SD)—54.2 (9.1) years; 98% were females; 64% pts were symptomatic; 47% pts had severe hypercalcemia. Mean duration of PHPT since diagnosis—47.7 weeks. Before PTX QoL by the majority of SF-36 scales was decreased. The lowest scores by SF-36 were for vitality, mental health and role functioning scales—their means (SD) varied from 48 (21) to 56 (40). Median PHPQoL score was 53 (IQR: 42–64). Symptomatic



pts had worse QoL than asymptomatic pts by PHPQoL total score: means (SE)—48 (3.0) vs 58 (3.8) (GLM,  $p < 0.05$ ). Remarkable QoL improvement was observed 3 mos after PTX. According to SF-36, the largest changes were found for role functioning scales, vitality and mental health ( $p < 0.05$ ):  $\Delta$  varied from 19 to 47 scores. PHPQoL score increased borderline significantly as compared to baseline: means (SD)—45 (14) vs 57(9) (effect size—1.03;  $p = 0.05$ ); difference was clinically significant (MID for PHPQoL  $\geq 9$  points). The sensitivity to change of the Russian version of PHPQoL was demonstrated. In addition, it was shown that the tool was informative to capture specific concerns of PHPT patients, especially in physical functioning. Conclusion: PTX leads to noticeable QoL improvement in PHPT patients. The Russian version of PHPQoL is a feasible and practical PHPT specific tool for QoL assessment in patients undergoing surgery.

### **(3016) When providers become the center of our attention: Can we comfort them? A qualitative thematic analysis from a hospital-wide cross-sectional survey**

Eliezer Oliveira, MBA, Children's National Medical Center, Washington, District Of Columbia, United States; Kathleen Ennis-Durstine, Children's National Medical Center, Washington, District Of Columbia, United States; Sabrina Figueiredo, PhD, The George Washington University, Washington, District Of Columbia, United States

Aims: Hospital healthcare providers (HCP), in addition to working long hours, deal with constant life-threatening situations. Not surprisingly, burnout reached epidemic levels and its manifestation spirals to a cascade of negative outcomes. HCP experience a myriad of symptoms that affects their empathy—a crucial component of person-centered care. Patients rate their hospital experience as poor. Institutions report increased errors and staff disengagement. Thus, to build resilience among hospital staff and potentially prevent burn-out, Chaplains from the Children's National Medical Center developed Comfort Corner (CC)—an emotional support strategy offered to hospital units. The aim of this study was to estimate how staff perceive the support of CC on behavior modification, patient safety and work environment. Methods: Comfort Corner is a 90-min weekly session offering emotional support, mindful activities, snacks, and hot beverages. Sessions are held in a quiet room within each hospital unit. Attendees stay on average 10–15 min. Cross-sectional survey measured staff perception. Deductive qualitative analysis mapped themes related to needs, behavior modification, patient safety and work environment. Inductive analysis mapped additional themes. Results: Convenience sample of 248 respondents representing all disciplines and participating hospital units. After attending CC, 97% of respondents said their needs were met. The most frequent need was having protected time for decompressing/relaxation (83%) followed by snacks and hot drinks (45%). Behavior change was indicated by 76% of respondents and the most prevalent activity incorporated into staff self-care routine was taking regular “me” breaks (21%) followed by aromatherapy and coloring (12%). Of great interest, 80% of respondents indicated providing safer care after attending CC because they think more clearly (41%); are less tired (19%); and feel less anxious (15%). Respondents also indicated that CC enabled team engagement (84%) and influenced their decision to remain in the hospital (12%). Conclusion: CC is perceived by staff as a great resource as there was a sense of improved morale and appreciation. It may also have a positive impact in reducing employment turnover. The Institute of Healthcare Improvement suggests that improving joy in work is one out of the seven innovations that can improve healthcare. Further studies should investigate CC's efficacy and cost-efficiency.

### **(3017) The patient experience with symptoms of hereditary angioedema (HAE) before, during, after, and between attacks**

Milenka Jean-Baptiste, MPH, Evidera, Bethesda, Maryland, United States; Dylan Supina, PhD, CSL Behring, King of Prussia, Pennsylvania, United States; Robbin Itzler, PhD, CSL Behring, King of Prussia, Pennsylvania, United States; Subhransu Prusty, MD, CSL Behring, Marberg, Germany; Mona L. Martin, RN MPA, Evidera, Seattle, Washington, United States

Aims: Hereditary Angioedema (HAE) is a genetic disorder characterized by episodic painful swelling attacks throughout the body. The treatment landscape for HAE patients is now changing with the introduction of new prophylactic therapies. Symptoms that occur during attacks have been well characterized but those occurring outside of a swelling attack are not well understood. This study aimed to describe the full range of the HAE patient symptom experience in order to identify whether current measurement tools adequately capture the changes patients are experiencing with new treatments. Methods: A targeted literature review identified published information on HAE symptoms. Interviews with expert clinicians described the symptoms patients reported to their doctors. Symptoms identified were categorized based on the type of symptom and when in relation to the swelling attack it was experienced (before, during, after, or in the longer periods between attacks). Results: Published literature reported 21 different groups of symptoms the authors heard frequently from patients. Additionally, eleven clinician interviews yielded 39 HAE symptoms. The symptoms reported most often were fatigue, malaise, weakness, erythema marginatum, non-itchy rash, paresthesia, stinging, tingling, skin tightness, prickliness, abdominal issues (cramps, GI, upset stomach, loose stool), nausea, vomiting, pain (general, abdominal, muscle, and joint), flu-like symptoms, pruritus, depression and anxiety. Facial, and extremity swelling occurred mostly during the attack. Abdominal swelling occurred at all time points. Flu-like symptoms and rash and paresthesia's were often seen as prodromal, but also reported following attacks. Other symptoms such as pain, headache, fatigue, abdominal and GI symptoms, depression and anxiety were reported across all time points and could vary greatly with gender, HAE severity, and type of treatment. Conclusion: Despite varied terminology, there are numerous symptoms reported for HAE. Most are associated with attacks or labeled as “prodromal.” Very little is said about the “between attack” experience of patients and how it varies with treatment. Qualitative interviews with HAE patients are needed to better understand their symptom burden and improve the ability to evaluate new treatments.

### **(3018) Social participation of people with chronic wounds—a systematic review**

Toni Maria Klein, M.Sc., University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Valerie Andrees, M.Sc., University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Natalia Kirsten, Dr., University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Matthias Augustin, Prof. Dr., University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Christine Blome, PD Dr., University Medical Center Hamburg-Eppendorf, Hamburg, Germany

Aims: Living with chronic ulcer can be burdensome and impose restrictions, not only regarding people's physical and mental health but also regarding their social life. Therefore, this review aims to analyze social participation in people with chronic wounds and to compare results across different wound aetiologies. Methods: A search string was applied in several electronic databases. Duplicates were removed and results screened in a two-step process comprising

title and abstract screening, and full-text assessment. Inclusion (e.g., original article, social participation major outcome) and exclusion (e.g., non-dermatological ulcer) criteria were pre-defined and applied in both steps of screening. Literature cited in relevant reviews was screened accordingly. Data of eligible articles were extracted and synthesized narratively. Results: The search revealed 42 eligible publications. The most frequently studied population were patients with venous leg ulcers, followed by any leg ulcers, diabetic foot ulcers, and pressure ulcer. In 16 studies, social participation was treated as distinct construct, whereas other studies regarded it as subdomain of health-related quality of life or as aspect of another construct. Included studies showed few differences across ulcer aetiologies. Overall, family members were the major social contacts for patients and often provided wound care and emotional support. Patients had few non-family relations, but those were of remarkably strong ties. Patients felt guilty as their condition led to burdens for family and friends. With nurses, a unique relationship was described when there was a continuous patient-nurse relation. Patients experienced restrictions in various activities, which were caused by direct and indirect consequences of the wound. Social support and social connections were reduced in ulcer patients compared to healthy controls. Inconsistent results were found with regard to whether social isolation was higher in cases than in controls. Conclusion: This review showed impairments in all aspects of social participation for people with chronic wounds. Especially family can be regarded as important social support. Furthermore, the special relationship with nurses should be acknowledged and might be strengthened by continuity in care. Additionally, comprehensive implementation of ulcer-specific projects offering both professional care and interaction with other people would allow all patients to uptake new activities and to meet fellow sufferers.

### **(3019) Burden of mental health symptoms and perceptions of their management in hemodialysis care: a mixed-methods study**

Kara Schick-Makaroff, PhD, University of Alberta, Edmonton, Alberta, Canada; Lisa Wozniak, MA, University of Alberta, Edmonton, Alberta, Canada; Hilary Short, MSc, University of Alberta, Edmonton, Alberta, Canada; Sara Davison, MD, FRCP(C), MHSc, University of Alberta, Edmonton, Alberta, Canada; Scott Klarenbach, MD, MSc, FRCPC, University of Alberta, Edmonton, Alberta, Canada; Robert Buzinski, Patient Partner, Edmonton, Alberta, Canada; Jeffrey Johnson, PhD, University of Alberta, Edmonton, Alberta, Canada

Aims: Chronic kidney disease requiring dialysis is associated with poor quality of life and a range of physical and mental symptoms. Unfortunately, mental health symptoms are often unrecognized and undertreated. We aimed to describe 1) the burden of depressive and anxiety symptoms reported by adults on in-center hemodialysis in Northern Alberta using routine patient-reported outcome measures (PROMs), and 2) patients' and nurses' perceptions of managing such symptoms. Methods: A mixed-methods approach was employed. We used baseline data from a randomized controlled trial to describe the prevalence of positive screens (i.e., scores  $\geq 3$ ) for depressive (PHQ-2) and anxiety (GAD-2) symptoms. We used interpretative description to describe patients' and nurses' perceptions of managing depressive and anxiety symptoms. Using purposeful sampling, we invited patients and nurses to participate in individual interviews. We also conducted site visits in dialysis units documenting our observations in field notes. We compiled both patients' responses to open-ended survey questions from the trial and nurses' electronic chart notes related to mental health. Qualitative data were managed using ATLAS.ti 8 and analyzed using thematic analysis. Results: The

average age of patients at baseline ( $n = 408$ ) was 64.0 years (SD 15.4), 57% were male, and 87% were not working. 29% screened positive for depressive symptoms, 21% for anxiety symptoms, and 15.7% for both. We conducted interviews with 10 patients, 8 nurses, undertook 6 observations in the dialysis units, and reviewed 510 open-ended survey responses from patients and 84 nurses' chart notes. We found that PROMs had the potential to identify mental health concerns. However, opinions differed between patients and nurses about whether mental health was within the scope of kidney care. Nonetheless, participants agreed there was a lack of resources to address mental health. Lack of privacy in dialysis units was identified, however, only nurses explicitly perceived this as a barrier to addressing mental health concerns in this setting. Conclusion: The prevalence of depressive and anxiety symptoms was similar to the existing literature. PROMs can prompt communication about mental health concerns. However, there needs to be more discussion and resources (e.g., private spaces) with clinicians about mental health in kidney care.

### **(3020) Qualitative interviews with clinicians who treat hyperphagia**

Glenn Phillips, PhD, Rhythm Pharmaceuticals, Boston, Massachusetts, United States; Timothy Howell, MA, Evidera, Bethesda, Maryland, United States; Louis Matza, PhD, Evidera, Bethesda, Maryland, United States; Yestle Kim, Rhythm Pharmaceuticals, Boston, Maryland, United States; Katie Stewart, MA, Evidera, Bethesda, Maryland, United States; Amy Gross, University of Minnesota, Bethesda, Minnesota, United States; Claudia Fox, University of Minnesota, Minneapolis, Minnesota, United States

Aims: Hyperphagia is the drive to eat excessively without reaching satiation. This condition is commonly associated with obesity-related genetic disorders such as Prader-Willi Syndrome, and it is also a factor in the global obesity epidemic. The purpose of this study was to identify symptoms and behaviors that clinicians consider to be associated with hyperphagia and to learn about their approach to assessing for hyperphagia. Methods: Telephone interviews were conducted with clinicians in the United States from July 2019 to May 2020. Clinicians were asked about their professional background, experience treating patients with hyperphagia, the definition of hyperphagia, and the symptoms/impacts they see in pediatric and adult patients with hyperphagia. Finally, clinicians were asked about how they assess hyperphagia. Results: Twelve clinicians were interviewed (8 MD, 2 NP, 1 PhD, 1 DO). Their specialties varied (e.g., obesity medicine, weight management, endocrinology), but all had experience treating hyperphagia in adults ( $n = 3$ ), children ( $n = 6$ ), or both ( $n = 3$ ). All clinicians agreed that a definition of hyperphagia should include two parts: (1) excess hunger with difficulty achieving and maintaining satiety, and (2) excessive food seeking behavior and behavioral problems related to food. Commonly reported symptoms of hyperphagia included being "hungry all the time" and "never full/satisfied." Clinicians reported that some symptoms are similar across age groups (excess hunger, difficulty achieving satiety), but the ability to report these symptoms improves as patients get older. Clinicians reported that pediatric patients often have behavioral problems related to food, such as sneaking food without parents' knowledge, throwing tantrums when denied food (even immediately after a meal), eating items off others' plates, eating food from the trash, or consuming non-meal items such as flour, ketchup, sugar, and syrup. Adult patients often try to avoid social situations involving food. Clinicians agreed that there are no standardized instruments for screening or assessment of hyperphagia, but that such an instrument could be useful in their practice. Conclusion: Clinicians generally agreed on the definition and most common symptoms/behaviors associated with hyperphagia.

Given the impact of hyperphagia, it would be useful to develop measurement tools to screen for this condition and assess its severity.

### Applied Clinical Research III

#### (3021) Both disease specific and generic domains contribute to overall health-related quality of life in patients with end stage kidney disease (ESKD)

Istvan Mucsi, MD PhD, University of Toronto Health Network, Toronto, Ontario, Canada; Heather Ford, University Health Network, Toronto, Ontario, Canada; Sara Macanovic, University Health Network, Toronto, Ontario, Canada; Ali Rezaeishahreza, University Health Network, Toronto, Ontario, Canada; Junayd Hussain, University Health Network, Toronto, Ontario, Canada; Nathaniel Edwards, University Health Network, Toronto, Ontario, Canada; Rabail Siddiqui, University Health Network, Toronto, Ontario, Canada; Melika Dastgheib, University Health Network, Toronto, Ontario, Canada; Gaaree Chawla, University Health Network, Toronto, Ontario, Canada; Marta Novak, MD, PhD, University Health Network, Toronto, Ontario, Canada; MD, PhD, University Health Network, Toronto, Ontario, Canada

**Aims:** Health-related quality of life (HRQoL) includes domains of physical, mental, emotional, and social functioning. In addition to these generic domains, condition-specific experiences are also captured. Preference-based estimates of overall HRQOL (health utility) are important for cost-utility analyses. Here we assess the association between the EQ5D5L health utility scores and generic and kidney disease specific HRQOL domains in patients with end stage kidney disease (ESKD) treated with maintenance dialysis or kidney transplant. **Methods:** Secondary analysis of cross-sectional data, obtained from adults with ESKD. A convenience sample completed the EQ5D5L, the Social Difficulties Inventory (SDI) and the Kidney Disease Quality of Life-36 (KDQOL36) on electronic data capture. Sociodemographic and clinical characteristics were collected from medical records. The association between the Canadian valuation set scores of the EQ5D5L and disease specific (Symptoms of Kidney Disease—Symptoms) and generic (SF12 Physical Component—SF12PC, SF12 Mental Component—SF12MC and SDI) scores were assessed. We built models with EQ5D5L as dependent variable, sociodemographic and clinical variables as covariables. Each HRQOL domain was added separately to assess the association and improvement of model fit. In a final model all generic domains (including SDI) and one kidney specific domain (Symptoms) were added. **Results:** For the 611 participants (59% male, 44% with diabetes, 40% dialysis, 54% Caucasian, mean[SD] age 56[16] years) the median (interquartile range) scores were: EQ5D5L 0.86 (0.72, 0.93); SF12PC 41 (31, 52); SF12MC 53 (42, 78); Symptoms 83 (73, 94); SDI16 score 5(1–11). When added individually to the multivariable model, each HRQOL domain was significantly associated with EQ5D5L score: coefficient (95% confidence interval): SF12PC 0.008 (0.007, 0.009); SF12MC 0.007 (0.006, 0.009); Symptoms 0.006 (0.005, 0.006) and SDI – 0.015 (– 0.016, – 0.014), respectively;  $p < 0.001$  for all. Adding each domain score resulted in substantial increase in the pseudo R<sup>2</sup> (range: 0.11–0.2). In the model including all HRQOL domains simultaneously, all were significantly associated with EQ5D5L. **Conclusion:** Scores reflecting physical, mental and social domains of generic HRQOL and disease specific domains are all strongly and independently associated with health utility. These results are compatible with the recommendations for multidisciplinary approach to the management of chronic medical conditions.

#### (3022) Item generation and cognitive interviewing for a patient-reported outcome measure of recovery after abdominal surgery

Fateme Rajabiyazdi, McGill University, Montreal, Quebec, Canada; Roshni Alam, McGill University, Montreal, Quebec, Canada; Haley Montgomery, McGill University, Montreal, Quebec, Canada; Charbel El Kefraoui, McGill University, Montreal, Quebec, Canada; Lawrence Lee, McGill University, Montreal, Quebec, Canada; Nancy Mayo, McGill University, Montreal, Quebec, Canada; Liane Feldman, McGill University, Montreal, Quebec, Canada; Julio Fiore Jr., McGill University, Montreal, Quebec, Canada

**Aims:** There is a lack of patient-reported outcome measures (PROMs) with robust measurement properties to assess recovery and support patient-centered care after abdominal surgery. Given this knowledge gap, we initiated a research program to develop a conceptually relevant and psychometrically sound recovery specific PROM. The aim of this study was to generate PROM items reflecting the process of postoperative recovery after abdominal surgery and to ensure patient understanding of the items. **Methods:** We conducted concept-elicitation interviews with patients undergoing abdominal surgery in four countries (Canada, Brazil, Japan, and Italy) to develop an ICF-based conceptual framework of recovery. Items reflecting the essence of each recovery domain were generated through an iterative process of drafting, evaluation, and revision. Items were created based on the statements made by patients during interviews; patient language was preserved as much as possible. Patient understanding of the items was assessed via cognitive debriefing interviews. Patients were asked to provide feedback on the meaning and clarity of the items, the relevance of the response options, and the appropriateness of the recall period. Interviews were recorded and transcribed for analysis and items were modified iteratively according to patient feedback. **Results:** Concept-elicitation interviews were conducted with 30 patients with diverse demographics and surgical characteristics (50% female, age 57 ± 18 years, 66% major or major extended surgery). Thirty-nine domains of recovery emerged from the interviews, 17 related to “Body Functions” and 22 related to “Activities and Participation.” Sixty-two items were generated based on statements made by patients. Two rounds of cognitive debriefing interviews were conducted, each including 12 patients (50% female, age 61 ± 14 years, 70% major or major extended surgery). Cognitive debriefing analyses resulted in the removal of 3 items, modifications of 5 items, and adjustments of one set of response options. A total of 59 items remained for further psychometric testing. **Conclusion:** The items generated in this research provide an essential step towards the development of a novel PROM to support patient-centered care and quality improvement initiatives in abdominal surgery. Rasch analysis will be used to further refine these items, assess the dimensionality structure, and support appropriate scoring.

#### (3023) Using routinely collected patient-reported outcome measures (PROMs) data in evaluating community rehabilitation services in Alberta, Canada

Fatima Al Sayah, University of Alberta, Edmonton, Alberta, Canada; Katie Churchill, Alberta Health Services, Calgary, Alberta, Canada; Lisa Warner, Alberta Health Services, Calgary, Alberta, Canada

**Aims:** There is a growing movement around the world towards the routine use of patient-reported outcome measures (PROMs) within healthcare systems, in an effort to incorporate patients’ perspectives into planning healthcare services delivery and evaluating the

performance and efficiency of the health system. Our aim was to examine routinely collected PROMs (EQ-5D-5L) data of patients undergoing community outpatient and specialized rehabilitation in the province of Alberta, Canada. Methods: Data from 889 patients who had an intake and end of care episode survey between December 2018 and November 2019 were included in this analysis. Results: Half of the patients (53.4%) were seniors ( $\geq 65$  years) and 58.1% were male. At intake, the majority of patients reported problems on at least one EQ-5D-5L dimension; 76% reported mild-extreme problems (levels 2-5) in mobility, 40.7% in self-care, 83.6% in usual activities, 84.3% in pain/discomfort, and 56.7% in anxiety/depression. The mean index score was 0.69 (SD 0.19), and mean VAS score was 64.4 (19.1). From the time of intake until the end of care episode, there was an increase of 6.3% in the proportion of patients reporting no problems on all EQ-5D-5L dimensions (i.e., health state 11111). Additionally, there was an increase in the proportion of patients reporting no problems on each of the EQ-5D-5L dimensions: 17% for mobility, 12.2% for self-care, 18.8% for usual activities, 8.5% for pain/discomfort, and 9.4% for anxiety/depression (Fig. 2). By the end of the care episode, there was an increase of 0.09 (SD 0.16) in the EQ-5D-5L index score (effect size = 0.5), and 9.7 (SD 18.7) in the VAS score (effect size = 0.5). The magnitude of change in these parameters was moderate; however, both reached the minimal important difference thresholds. Conclusion: Levels of problems in this patient population are much higher than those reported by the Alberta general population. Also, the index and VAS scores are much lower than those for the general population. Despite changes in EQ-5D-5L, the health status of patients at the end of care episode was still much lower than that of the general population in all dimensions.

#### **(3024) “I did not take to all the instructions”: Exploring reasons for delay in seeking medical help with diabetic Foot Problems**

Michael Ogunlana, PhD, Federal Medical Centre Abeokuta & University of KwaZulu Natal Durban South Africa, Abeokuta and Durban, Nigeria; Pragashnie Govender, PhD Occupational Therapy, University of KwaZulu-Natal, Westville, Durban, South Africa; Olufemi Oyewole, PhD Physiotherapy, Olabisi Onabanjo University Teaching Hospital Sagamu, Sagamu Ogun State, Nigeria; Adesola Odole, PhD Physiotherapy, University of Ibadan, Ibadan, Oyo State, Nigeria; Love Falola, BMR(PT), Federal Medical Centre Abeokuta, Abeokuta, Nigeria; Olubiyi Adesina, MBChB, Federal Medical Centre Abeokuta, Abeokuta, Nigeria; Jabez Akindipe, MBBS, Federal Medical Centre Abeokuta, Abeokuta, Nigeria

Aims: Delay in reporting foot symptoms in patients with diabetes to health-care professionals is said to be responsible for limb amputation. While reasons for these delays have been investigated elsewhere, they are not well documented in Nigeria. This study explored the causes of delayed presentation in a Nigerian sample of patients with diabetic foot ulcers. Methods: The study followed a descriptive phenomenological qualitative design in which the lived experience of eight participants with diabetes was explored. The participants completed in-depth interviews which were digitally audio-recorded and transcribed verbatim. Data were analyzed thematically using deductive reasoning. Results: The study identified four themes which included knowledge and awareness of foot challenges, risk perception, health seeking triggers and behaviors and competing priority as the factors responsible for delay in presentation of diabetic foot complications. Conclusion: Limited knowledge and awareness and negative health seeking behaviors including self-management and consultation of traditionalists were the major reasons for delays.

#### **(3025) Patient-generated and reported wound images for surgical site infection outcome assessment: feasibility of a method for use in clinical practice, research and audit**

Rhiannon Macefield, PhD, University of Bristol, Bristol, United Kingdom; Kerry Avery, University of Bristol, Bristol, United Kingdom; Jonathan Rees, University of Bristol, Bristol, United Kingdom; Anne Pullyblank, North Bristol NHS Trust, Bristol, United Kingdom; Barnaby Reeves, University of Bristol, Bristol, United Kingdom; Anni Skilton, University of Bristol, Bristol, United Kingdom; Jane Blazeby, University of Bristol, Bristol, United Kingdom

Aims: Wound infection after surgery (surgical site infection; SSI) can result in substantial patient morbidity and health service cost. SSI is an important outcome in research and routine clinical practice, but accurate assessment is challenging because problems often occur after hospital discharge. Patient-generated images of wounds may be valuable to supplement other patient-reported data to identify SSI remotely, minimizing the need for face-to-face follow-up, reducing costs and facilitating blinded outcome assessment. The aim of this study was to develop and evaluate the feasibility, usability and acceptability of a method for patients to take and transmit a standardized wound image after hospital discharge using their own mobile device. Methods: A review of wound-photography literature informed the development of photography instructions for patients. Existing documents ( $n = 11$ ; clinical photography guidelines, trial protocols) were purposefully sampled and key features for taking standardized wound images extracted. Existing software was adapted to design a secure process (web-based survey with image upload) for transmitting images. Cognitive interviews with patients ( $n = 16$ ) were conducted to pre-test and refine the photography instructions/process for transmitting images. Feasibility, usability, and acceptability were explored with a larger group of patients ( $n = 89$ ) field-testing the method remotely, including follow-up telephone interviews. Image quality was examined by three independent clinical assessors. Results: 21 key features (e.g., lighting, camera angle) were identified and informed provisional photography instructions. Three iterations to the instructions/process for transmitting images during pre-testing improved understanding and ease of use. During field-testing, 52/89 (58.4%) participants took an image of their wound(s). Of these, 46 (88.5%) successfully transmitted images. Most common reasons for not taking/transmitting images included further health problems, not having time or no longer being interested in participating ( $n = 11$ ; 12.4%). Problems relating to usability (e.g., technical/competency issues) were reported by a minority ( $n = 4$ ; 4.5%). Some 87/102 (85.3%) images were judged as sufficient to assess the wound for SSI by at least two of the three assessors. Conclusion: Findings demonstrated a method for obtaining patient-generated and reported images for SSI outcome assessment is feasible, usable and acceptable and produces high quality images. Further evaluation of the method in a clinical trial or routine surgical follow-up is now warranted.

#### **(3026) Health-related quality of life and academic performance among undergraduate students at a Colombian University**

Martha Rodriguez, MSc, Universidad Santo Tomás, Bucaramanga, Colombia; Valentina Navarro, DDS, Universidad Santo Tomás, Bucaramanga, Colombia; Juliana Niz, DDS, Universidad Santo Tomás, Bucaramanga, Colombia; Pamela Isaza, DDS, Universidad Santo Tomás, Bucaramanga, Colombia; Silvia J Rueda, DDS, Universidad Santo Tomás, Bucaramanga, Colombia

**Aims:** To establish Health-Related Quality of Life (HRQL) among Health (dentistry, optometry, physical education, and dental laboratory technology) and Law undergraduate students at a Colombian University and its relation to their academic performance. **Methods:** An observational cross-sectional study was conducted with 679 legal age students who were chosen by simple random sampling. All of them filled out the WHOQOL-BREF questionnaire. The variables analyzed were the four domains (Physical and Psychological Health, Social Relations, and Environment), socio-demographic characteristics (sex, age, socio-economic condition, marital status, and presence of illness), and academic performance (period grade average). The four domains were treated as continuous variables, frequencies, percentages, ranges, means, and standard deviations were obtained. A *t* test was used to compare de mean scores, and multiple linear regression was conducted. A  $p < 0.05$  was considered statistically significant. **Results:** The participants' median age was 20 years (range 19–22 years) and 426 (62.7%) were female. There was no statistically significant difference in age by sex ( $p = 0.1519$ ). Less than half (42.6%) of the students rated their HRQOL as “very good,” and 274 (40.4%) were “very satisfied” with their health. In the physical and psychological health domains, the variables that had the greatest influence were doing physical activity ( $\beta$ : 4.5 95% CI 2.3–6.7/ $\beta$ : 4.1 95% CI 1.6–6.5), suffering a disease ( $\beta$ : - 7.7 95% CI - 10.7 to - 4.6/ $\beta$ : - 7.2 95% CI: - 10.6 to - 3.8), and having a good semester grade average ( $\beta$ : 5.3 95% CI 3.2–7.4/ $\beta$ : 4.8 95% CI 2.5–7.2). In the social relations and environmental domains, the greatest influence observed was having a high socio-economic condition ( $\beta$ : 1.7 95% CI 0.2–3.2/ $\beta$ : 5.2 95% CI 4.1–6.2), doing physical activity ( $\beta$ : 3.5 95% CI 0.4–6.5/ $\beta$ : 2.6 95% CI 0.4–4.7), and having a good semester grade average ( $\beta$ : 4.5 95% CI 1.6–7.3/ $\beta$ : 4.0 95% CI 2.0–6.0). **Conclusion:** Being male was related to the Psychological Health and Environment domains. Doing physical activity had an important influence on HRQOL, this result is consistent with other reports.

### (3027) Symptom experience and content validity of the Psoriasis Symptom Scale (PSS) in patients with general pustular psoriasis (GPP) and palmoplantar pustulosis (PPP)

Anne Rentz, Evidera, Kailua, Hawaii, United States; Anne Skalicky, Evidera, Bethesda, Maryland, United States; Dirk Esser, Boehringer Ingelheim GmbH, Ingelheim am Rhein, Germany; Christian Thoma, Boehringer Ingelheim International GmbH, Biberach, Germany; Tristan Glode, MD, Boehringer Ingelheim International GmbH, Ingelheim, Germany

**Aims:** To understand key symptoms of generalized pustular psoriasis (GPP) and palmoplantar pustulosis (PPP) and to confirm the relevance and content validity of the Psoriasis Symptom Scale (PSS) in GPP and PPP. **Methods:** A literature review, clinical expert interviews, and patient interviews were conducted to determine disease-specific symptoms important to patients with GPP and PPP. Combined concept elicitation and cognitive interviews with adults who met the study eligibility criteria were conducted in person and by telephone. **Results:** Seven (27%) participants had a GPP diagnosis, 19 (73%) participants had a PPP diagnosis based on clinician verification (one patient had GPP involvement on palms and soles and is reported in both samples). The median age of study participants was 55.7 years (range 27–72). Most were female ( $n = 21$ , 81%), not Hispanic or Latino ( $n = 25$ , 96%), and White ( $n = 20$ , 77%); 7 (27%). Thirty-nine percent of the respondents reported their symptoms as moderate ( $n = 10$ ), 23% reported severe ( $n = 6$ ), and 12% reported very severe ( $n = 3$ ). During concept elicitation, both GPP and PPP participants indicated that pustules are the underlying cause of their symptoms. Frequently reported GPP symptoms or signs were pain, redness, discomfort, and inflammation/swelling ( $n = 7$ , 100%), followed by

itching, burning, irritation, dryness/dry skin, and soreness ( $n = 6$ , 86%), flaky/peeling skin ( $n = 4$ , 57%), and fissures/cracks ( $n = 2$ , 29%). Frequently experienced PPP symptoms or signs were redness, itching and discomfort ( $n = 19$ , 100%), pain ( $n = 18$ , 95%), burning and irritation ( $n = 17$ , 89%), flaky/peeling skin ( $n = 16$ , 84%), and inflammation/swelling ( $n = 15$ , 79%). The symptoms included in the PSS—burning, itch, pain, and redness—were considered important to GPP and PPP patients. The symptoms best reported by GPP and PPP patients are pain (i.e., pain, discomfort, soreness), itching, and burning. Given the complexity of patient descriptions of redness, redness should be reported by patients and also assessed by clinicians. **Conclusion:** Participants provided positive feedback on the PSS instrument and found the measure to be relevant, straightforward, and easy to understand. **Results:** from this qualitative study provide support for the content validity of the PSS for use as clinical trial endpoints among patients with GPP and PPP.

### (3028) Health-related quality of life of health professionals in a Nigerian tertiary health institution: a cross-sectional survey

Marufat Olujemisi Odetunde, Ph.D, Obafemi Awolowo University, Ile-Ife, Nigeria; Benjamin Ajibola, BMR (PT), Obafemi Awolowo University, Ile Ife, Nigeria; Nurain Odetunde, MBBS, General Hospital, Gusau, Zamfara State, Nigeria; Adaobi Okonji, M.Sc, Obafemi Awolowo University, Ile Ife, Nigeria

**Aims:** Health professionals (HP) are a vulnerable group of workers who are frequently exposed to high number of hazards at the workplace, thus their health-related quality of life (HRQOL) is an important factor that determines the quality of health services they provide. **Methods:** This was a cross-sectional survey of 385 HP purposively recruited from the Obafemi Awolowo University Teaching Hospital (OAUTH), Ile-Ife, Nigeria. All respondents completed the Medical Outcomes Study short form (SF-36) questionnaire, comprising eight subscales and two summary scores. Also, socio-demographic information on the respondents was obtained. Descriptive statistics of mean, standard deviation, frequency and percentages; and inferential statistics of independent t-test and one way ANOVA was used to analyze data. Alpha level was set at  $p < 0.05$ . **Results:** A total of the 385 health professionals (273 males and 112 females) consented for this survey. Majority of the respondents were in the 20–39 years age group (80%), married (52.8%) and hold bachelor's degree (65.7%). Respondents scored well above average in all but two SF-36 subscales; with the highest score on Role physical ( $91.41 \pm 21.95$ ) and lowest on vitality ( $49.72 \pm 15.22$ ) subscales ( $p < 0.05$ ). The Physical Component Score (PCS) ( $76.19 \pm 11.1$ ) was higher than the Mental Component Score (MCS) ( $63.93 \pm 8.8$ ). Male respondents scored higher on three subscales and MCS; and single respondents on five subscales and MCS ( $p < 0.05$ ). There were significant differences across each of age groups and academic qualifications on six subscales ( $p < 0.01$ ). Respondents with Master's degree scored highest on three subscales and PCS ( $p < 0.001$ ), those between 11 and 20 years in practice scored highest on two; and MCS ( $p < 0.05$ ); while those on lower cadre scored lowest on five; and MCS. Significant differences also exist among the HP on all SF 36 subscales; PCS and MCS ( $p < 0.01$ ). Physiotherapists scored highest on four subscales, PCS and MCS; while Medical Laboratory Scientists scored lowest on three ( $p < 0.01$ ). **Conclusion:** Considerable inequalities in HRQOL exist among the HP in OAUTH. These inequalities are gender biased in favor of male HP, those who are single and those with post graduate degree. There is need for initiatives to provide necessary support to the HP and encourage post graduate training.

## Cancer Clinical Care

### (3029) Development of an E-Learning Course on EORTC Patient-Reported Outcome Measures in Clinical Practice: Rationale and study design of Phase I content generation

Heike Schmidt, Institute of Health and Nursing Science and University Clinic and Outpatient Clinic for Radiotherapy, Medical Faculty, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany; Lisa M. Wintner, University Hospital for Psychiatry II Medical University of Innsbruck, Innsbruck, Austria; Sigrid Roggendorf, Institute of Health and Nursing Science, Medical Faculty, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany; Thomas Nordhausen, Institute of Health and Nursing Science, Medical Faculty, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany; Timo Kachel, University Hospital for Psychiatry II Medical University of Innsbruck, Innsbruck, Austria; Linda Dirven, Leiden University Medical Center, The Netherlands and Haaglanden Medical Center, Innsbruck, Netherlands; Martin J.B. Taphoorn, Leiden University Medical Center, The Netherlands and Haaglanden Medical Center, The Hague, Netherlands; Irma M. Verdonck- de Leeuw, Amsterdam University Medical Centers (location VUmc), Amsterdam, Netherlands; Galina Velikova, University of Leeds, Leeds, United Kingdom; Andrew Bottomley, European Organisation for Research and Treatment of Cancer (EORTC), Quality of Life Department, Brussels, Belgium; Dagmara Kuli?, European Organisation for Research and Treatment of Cancer (EORTC), Quality of Life Department, Brussels, Belgium; Dagmara Kuli?, European Organisation for Research and Treatment of Cancer (EORTC), Quality of Life Department, Brussels, Belgium; Monika Sztankay, University Hospital for Psychiatry II Medical University of Innsbruck, Innsbruck, Austria; submitted on behalf of the EORTC Quality of Life Group

**Aims:** Integrating patient-reported outcomes (PROs) into clinical practice can optimize symptom management and identification of patients who are in need for supportive care, thereby increasing the quality of patient-centered care. However, broad implementation into clinical routine remains a challenge. One barrier is a scarcity of materials for training of health care professionals (HCPs) on the use of PRO measures in clinical care. E-learning has shown to be effective in optimizing knowledge, competence and behavior, with the added advantage of individually plannable access and utilization. Hence, we present the study design for a project that aims to provide the groundwork for the development of a specialized e-learning course on implementing PRO measures in routine clinical practice. **Methods:** Development of the e-learning course content follows a participatory approach in a stepwise mixed-methods design. Based on a systematic literature review on clinical use of PRO measures in cancer care and available (online) training concepts, semi-structured interviews with HCPs (physicians, nurses and allied health professions), IT specialists and patient representatives are conducted to explore educational needs and preferences regarding content and teaching methods. A convenience sampling strategy is applied to recruit interviewees. Qualitative results inform a subsequent online survey to substantiate and extend the results by addressing a greater number of potential users. The survey is distributed through the EORTC, EORTC disease-oriented groups, collaborating professional organizations in Europe and scientific networks internationally. **Results:** Literature search showed that although the demand for e-learning courses is rising, there is still a lack of scientifically developed courses addressing the complexity of the implementation of PROs in clinical routine and the specific educational needs of multi-professional users. The chosen methodology takes these factors into account, thus providing the groundwork for the development of the first comprehensive e-learning course addressing implementation of PROs in clinical practice. **Conclusion:** Results: of these preliminary

steps will inform content, structure and methods for the e-learning course. The preliminary course will be pretested regarding usability and then pilot-tested with an inter-professional convenience sample of HCPs.

### (3030) Creating User-Relevant Content for an E-Learning Course on EORTC Patient-Reported Outcome Measures in Clinical Practice: Results of the Qualitative Analyses of Stakeholder Perspectives

Monika Sztankay, PhD, Medical University of Innsbruck, Department of Psychiatry, Psychotherapy and Psychosomatics; University Hospital of Psychiatry II, Innsbruck, Austria; Lisa M. Wintner, Medical University of Innsbruck, Innsbruck, Austria; Sigrid Roggendorf, Institute of Health and Nursing Science, Medical Faculty, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany; Thomas Nordhausen, Institute of Health and Nursing Science, Medical Faculty, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany; Timo Kachel, Medical University of Innsbruck, Department of Psychiatry, Psychotherapy and Psychosomatics; University Hospital of Psychiatry II, Innsbruck, Austria; Linda Dirven, Leiden University Medical Center, The Netherlands and Haaglanden Medical Center, Innsbruck, Netherlands; Martin J.B. Taphoorn, Leiden University Medical Center, The Netherlands and Haaglanden Medical Center, Leiden, The Hague, Netherlands; Irma M. Verdonck- de Leeuw, Amsterdam University Medical Centers (location VUmc), Amsterdam, Netherlands; Galina Velikova, University of Leeds, Leeds, United Kingdom; Andrew Bottomley, European Organisation for Research and Treatment of Cancer (EORTC), Quality of Life Department, Brussels, Belgium; Dagmara Kulis, European Organisation for Research and Treatment of Cancer (EORTC), Quality of Life Department, Brussels, Belgium; Dagmara Kulis, European Organisation for Research and Treatment of Cancer (EORTC), Quality of Life Department, Brussels, Belgium; Heike Schmidt, Institute of Health and Nursing Science and University Clinic and Outpatient Clinic for Radiotherapy, Medical Faculty, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany; submitted on behalf of the EORTC Quality of Life Group

**Aims:** Key components for a successful implementation of patient-reported outcome measures (PROs) in clinical routine are knowledge and competences regarding their use. Development of an e-learning course for health care professionals (HCPs) on the utilization of EORTC PRO measures in clinical practice is currently ongoing, aiming to facilitate implementation. The aim of this qualitative study was to obtain insight into future users' educational needs for the e-learning course. **Methods:** A systematic literature review on the clinical use of PROs informed the qualitative interview guideline. To cover a broad range of potential users and capture a variety of individual perspectives, HCPs (physicians, nurses, allied health professions), IT specialists and patient representatives were recruited for semi-structured interviews applying a convenience sampling strategy using the scientific networks of the investigators. Data analysis followed the rules of qualitative content analysis. **Results:** Seventy-three interviews were conducted with participants from Austria ( $n = 20$ ), The Netherlands ( $n = 21$ ), Germany ( $n = 20$ ), UK ( $n = 7$ ), France, Norway, Denmark, Israel, Malaysia ( $n = 1$  each). Qualitative content analysis revealed the interviewees' interest in the following topics: Basic information on PROs in clinical routine ("definition of the PRO concept," "available PRO instruments"); benefit and clinical use of PRO assessments ("interpretation of scores," "integration of PROs into the medical consultation," "PROs guiding communication and care"); planning of implementation of PRO assessments including barriers and facilitators, clinical workflow and technical issues. Regarding the attitude of HCPs, the use of PROs

in clinical routine was associated with “a change of culture,” a “holistic approach” and “engaging the patient.” The importance of “motivating examples,” electronic data assessment and integration of PRO systems into the medical records was highlighted. The e-learning course should preferably have a modular structure. Conclusion: Data imply HCPs’ need for standardization of clinical procedures of PRO assessment. Educational needs regarding knowledge and competences about the integration of PROs in communication and clinical care and implementation issues became evident, revealing varying perspectives across professions. Complementing these insights on stakeholders’ perspectives an online survey with a broader sample will inform the development of the e-learning course. Subsequent pilot-testing regarding usability, acceptance and perceived benefit is planned.

### **(3031) Integrating patients’ quality of life (QoL) into clinical practice: a Delphi method-based consensus among French physicians managing lung cancer patients**

Virginie Westeel, Unité de Méthodologie et de Qualité de Vie en Cancérologie, Centre Hospitalier Universitaire de Besançon, UMR1098, Université de Franche-Comté, Besançon, France; Marianne Bourdon, Institut de Cancérologie de l’Ouest; UMR INSERM 1246 SPHERE, Université de Nantes, Université de Tours, Nantes, France; Alexis Cortot, Université de Lille, CHU Lille, F-59000, Lille, France; Didier Debievre, Centre Hospitalier Emile Muller, Mulhouse, France; Anne-claire Toffart, Centre Hospitalier Universitaire de Grenoble Alpes, Grenoble, France; Michael Acquadro, Patient Centred Outcomes, ICON plc, Grenoble, France; Benoit Arnould, Patient Centred Outcomes, ICON plc, Lyon, France; Jérémy Lambert, Patient Centred Outcomes, ICON plc, Lyon, France; François-Emery Cotte, Bristol-Myers Squibb, Rueil-Malmaison, France; Anne-Françoise Gaudin, Bristol-Myers Squibb, Rueil-Malmaison, France; Hervé Lemasson, Bristol-Myers Squibb, Rueil-Malmaison, France

**Aims:** QoL assessment has become standard in oncology clinical trials, while its management in routine practice remains subject to many questions. This study aimed to reach a consensus from physicians involved in lung cancer on the patients’ QoL management (assessment and discussion) in daily practice. **Methods:** 747 physicians involved in lung cancer (oncologists, pulmonologists, radiotherapists) were solicited to take part in a Delphi-method-based consensus approach. Based on 3 rounds of iterated queries, this explored 7 QoL management aspects (from specification to assessment ways). Consensus was obtained in the case of 70% responders’ agreement. A scientific committee composed of clinicians and a psycho-oncologist analyzed results following each round. **Results:** A representative panel of 60 physicians (13 oncologists, 43 pulmonologists, 4 radiotherapists) participated in at least one round (53 at round 1, 46 at round 2, 39 at round 3). Consensus elements were reached for 6 aspects. Consensus was obtained for a QoL management all along the patient journey. Three key time points were identified: “diagnosis,” “tumor evaluation showing progressive disease or start of a new treatment” and “palliative and end of life care.” A consensus was reached for a multidimensional QoL discussion with specificities at particular points such as spirituality in palliative care. QoL discussion must occur mainly during routine visits or hospitalization. The need to involve patient’s relatives at all time points (except when discussing side effects) and for a relay by a multidisciplinary team beyond this discussion were consensually recognized.

QoL assessment before the visit could be of interest, however its systematization for all patients at all time points was not consensual. No consensus was reached on the type of tool (interview guide, questionnaire) needed to support the QoL assessment. **Conclusion:** QoL was considered by French physicians as a part of routine clinical visits in thoracic oncology, and was deemed key in the patient-physician interaction. Further work should be conducted to harmonize how to best implement and use QoL assessment.

### **(3033) Comorbidity and risk of reoperation or mortality within 90 days of esophageal cancer surgery**

Zhao Cheng, Karolinska Institutet, Solna, Sweden; Asif Johar, Karolinska Institutet, Stockholm, Sweden; Eivind Gottlieb, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden; Magnus Nilsson, Karolinska University Hospital, Stockholm, Sweden; Jesper Lagergren, Karolinska Institutet, Stockholm, Sweden; Pernilla Lagergren, Karolinska Institutet, Stockholm, Sweden

**Aims:** To assess the influence of preoperative comorbidity on the risk of reoperation or mortality within 90 days of surgery for esophageal cancer. **Methods:** This population-based and nationwide cohort study enrolled 98% of all patients who underwent esophageal cancer surgery in 1987–2010 in Sweden. Log-binomial regression models provided risk ratios (RRs) with 95% confidence intervals (CIs) to estimate associations between comorbidity and risk of reoperation or death within 90 days of esophagectomy. The RRs were adjusted for age, sex, education level, pathological tumor stage, neoadjuvant therapy, surgeon volume, tumor histology, and calendar period of surgery. **Results:** Among 1820 patients, 354 (19.5%) underwent reoperation or died within 90 days of esophagectomy. Patients with Charlson comorbidity index  $\geq 2$  had an increased risk of reoperation or death compared with those with comorbidity index 0 (RR = 1.53; 95% CI 1.21–1.94), and the risk increased by 21% for each point increase of the index (RR = 1.21; 95% CI 1.11–1.31). The RRs were increased in patients with pulmonary diseases (RR = 1.60; 95% CI 1.27–2.02), cardiac diseases (RR = 1.35; 95% CI 1.04–1.77) and diabetes (RR = 1.47; 95% CI 1.07–2.02). The associations were stronger when low-volume surgeons conducted the surgery and in earlier calendar periods. **Conclusion:** Comorbidity in general, and pulmonary diseases, cardiac diseases and diabetes in particular, seems to increase the risk of reoperation or death within 90 days of esophageal cancer surgery. This highlights the value of tailored patient selection, preoperative preparation, and postoperative follow-up.

### **(3034) Does the long-term experience of rectal cancer survivors differ from the wider experience of colorectal cancer survivors? A qualitative review**

Stephanie Philpott, MSc, IQVIA, Reading, United Kingdom; Laurie Batchelder, MSc PhD, IQVIA, Reading, United Kingdom; Ana Maria Rodriguez, PhD MSc BScPT BSc, IQVIA, Madrid, Spain

**Aims:** Colorectal cancer (CRC) is an umbrella term used to describe cancers located in the start of the colon to the end of the rectum. Colon (CC) and rectal cancer (RC) are often grouped together because of similar features and treatments. The long-term impact of CC and RC are also often investigated together as long-term impacts of CRC. The aim of this qualitative review is to investigate the long-

term experiences of RC survivors specifically and how these differ from the wider long-term experiences of CRC survivors. **Methods:** Two online database searches (PubMed and Google Scholar) were conducted. Search 1 identified qualitative articles referencing ‘colorectal’ cancer survivorship, whereas Search 2 identified qualitative articles referencing ‘rectal’ cancer survivorship. Qualitative themes arising from both searches were collated and thematically explored. **Results:** Search 1 yielded 29 articles, of which 14 were included, Search 2 yielded 14 articles, of which 9 were included. Descriptions of physical impacts such as poor bowel function, psychological impacts such as anxiety around bowel function, and impacts on social life and self-care strategies were similar in papers identified through both searches. Descriptions of sexual dysfunction differed. In Search 1 (broad CRC population), sexual dysfunction was generally about avoidance due to concerns around engaging in sexual activity with an ostomy and body image concerns. In Search 2 (RC population), sexual dysfunction was more about physical inability to engage due to long-lasting vaginal stenosis and unhealed radiation burns. This was described as potentially resulting from different surgical procedures—either from an ostomy procedure or pelvic radiation. **Conclusion:** While combining CC and RC under the umbrella of CRC is largely justified to understand impacts of the cancers beyond the acute phase, there may be some important differences between RC and other CRC patients which may not be clearly seen when only looking in general heterogeneous CRC populations. For example, women with RC who undergo pelvic radiation therapy are more likely to experience inability to engage in sexual activities. Oncologists may need to engage in different conversations to set expectations about treatment with RC and CC patients.

### **(3035) Postoperative complications and cancer-related fatigue for survivors 1 year after oesophageal cancer surgery in Sweden**

Zhao Cheng, Karolinska Institutet, Solna, Sweden; Magnus Nilsson, Karolinska University Hospital, Stockholm, Sweden; Asif Johar, Karolinska Institutet, Stockholm, Sweden; Pernilla Lagergren, Karolinska Institutet, Stockholm, Sweden

**Aims:** To clarify the prevalence and characteristics of cancer-related fatigue at 1 year after oesophagectomy in Sweden, and to examine the impact of complications on cancer-related fatigue and the trajectory of cancer-related fatigue from 1 year to 2 year postoperatively. **Methods:** This nationwide cohort study enrolled surgically treated esophageal cancer survivors in Sweden from 2013 onwards. Cancer-related fatigue is measured 1, 1½, and 2 year after surgery. The fatigue scale of the European Organization for Research and Treatment of Cancer Quality of Life Fatigue Module (EORTC QLQ-FA12) questionnaire is used to assess fatigue. Repeated measure mixed effect models will be used to calculate the mean difference of fatigue scores with 95% confidence intervals (CIs) between patients with and without complications, adjusting for age at surgery, sex, education years, pathological tumor stage, neoadjuvant therapy, Charlson comorbidity index, tumor histological type, body mass index. **Results:** In total 248 patients were enrolled. Fatigue scores were 19.6 (SD 18.2) at 1 year, 23.7 (20.8) at 1½ year, and 24.7 (22.2) at 2 year postoperatively. The mean physical fatigue subscale score was highest [31.1 (24.7)], followed by emotional [18.9 (24.7)] and cognitive subscale [9.5 (18.7), 26.3]. Patients with higher complication burden had more fatigue at 1, 1½, and 2 year after oesophagectomy. Similar patterns were also seen in the three subscales. **Conclusion:** Postoperative complications are associated with a higher level of

persistent cancer-related fatigue up to 2 years after esophageal cancer surgery.

### **(3036) A Discrete Choice Experiment (DCE) to assess cancer patient preferences for the features of an assessment tool for Chemotherapy Induced Peripheral Neuropathy (CIPN)**

Alice Yu, PhD candidate, University of Technology Sydney, Sydney, Australia; Deborah Street, University of Technology Sydney, Sydney, Australia; Rosalie Viney, University of Technology Sydney, Sydney, Australia; Stephen Goodall, University of Technology Sydney, Sydney, Australia; Alison Pearce, University of Sydney, Sydney, Australia; Philip Haywood, University of Technology Sydney, Sydney, Australia; Marion Haas, University of Technology Sydney, Sydney, Australia; Hannah Timmins, University of Sydney, Sydney, Australia; Susanna Park, University of Sydney, Sydney, Australia

**Aims:** Up to 40% of cancer patients experience damage to the peripheral nerves as a result of chemotherapy treatment, commonly known as CIPN. In addition, this side-effect can persist years after chemotherapy treatment has ended. However, an assessment for CIPN is not routinely carried out, despite the existence of many assessment tools within the literature. To assist in deciding on a ‘best practice’ CIPN assessment tool, it is important to consider what is relevant and meaningful to both the clinician and the patient. However, there is little to no information in the literature on patient preferences for such assessments. This study aims to address this gap by identifying the features of a CIPN assessment tool that cancer patients’ value. **Methods:** An online DCE was implemented. Respondents were randomly assigned to a block of 8 choice sets to complete. In each choice set, respondents were asked to choose between two hypothetical CIPN assessment tools, described in terms of 6 attributes or features. DCE results were analyzed with a multinomial logit model and a mixed logit model. **Results:** The survey was completed by 117 current and former cancer patients. Respondents were strongly against a situation where the doctor, alone, decides how results of the CIPN assessment influence their care and especially their chemotherapy treatment. Respondents preferred an assessment tool that asks about how CIPN symptoms impact on their ability to perform usual activities. They were also concerned about small changes in their CIPN regardless of whether it is clinically meaningful or not. The inclusion of a physical test of some type was positively received, and respondents did not mind having their usual clinic time extended by up a half hour to accommodate a CIPN assessment. **Conclusion:** Patients desire shared decision-making when it comes to the impact of CIPN assessment results on their general care and especially their chemotherapy treatment. Clinicians should take this into account when evaluating assessment results. The findings from this DCE will also assist clinicians in choosing an assessment tool for CIPN that is satisfactory to both the clinician and the patient.

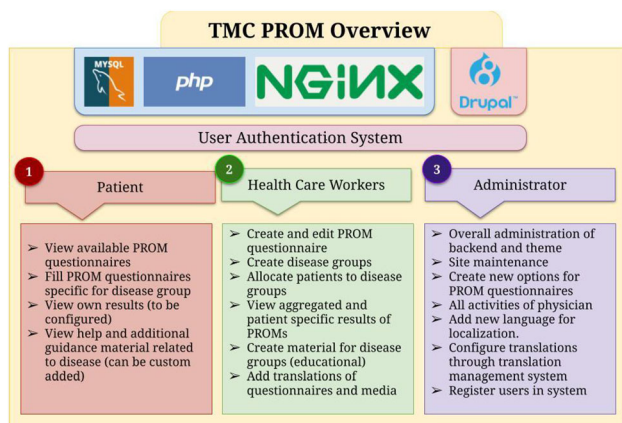
### **(3037) Development and user experience testing of an electronic system for routine collection and use of electronic patient-reported outcome measures**

Santam Chakraborty, MD Radiotherapy, Tata Medical Center, Kolkata, India; Indranil Mallick, MD Radiotherapy, Tata Medical Center, Kolkata, India; Tapeshe Bhattacharyya, MD, Tata Medical Center, Kolkata, India; Moses Arunsingh, MD, FRCR, Tata Medical Center, Kolkata, India; Rimpa Basu Achari, MD, Tata Medical



Center, Kolkata, India; Sanjoy Chatterjee, FRCC, Tata Medical Center, Kolkata, India

**Aims:** Electronic collection of patient-reported outcomes (e-PROM) allows accurate recording of data. They also enable the visualization of longitudinal trends in domain-specific scores for a patient, and may improve patient-physician communication. Several commercial offerings are available but deploying them in countries like India is challenging due to language barriers and literacy levels. Additionally, costs involved remain a major problem. We propose to develop an open-source tool to serve the same purpose. **Methods:** After an exhaustive requirement analysis for a minimum viable product, we decided to proceed with developing the system using an open-source content management system. Additional contributed modules like Webform and Media were used to provide additional functionality. Three tiers of user roles with role-specific privileges were defined. A preliminary user experience testing was done for the patient role. **Results:** All requirements identified in the requirement analysis section have been met. The system allows users with a patient role to fill in questionnaires presented to them. Questionnaires can be translated to the desired language and additional media elements like voice-overs and video can be added to aid the persons filling the questionnaire. To ensure that diverse groups of patients can be targeted with specific questionnaires, patients are grouped according to disease groups. Health care workers can visualize the results of the questionnaire as well as develop new questionnaires using a graphical interface. Of 48 persons who expressed an interest in user experience testing, 31 (64.5%) participated. Most participants were able to complete the allocated tasks in the testing process. Initial user experience testing shows that 93.5% of the users (playing the role of patients) were able to use the website without additional help. **Conclusion:** An open-source system to collect electronic PROM has been developed with localization in Indian languages. We aim to continue developing, validating and extending the system in the future.



### (3038) The electronic assessment of patient-reported outcomes and quality of life in radio-oncology—implementation and process evaluation

Thomas Nordhausen, M.Sc., Institute for Health and Nursing Science, Medical Faculty of the Martin Luther University Halle-Wittenberg, Halle (Saale), Germany; Dirk Vordermark, Prof., Department of Radiation Oncology, University Hospital Halle, Halle (Saale), Germany; Bernhard Holzner, Prof., University Hospital of Psychiatry II, University Hospital Innsbruck, Innsbruck, Austria; Haifa Kathrin Al-Ali, PD Dr., Krukenberg Cancer Center, University Hospital Halle, Halle (Saale), Germany; Gabriele Meyer, Prof., Institute for Health and Nursing Science, Medical Faculty of the Martin Luther

University Halle-Wittenberg, Halle (Saale), Germany; Heike Schmidt, Dr., Department of Radiation Oncology, University Hospital Halle (Saale), Germany, Halle (Saale), Germany

**Aims:** Implementation of Patient-Reported Outcome Measures (PROM) including Health-Related Quality of Life (HRQOL) in clinical routine is still rare although studies have proven their clinical benefit. A successful implementation in clinical routine needs a structured, science-based approach. **Aims:** of the study are development, piloting and evaluation of the implementation of electronic HRQOL assessments in a radio-oncology clinic. **Methods:** 1. Analysis of current practice, assessment of facilitators/barriers of implementation; 2. Multi-professional focus groups to decide on PROMs, assessment times and clinical procedures; 3. Development of face-to-face training and e-learning for clinical staff regarding communication and clinical use of PRO; 4. Stepwise implementation for in- and out-patients; 5. Process optimization and evaluation regarding outcomes related to the success of the implementation. **Results:** Decision on PROMs comprised screening of HRQOL (EORTC QLQ-C30) prior therapy and prior discharge and symptom monitoring during therapy (EORTC single items). Treatment procedures have been developed to ensure guideline-based reaction to reported symptoms. First results of HRQOL screening ( $n = 304$ ) prior therapy revealed moderate or severe symptom burden for fatigue (69.7%), nausea/vomiting (21.4%), pain (55.6%), dyspnea (52.0%), insomnia (66.8%), appetite loss (51.6%), constipation (31.3%), diarrhea (21.7%), financial problems (42.4%) and impact on functioning for physical (54.6%), role (58.9%), emotional (60.9%), cognitive (33.9%) and social function (56.9%) and reduced global quality of life (81.9%). Process evaluation resulted in an average time of 12.3 (SD: 6.8; Min: 1.0; Max: 30.0) minutes for explaining and supporting the patients to complete the assessments. **Conclusion:** PROMs provide relevant information complementing clinical records. Access is facilitated through available languages, interpretation of results is supported by graphical presentation for screening and monitoring including configuration options for comparator and cut-off values. However, the integration of electronic assessment of PROMs into clinical practice is challenging due to the complexity of the intervention. Clearly defined responsibilities, inter-professional training, IT support, and smoothly operating systems as well as clinical role models valuing the benefit of the PROs are needed to keep up motivation and achieve changes of clinical routine and culture. A manual based on the results of the process evaluation will be developed to support future implementation.

### (3039) Practical application of the HM-PRO in clinical practice: views from the trenches

Boris Afanasyev, Raisa Gorbacheva Memorial Research Institute of Children Oncology, Hematology and Transplantation, I. P. Pavlov First Saint Petersburg State Medical University, Ministry of Health of Russia, Saint-Petersburg, Russia; Madina Avtorhanova, Raisa Gorbacheva Memorial Research Institute of Children Oncology, Hematology and Transplantation, I. P. Pavlov First Saint Petersburg State Medical University, Ministry of Health of Russia, Saint-Petersburg, Russia; Anna Bannikova, N.I. Pirogov National Medical Surgical Center, Moscow, Russia; Anastasia Bykova, National Research Center for Hematology, Moscow, Russia; Andrey Chekalov, Raisa Gorbacheva Memorial Research Institute of Children Oncology, Hematology and Transplantation, I. P. Pavlov First Saint Petersburg State Medical University, Ministry of Health of Russia, Saint-Petersburg, Russia; Sergey Gritsaev, Russian Research Institute of Hematology and Transfusiology, Saint-Petersburg, Russia; Ivan Kostroma, Russian Research Institute of Hematology and Transfusiology, Saint-Petersburg, Russia; Alexander Kulagin, Raisa Gorbacheva Memorial Research Institute of Children Oncology,

Hematology and Transplantation, I. P. Pavlov First Saint Petersburg State Medical University, Ministry of Health of Russia, Saint-Petersburg, Russia; Yuriy Kuznetsov, Raisa Gorbacheva Memorial Research Institute of Children Oncology, Hematology and Transplantation, I. P. Pavlov First Saint Petersburg State Medical University, Ministry of Health of Russia, Saint-Petersburg, Russia; Elena Lepik, Raisa Gorbacheva Memorial Research Institute of Children Oncology, Hematology and Transplantation, I. P. Pavlov First Saint Petersburg State Medical University, Ministry of Health of Russia, Saint-Petersburg, Russia; Kirill Lepik, Raisa Gorbacheva Memorial Research Institute of Children Oncology, Hematology and Transplantation, I. P. Pavlov First Saint Petersburg State Medical University, Ministry of Health of Russia, Saint-Petersburg, Russia; Vladimir Melnichenko, N.I. Pirogov National Medical Surgical Center, Moscow, Russia; Natalia Mikhailova, Raisa Gorbacheva Memorial Research Institute of Children Oncology, Hematology and Transplantation, I. P. Pavlov First Saint Petersburg State Medical University, Ministry of Health of Russia, Saint-Petersburg, Russia; Ivan Moiseev, Raisa Gorbacheva Memorial Research Institute of Children Oncology, Hematology and Transplantation, I. P. Pavlov First Saint Petersburg State Medical University, Ministry of Health of Russia, Saint-Petersburg, Russia; Nikita Mochkin, N.I. Pirogov National Medical Surgical Center, Moscow, Russia; Irina Nemchenko, National Research Center for Hematology, Saint-Petersburg, Russia; Tatiana Nikitina, University Hospital, Saint-Petersburg State University, Saint-Petersburg, Russia; Anna Petrova, National Research Center for Hematology, Moscow, Russia; Natalia Porfirieva, Multinational Center for Quality of Life Research, Saint-Petersburg, Russia; Olga Pospelova, National Research Center for Hematology, Moscow, Russia; Anastasia Samoylova, N.I. Pirogov National Medical Surgical Center, Moscow, Russia; Anna Smirnova, Raisa Gorbacheva Memorial Research Institute of Children Oncology, Hematology and Transplantation, I. P. Pavlov First Saint Petersburg State Medical University, Ministry of Health of Russia, Saint-Petersburg, Russia; Elena Smirnova, N.I. Pirogov National Medical Surgical Center, Moscow, Russia; Elnara Tagieva, Raisa Gorbacheva Memorial Research Institute of Children Oncology, Hematology and Transplantation, I. P. Pavlov First Saint Petersburg State Medical University, Ministry of Health of Russia, Saint-Petersburg, Russia; Ivan Tsygankov, Raisa Gorbacheva Memorial Research Institute of Children Oncology, Hematology and Transplantation, I. P. Pavlov First Saint Petersburg State Medical University, Ministry of Health of Russia, Saint-Petersburg, Russia; Anna Turkina, National Research Center for Hematology, Moscow, Russia; Nikita Volkov, Raisa Gorbacheva Memorial Research Institute of Children Oncology, Hematology and Transplantation, I. P. Pavlov First Saint Petersburg State Medical University, Ministry of Health of Russia, Saint-Petersburg, Russia; Irina Zapreeva, Russian Research Institute of Hematology and Transfusiology, Saint-Petersburg, Russia; Anastasia Zherniakova, Russian Research Institute of Hematology and Transfusiology, Saint-Petersburg, Russia; Sam Salek, School of Life and Medical Sciences, University of Hertfordshire, Hatfield, United Kingdom; Esther Olivia, Hematology Unit, Grande Ospedale Metropolitano, Reggio Calabria, Italy; Tatyana Ionova, University Hospital, Saint-Petersburg State University, Saint-Petersburg, Russia

**Aims:** The HM-PRO is a new patient-reported outcomes measure in hematological malignancy (HM) for use in clinical practice. It consists of two scales: Part A measuring the ‘impact on patients’ quality

of life (QoL)’ grouped into 4 domains (physical behavior—PB, social well-being—SW, emotional behaviour—EB, eating and drinking—ED); and Part B—‘signs and symptoms (S&S).’ The aim of the study was to determine the value of HM-PRO in daily clinical practice in Russia. **Methods:** Adult patients with different HMs completed the Russian version of the HM-PRO before seeing the hematologist in an in- or out-patient clinic. Physicians then reviewed patients’ responses during consultation, as well as completed the clinical and demographic information. They also reported about the feasibility of implementing the HM-PRO in clinical practice and its usefulness during treatment decision-making. Descriptive statistics and non-parametric tests were employed with  $p$  value  $< 0.05$ . **Results:** 192 patients from 5 tertiary hospitals were recruited into a cross-sectional study: 93 patients were male; mean age—50.7 ( $\pm 14.8$ ) yrs. Diagnoses were acute myeloid leukemia (43), chronic myeloid leukemia (47), multiple myeloma (22), aggressive non-Hodgkin lymphoma (NHL) (21), Hodgkin lymphoma (21), myelodysplastic syndromes (14), chronic lymphocytic leukemia (10), indolent NHL (7), myeloproliferative neoplasms (4) and others (3). Ten (34.5%) out of 29 physicians were male; mean age was 36 ( $\pm 11$ ) years; mean years of practice—11 ( $\pm 10$ ) years. PB and S&S were significantly worse (higher scores) in patients with progressive disease (Table 1). Physicians reported that the HM-PRO was useful for patient management in 155 cases (81%): 10% at diagnosis, 57% undergoing treatment, 33% at follow-ups. Physicians reported that HM-PRO was useful for all patients with progressive disease and for the majority in remission/stable: PB was more impaired and S&S was more pronounced in progressive disease. All the physicians stated that the use of HM-PRO is of value in clinical practice for managing patients with HMs. **Conclusion:** The HM-PRO is a practical and valuable tool to capture patients’ needs regardless of the stage and state of the disease. The implementation of the HM-PRO in daily clinical practice enhances quality of patient management in HMs.

Table 1 HM-PRO median scores and interquartile ranges

HM-PRO	Remission/Stable (n=178)	Progressive disease (n=14)	p
Part A Total Score	18.19 (9.09–31.36)	24.58 (21.02–36.36)	0.060
Physical behaviour	14.29 (0–35.71)	50.00 (28.57–78.57)	<0.001
Social well-being	0 (0–16.67)	0 (0–16.67)	0.437
Emotional behaviour	27.27 (13.64–45.45)	31.82 (9.09–45.45)	0.603
Eating and drinking	25.00 (0–25.00)	12.50 (0–50.00)	0.893
Part B Total Score	14.71 (5.88–23.53)	27.94 (14.71–35.29)	0.010

## Cancer Research

### (3040) Melanoma patient registries: a targeted review and comparison of datasets

Katie Hall, Vitaccess Ltd, Oxford, United Kingdom; Samuel Llewellyn, Vitaccess Ltd, Oxford, United Kingdom; Catherine Bottomley, Vitaccess Ltd, Oxford, United Kingdom; Mark Larkin, Vitaccess Ltd, Oxford, United Kingdom

**Aims:** Health-care providers, payers and regulators use patient registries as a valuable source of real-world data for evidence-based decision-making. This review aimed to identify melanoma-specific registries and compare the patient data reported within them. **Methods:** Melanoma registries were identified through targeted searches (no date restrictions applied) of the ClinicalTrials.gov, European Network of Centres for Pharmacoepidemiology and Pharmacovigilance, PubMed databases, and Google. Registry publications and websites were assessed for availability of information on location, patient characteristics, diagnosis and disease classification, mortality, treatments and hospitalizations. **Results:** Of 21 identified registries

reporting data for melanoma patients, 15 were active and one in development. Most registries were national-level initiatives (52%), 82% of which were based in Europe. Number of patients varied from 89 to 138,000. Six registries specified trial participation, test, or treatment in their eligibility; of which all were based in the USA or international. Eleven in total specified the type or stage of melanoma of participants. Patient demographics, mortality, disease classification and treatments were the most widely available data (95%, 81%, 76% and 67%, respectively). Less than a third of registries collected data on quality of life, comorbidities, patient-reported outcomes, or hospitalization, and only 10% collected productivity loss data. Only one national registry collected each data variable with no eligibility criteria regarding stage or type of melanoma. Fifty-seven percent of registries were initiated during the last 10 years, of which 45% were industry funded. Eighty-two percent of registries initiated during the last 10 years are still known to be active. Conclusion: As rates of melanoma continue to rise, melanoma patient registries are an important source of real-world information. A rise in the number of industry funded registries during the last 10 years indicates increased interest of stakeholders in the data. There is a paucity in registries collecting quality of life, patient-reported outcomes and productivity loss data, especially those inclusive of all patients with melanoma.

#### **(3041) Symptoms on the first day post surgery predicting major complications in patients with lung cancer**

Qingsong Yu, Postgraduate, School of Public Health and Management, Chongqing Medical University, Chongqing, China; Qiuling Shi, School of Public Health and Management, Chongqing Medical University, Chongqing, China; Wei Xu, School of Public Health and Management, Chongqing Medical University, Chongqing, China; Yang Pu, School of Public Health and Management, Chongqing Medical University, Chongqing, China; Hongfan Yu, School of Public Health and Management, Chongqing Medical University, Chongqing, China; Wei Dai, Department of Thoracic Surgery, Sichuan Cancer Hospital, School of Medicine, University of Electronic Science and Technology of China, Chongqing, China

**Aims:** Postoperative complications are common problems after lung cancer surgery, resulting in delayed recovery, impaired quality of life and shortened survival. We aim to determine symptoms on postoperative day 1 (POD1) that can inform postoperative complications during hospitalization for patients with lung cancer surgery. **Methods:** Data extracted from a prospective cohort study of 512 patients who had received lung cancer surgery (November 2017–November 2019). Symptoms were assessed by M.D. Anderson Symptom Inventory begin with admission and end with 2 years after discharge in patients. Types and grades of complications during hospitalization were recorded. Individual symptoms were tested for a possible association with cardiopulmonary complications by univariate analyses. To established a clinically meaningful thresholds of symptoms to inform the onset of major complications, optimal cutpoints were derived using analysis of calibration method. **Results:** 493(96.3%) patients with complete POD1 MDASI-LC and records on postoperative complications were included in this analysis. Grade 1–4 (Clavien-Dindo grading) complications developed in 119 (24.1%) patients, including pneumonia (5.7%, 28/493), prolonged air leak (> 7 days) (4.7%, 23/493), incision dehiscence (4.5%, 22/493), arrhythmia (3.7%, 18/493), pleural effusion (3.2%, 16/493) and hoarseness (2.4%, 12/493). Univariate logistic regression showed that shortness of breath independently predicted with prolonged air leak (> 7 days), while dry mouth was related to hoarseness. Cutpoints for shortness of breath was 6 and dry mouth was 5. Patients reporting 6–10 in

shortness of breath showed 3.5 times of risk to develop prolonged air leak (OR 3.59, 95% CI 1.48–8.70). Patients reporting 5–10 of dry mouth are more likely to have hoarseness (OR 5.78, 95% CI 1.54–21.64). **Conclusion:** Shortness of breath and dry mouth on the first day after surgery can inform the onset of major complications in patients with lung cancer. Our results suggest that PRO-based symptom assessment with clinically meaningful threshold could alert clinicians for early prevention of postoperative complications.

#### **(3042) Patient-reported erectile and urinary function after repeat biopsies in men on active surveillance. Results of the Movember Foundation's GAP3 cohort**

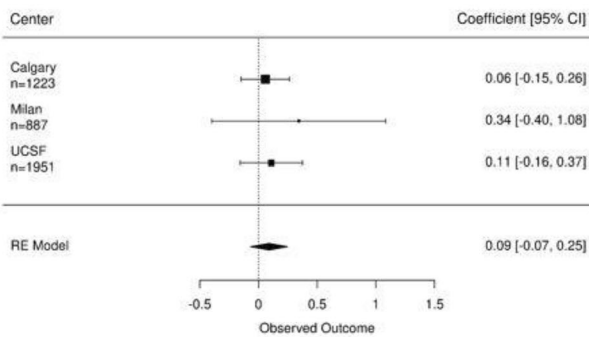
Lionne Venderbos, Department of Urology, Erasmus University Medical Center Rotterdam, Rotterdam, Netherlands; Trafford Crump, Department of Surgery, University of Calgary, Calgary, Alberta, Canada; Sebastiaan Remmers, Department of Urology, Erasmus University Medical Center Rotterdam, Rotterdam, Netherlands; Mieke Van Hemelrijck, King's College London, Faculty of Life Sciences and Medicine, Translational Oncology & Urology Research (TOUR), London, United Kingdom; Jozien Helleman, Department of Urology, Erasmus University Medical Center Rotterdam, Rotterdam, Netherlands; Daan Nieboer, Department of Urology, Erasmus University Medical Center Rotterdam, Rotterdam, Netherlands; Monique Roobol, Department of Urology, Erasmus University Medical Center Rotterdam, Rotterdam, Netherlands; submitted on behalf of Movember Foundation's Global Action Plan Prostate Cancer Active Surveillance (GAP3) consortium

**Aims:** Active surveillance (AS) for low-risk prostate cancer (PCa) is intended to overcome potential side-effects of definitive treatment. Frequent prostate biopsies during AS may, however, impact erectile (EF) and urinary function (UF). The objective of this study was to test the influence of prostate biopsies on patient-reported EF and UF using multicenter data from the largest to-date AS-database. **Methods:** Data analyses were performed using the Movember GAP3 database (v3.2), containing data from 21,169 AS participants from 27 AS-cohorts worldwide. Men were included in the study if they had at least one follow-up prostate biopsy and completed at least one patient-reported outcome measure (PROM) related to EF (Sexual Health Inventory for Men [SHIM]) or UF (International Prostate Symptom Score [IPSS]) during follow-up. The effect of the number of biopsies on either SHIM or IPSS was analyzed using linear mixed models to adjust for clustering at patient level. Analyses were stratified by center; covariates included age and Gleason Grade group at diagnosis, and time on AS. **Results:** 696 men completed the SHIM 3175 times, with a median follow-up of 36 months (IQR 20–55). 845 men completed the IPSS 4061 times, with a median follow-up of 35 months (IQR 19–56). The intraclass correlation was 0.74 for the SHIM and 0.68 for the IPSS, indicating substantial differences between men in their PROMs. Limited heterogeneity between cohorts in the estimated effect of the number of biopsies on either PROM was observed (Fig. 1). A significant relationship was observed between the number of biopsies and the SHIM score, but not for the IPSS score. Every biopsy reduces the SHIM score an average 0.67 (95% CI 0.47–0.88) points. **Conclusion:** The results from this study suggest that repeated biopsies may have a potential detrimental effect on patient-reported EF, but not on UF. However, the observed effect on the SHIM is less than the minimal clinically important difference, so whether this is a meaningful change to patients is not known. Future studies could investigate whether clinicians should consider less invasive

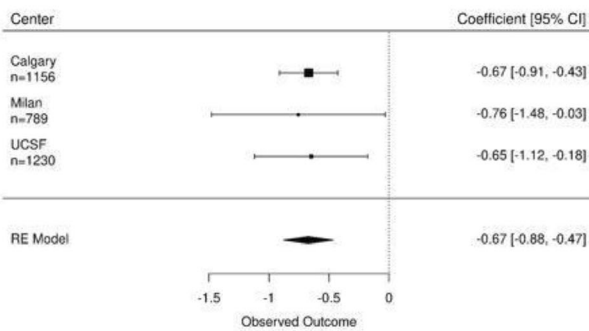
monitoring (e.g., imaging) for men on AS to help reduce potential negative consequences on EF as much as possible.

Figure 1. Result of linear mixed model on the effect of repeat biopsy on UF and EF.

### Urinary functioning



### Erectile functioning



### (3043) Longitudinal changes in sleep quality among premenopausal women with breast cancer

Chloe M. Beverly Hery, MS, The Ohio State University, Columbus, Ohio, United States; Sarah A. Janse, The Ohio State University, Columbus, Ohio, United States; Kimberly J. Van Zee, Memorial Sloan-Kettering Cancer Center, New York City, New York, United States; Elizabeth Z. Naftalis, HealthTexas Community Health Services, Dallas, Texas, United States; Electra D. Paskett, The Ohio State University, Columbus, Ohio, United States; Michelle J. Naughton, The Ohio State University, Columbus, Ohio, United States

**Aims:** Sleep disruptions are common in women with breast cancer, yet longitudinal studies of sleep in premenopausal patients are lacking. We examined trends and predictors of sleep quality over 3 years in a cohort of premenopausal breast cancer patients. **Methods:** 836 women,  $\leq 45$  years old with stage I-III breast cancer, were recruited to the Menstrual Cycle Maintenance and Quality of Life After Breast Cancer Study within 8 months of a breast cancer diagnosis. The Women's Health Initiative Insomnia Rating Scale (WHIIRS), a measure of sleep quality and insomnia symptoms, was completed at 6-month intervals from baseline through 3 years post-recruitment.

Higher WHIIRS scores (range 0–20) indicate more sleep disruption, with scores  $\geq 9$  indicative of a clinical diagnosis of insomnia. Linear mixed-effects models were used to examine sleep quality changes over time. Demographic, clinical, and behavioral factors were evaluated as predictors of sleep quality. Results: Mean age of participants was 38 years ( $\pm 4.9$  years) at enrollment. The mean baseline WHIIRS score was 9.5 ( $\pm 5.2$ ), with 37%–57% reporting sleep disruption consistent with insomnia symptoms (WHIIRS  $\geq 9$ ) across assessments. Sleep quality improved significantly over the 3 year period ( $p < 0.001$ ). Baseline sleep quality was significantly worse compared to all other assessment points ( $p < 0.0001$ ). Significant univariate predictors of poorer sleep quality over time were racial/ethnic minority status ( $p = 0.02$ ), lower education ( $p = 0.03$ ), chemotherapy ( $p = 0.02$ ), hot flashes ( $p < 0.0001$ ), lower physical activity ( $p < 0.01$ ), lower social support ( $p < 0.0001$ ), and greater depressive symptoms ( $p < 0.0001$ ). Older age at diagnosis and higher body mass index (BMI) were also related to poor sleep quality ( $p$  values  $< 0.01$ ), but did not change over time ( $p$  values  $> 0.33$ ). In multivariable models, which included all of the above significant univariable predictors, non-white race ( $p = 0.03$ ), hot flashes ( $p < 0.0001$ ), and higher depressive symptoms ( $p < 0.0001$ ) remained significant. Conclusion: Sleep quality and insomnia symptoms were worse closer to breast cancer diagnosis and treatment, yet poor sleep quality and insomnia symptoms persisted over 3 years for some young women. Survivorship care should include inquiring about and addressing sleep problems, particularly during and immediately after cancer treatment. Implementing early interventions for sleep disruption and providing information on healthy sleep habits may benefit young breast cancer survivors.

### (3044) Exploring the use of wearables in patients with cancer—a review

Cecilie Holländer-Mieritz, MD, Rigshospitalet, Department of Oncology, Copenhagen, Denmark; Ulrikke Lyng Beauchamp, BSc, Rigshospitalet, Department of Oncology, Copenhagen, Denmark; Helle Pappot, Professor, Rigshospitalet, Department of Oncology, Copenhagen, Denmark

**Aims:** The use of wearable sensor technology is expanding. In the health care setting trials investigating wearable devices are slowly increasing. Wearables can be used to monitor variables such as vital signs, sleep and physical activity. The clinical meaning of the wearables outcomes needs to be explored. The purpose of this literature-review was to explore the present use of wearables during cancer treatment in clinical trials. **Methods:** A systematic search of Pubmed and Embase was performed. Inclusion criteria were studies containing a wearable objective measure for cancer patients undergoing active anti-neoplastic treatment. Furthermore, the study had to address adherence to the wearable. Two independent reviewers screened articles for eligibility. Data on study design, type of wearable, adherence and device outcome were recorded. The outcome topics were thematically grouped. **Results:** Our systematic search identified 992 studies of which 25 studies met our inclusion criteria's. The cancer types represented in the studies were breast ( $n = 7$ ), gastrointestinal ( $n = 4$ ), lung ( $n = 4$ ), gynecologic ( $n = 1$ ), nine studies had multiple cancer types. Oncologic treatment was primarily chemotherapy ( $n = 16$ ). Study type distribution was; pilot/feasibility study ( $n = 12$ ), observational study ( $n = 10$ ), randomized controlled trial ( $n = 3$ ). Median sample size were 40 patients (7–180). All studies used a wearable with an accelerometer. The most frequent planned monitoring duration was 8–30 days ( $n = 13$ ). Topics for

wearable outcome were; physical activity ( $n = 18$ ), circadian rhythm ( $n = 8$ ), sleep ( $n = 6$ ), skin temperature ( $n = 2$ ). Sixteen studies also used patient-reported outcomes; quality of life ( $n = 9$ ), physical activity ( $n = 7$ ), mental health ( $n = 7$ ), specific symptom monitoring ( $n = 7$ ), others ( $n = 7$ ). We found that definitions of outcome measures and adherence varied across studies, and no consensus among studies existed on which variables to monitor during treatment. Conclusion: This review provides an overview of the use of wearables devices during cancer treatment. Physical activity was the most used wearable outcome. Better consensus of terms in use and establishing standards for definitions of wearable outcome and adherence would improve comparisons of outcomes from studies using wearables. Research using advanced wearable devices and active use of the data are encouraged to further explore the potential of wearable devices in oncology.

#### **(3045) Characterizing pain reductions associated with opioid prescription in cancer against non-cancer chronic pain patients**

Tito Mendoza, University of Texas MD Anderson Cancer Center, Houston, Texas, United States; Qiuling Shi, University of Texas MD Anderson Cancer Center, Houston, Texas, United States; Natalie Moryl, Memorial Sloan Kettering Cancer Center, New York, New York, United States; Susan Horn, University of Utah School of Medicine, Salt Lake, Utah, United States; Charles Cleeland, University of Texas MD Anderson Cancer Center, Houston, Texas, United States; Charles Inturrisi, Weill Cornell Medical College, Houston, New York, United States

**Aims:** Opioids have long been used to treat acute severe pain and chronic pain in cancer but its use in the treatment of chronic non-cancer pain remains controversial because of effectiveness and safety concerns. We leveraged two existing pain registries to (1) compare cancer versus non-cancer chronic pain patients on how opioid use is associated with pain improvement via their pain severity reports, (2) evaluate whether opioid side effects are more pronounced in patients who respond to opioid therapy, and (3) compare the changes in self-reported symptoms between cancer and non-cancer pain patients. **Methods:** This study was a secondary analysis of the Memorial Sloan Kettering Cancer Center (MSKCC) Pain Registry (cancer,  $n = 444$ ) and the Weill Cornell Medical College (WCMC) Pain Registry (non-cancer,  $n = 432$ ). Patients completed the Brief Pain Inventory (BPI) and the Condensed Memorial Symptom Assessment Scale (CMSAS) at two time points. We defined pain responders as those patients who reported a pain worst score decrease of 2 points or greater. We used either paired t-test or the Wilcoxon Signed Rank Test to examine whether there were significant changes in symptoms and in morphine equivalent (MEQ) dose. **Results:** We found that (101/444) 23% of cancer and (56/432) 13% of non-cancer patients showed improvement in pain. There were significantly more cancer patients with pain improvement compared with non-cancer patients ( $p$  value  $< .0002$ ). Of those patients who are pain responders, we observed no significant changes in morphine equivalent (MEQ) dose regardless of whether they were cancer (182 vs 186 MEQ) or non-cancer patients (131 vs 125 MEQ). However, MEQ prescription levels are significantly higher for cancer patients compared with non-cancer patients either at the first MEQ time point (182 vs 131,  $p < 0.008$ ) or at the 2nd MEQ time point (186 vs 125,  $p < 0.002$ ). Pain improvement was not associated with worsening of constipation or dry mouth. In comparing the non-cancer chronic pain patients against cancer patients, cancer pain responders reported significantly better physical functioning, less constipation, and better sleep when compared with non-cancer pain

responders. Conclusion: Our results support the notion that pain management using opioids is an effective approach for cancer patients.

#### **(3046) Predictive value of health-related quality of life on radiotherapy related toxicities in head and neck cancer patients**

Oumar Billa, Epidemiology and Quality of Life Research Unit, Inserm U1231, Georges Francois Leclerc Centre, Dijon, France; Franck Bonnetain, Methodology and Quality of Life in Oncology Unit, Inserm UMR 1098, University Hospital of Besançon, Dijon, France; Jérôme Chamois, Centre Hospitalier Saint Gregoire, Rennes, France; Angeline Ligey, Centre Hospitalier Fleriat, Bourg-en-Bresse, France; Valérie Ganansia, Paul strauss Cancer Centre-Unicancer, Strasbourg, France; Sophie Renard, Institut de cancérologie de Lorraine, Strasbourg, France; Sophie Maillard, Centre Bourgogne, Lille, France; Magali Quivrin, Radiotherapy department, Georges-François Leclerc Cancer Centre-UNICANCER, Dijon, France; Noémie Vulquin, Radiotherapy department, Georges-François Leclerc Cancer Centre-UNICANCER, Dijon, France; Pierre Truntzer, Paul strauss Cancer Centre-Unicancer, Strasbourg, France; Georges Noel, Paul strauss Cancer Centre-Unicancer, Strasbourg, France; Georges Noel, Paul strauss Cancer Centre-Unicancer, Strasbourg, France; Philippe Maingon, Radiotherapy Unit-Hôpital de la Pitié-Salpêtrière-APHP, Paris, France; Tienhan Sandrine DABAKUYO-YONLI, Epidemiology and Quality of Life Research Unit, Inserm U1231, Georges Francois Leclerc Centre, Dijon, France

**Aims:** To assess the association between baseline health-related quality of life (HRQOL) and occurrence of major toxicities in patients with head and neck cancers (HNC). **Methods:** Patients with HCN treated by radiotherapy were included from May 2009 to September 2014 in 4 centers in France. Patients were asked to complete the EORTC-QLQ-C30, the EORTC-QLQ-H&N35 and the Europol EQ-5D questionnaire at baseline and 1 and 2 years after. Toxicities were assessed by the NCI-CTC- scale version 4. Major toxicities, including skin, digestive, infection, general, hematological and head neck toxicities, were defined as adverse event with a grade  $\geq 2$ . Multivariable cox proportional-hazards regression analysis adjusted on sex, center, stage, type of tumor, age, smoking status, comorbidities, surgery and chemotherapy was used to assess association between baseline global health status[GHS], functional dimensions from EORTC-QLQ-C30 (physical[PF], role [RF], emotional [EF], cognitive [CF], social [SF]) and major toxicities. We defined a difference at least 10 points (on a 0 to 100 scale) in HRQOL as the minimal clinically important change.  $P$ value  $< 0.01$  were considered statistically significant. **Results:** 200 patients were included. Median age was 59 years [52 to 65], 79.5% of patients had comorbidities, 55.6% received adjuvant chemotherapy and 90.4% underwent surgery. At baseline, patients reported a mean GHS of 67 (SD = 19.3), PF (mean = 87.9; SD = 14.8), RF (mean = 82.7; SD = 24), EF (mean = 73; SD = 22, 9), CF (mean = 89.1; SD = 17.6) and SF (mean = 83.8; SD = 23.8). The median follow-up was 24 months [13.8 to 24.3]. Forty-one (20.5%) patients developed at least one major toxicity during the follow-up. In multivariable analysis, for a 10-point increase in GHS, patients had 23% chance do not develop a major toxicities (HR [99% CI] = 0.77 [0.60–0.98],  $p = 0.0073$ ). EF, CF and SF were not associated with major toxicities. At the limit of statistical significance, a 10-point increase in FP, RF showed a trend of protection for the occurrence of toxicities (FP: HR [99% CI] = 0.76[0.57–1.02],  $p = 0.01$ ), (RF: HR [99% CI] = 0.85[0.72–1.01]  $p = 0.01$ ). Conclusion: GHS was independent

prognostic factor of major toxicities. HRQOL should be assessed before treatment to identify patients at risk to develop radiotherapy related toxicities in HNC patients.

### (3047) Core symptoms and their impact on daily functioning at four weeks post-discharge for patients with early-stage lung cancer surgery

Yaqin Wang, Sichuan Cancer Hospital, Chengdu, China; Xiaoqin Liu, Sichuan Cancer Hospital, Chengdu, China; Wei Dai, Sichuan Cancer Hospital, Chengdu, China; Jia Liao, Sichuan Cancer Hospital, Chengdu, China; Xing Wei, Sichuan Cancer Hospital, Chengdu, China; Qiang Li, Sichuan Cancer Hospital, Chengdu, China; Qiuling Shi, Sichuan Cancer Hospital, Chengdu, China

**Aims:** To understand the core symptom burden and its interference with daily functioning of early-stage lung cancer surgical patients 4 weeks after discharge. **Methods:** Data were obtained from a registered prospective cohort study conducting in 6 hospitals from November 2017 to January 2020. Surgical patients with early-stage lung cancer (0-IB) without adjuvant therapy were extracted. The symptoms and their interference with daily life were assessed by the MD Anderson Symptom Inventory-lung cancer module (MDASI-LC) and quality of life (QOL) was measured by the single-item quality of life scale (UNISCALE) four weeks after discharge. By averaging MDASI-LC interference items, we generated two functional scales: activity-related (WAW: walking, activity, and work) and mood-related (REM: relation-with-others, enjoyment-of-life and mood), using 2 (in a 0-10 scale) as the cutpoint for a impaired functional status. **Results:** Total 158 patients had completed MDASI and UNISCALE data (female 62.7%; mean age 55.49 ± 9.92; adenocarcinoma 94.9%; IA stage 70.2%). The 5 most reported symptoms (scored > 0 in a 0-10 scale) were shortness of breath (89.9%), coughing (88.6%), pain (84.8%), fatigue (83.5%) and disturbed sleep (77.8%). The 3 most prevalent moderate to severe symptoms (scored 4-10) were coughing (29.7%), shortness of breath (25.9%) and disturbed sleep (22.2%). The mean score of quality of life was 6.3 ± 2.36. Prevalence of impaired WAW was 53.2%, and REM was 38.6%. Multivariate logistic regression demonstrated that activity-related interference was predicted by shortness of breath (OR 8.42; 95% CI 1.83, 38.76) and difficulty remembering (OR 6.01; 95% CI 1.04, 34.94). Mood-related interference was predicted by pain (OR 9.45; 95% CI 1.69, 52.97), distress (OR 33.31; 95% CI 4.33, 256.37), coughing (OR 4.00; 95% CI 1.35, 11.89) and gender (OR 2.86; 95% CI 1.03, 7.91). High distress score was associated with low QOL ( $B = -1.38$ ,  $SE = 0.65$ ,  $p = 0.037$ ). Female patients had lower QOL than male ( $B = -1.39$ ,  $SE = 0.40$ ,  $p = 0.001$ ). **Conclusion:** Over three-fourths early-stage lung cancer surgical patients reported symptom burden 4 weeks after discharge. The most frequent and most common moderate to severe symptoms are coughing, shortness of breath and disturbed sleep. Interventions for female and specific symptoms (pain, distress, coughing, shortness of breath, difficulty remembering) are warranted to improve daily functioning.

**Table 1. Demographic and clinical characteristics of the patients**

Characteristic	n (%)
Age, yrs, mean(SD)	55.49 (9.92)
BMI(kg/m <sup>2</sup> )	
< 18.5	5 (3.2)
18.5-23.9	93 (58.9)
≥24.0	60 (38.0)
Gender	
Male	59 (37.3)
Female	99 (62.7)
Ethnicity	
Han nationality	156 (98.7)
Hui nationality	2 (1.3)
Highest level of education	
Middle school and below	57 (36.1)
Above middle school	101 (63.9)
Tumor pathological type	
Adenocarcinoma	150 (94.9)
Others	8 (5.1)
Tumor pathological stage	
0	5 (3.2)
I A	111 (70.2)
I B	42 (26.6)
Perioperative complications	
No	116 (73.4)
Yes	42 (26.6)
How to complete the MDASI	
Paper	2 (1.3)
Telephone	7 (4.4)
e-questionnaire	149 (94.3)
How to complete the QOL	
Paper	2 (1.3)
Telephone	6 (3.8)
e-questionnaire	150 (94.9)

**Table 2. Symptom burden of the patients**

Items	Prevalence (rated ≥ 1), n (%)	Moderate to severe symptoms (rated ≥ 4), n (%)	Interquartile range		
			P <sub>25</sub>	P <sub>50</sub>	P <sub>75</sub>
Pain	134 (84.8)	24 (15.2)	1	2	3
Fatigue	132 (83.5)	27 (17.1)	1	2	3
Nausea	44 (27.8)	6 (3.8)	0	0	5
Disturbed sleep	123 (77.8)	35 (22.2)	1	2	3
Distress	89 (56.3)	27 (17.1)	0	1	2
Shortness of breath	142 (89.9)	41 (25.9)	1	2	4
Difficulty remembering	105 (66.5)	27 (17.1)	0	1	3
Lack of appetite	96 (60.8)	21 (13.3)	0	1	2
Drowsiness	93 (58.9)	18 (11.4)	0	1	2
Dry mouth	97 (61.4)	21 (13.3)	0	1	2
Sadness	82 (51.9)	18 (11.4)	0	1	2
Vomiting	32 (20.3)	2 (1.3)	0	0	0
Numbness or tingling	77 (48.7)	16 (10.1)	0	0	2
Coughing	140 (88.6)	47 (29.7)	1	2	4
Constipation	64 (40.5)	11 (7.0)	0	0	1
Sore throat	75 (47.5)	9 (5.7)	0	0	2

**Table 3. Life interference of the patients**

Items	Prevalence (rated ≥ 1), n (%)	Interquartile range		
		P <sub>25</sub>	P <sub>50</sub>	P <sub>75</sub>
Activity-related interference				
General activity	125 (79.1)	1	1.5	3
Walking	122 (77.2)	1	2	3
Work	138 (87.2)	1	2	4
Mood-related interference				
Relations with others	91 (57.6)	0	1	3
Mood	101 (63.9)	0	1	3
Enjoyment of life	107 (67.7)	0	1	3

Items	P value	OR	95%CI
<b>Activity-related interference</b>			
Shortness of breath (0-3 VS 4-10)	0.006	8.42	1.83, 38.76
Difficulty remembering (0-3 VS 4-10)	0.046	6.01	1.04, 34.94
<b>Mood-related interference</b>			
Pain (0-3 VS 4-10)	0.011	9.45	1.69, 52.97
Distress (0-3 VS 4-10)	0.001	33.31	4.33, 256.37
Coughing (0-3 VS 4-10)	0.013	4.00	1.35, 11.89
Gender (1=Male, 2=Female)	0.043	2.86	1.03, 7.91

Variables	B	SE	P value
Constant	8.48	1.29	0.000
Distress (0-3 VS 4-10)	-1.38	0.65	0.037
Gender (1=Male, 2=Female)	-1.39	0.40	0.001

### (3048) Symptoms interfering with activities of daily living in prostate cancer patients receiving oncological treatment—what can we learn from a PRO-study?

Dag Rune Stormoen, MD, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark; Christina Witt Bæksted, MPH, Danish Cancer Society, Copenhagen, Denmark; Gry Assam Taarnhøj, MD, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark; Helle Pappot, MD, DMSc, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark

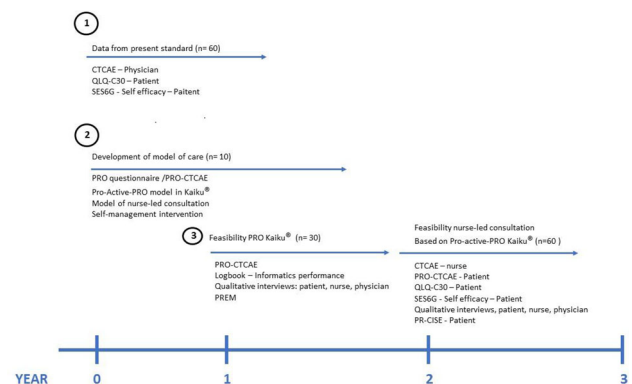
**Aims:** Patient-reported outcome (PRO) can give information to caregivers and doctors, about symptoms during treatment that affect activities of daily living (ADL). By monitoring these symptoms, countermeasures can be undertaken, to secure the patient's ADL and hence quality of life. We here report PRO responses from patients with metastatic castration resistant prostate cancer (mCRPC) during oncological treatment. The aim was to investigate how PRO data can help understand patient experiences during antineoplastic treatment by looking at level of interference with ADL (iADL) for PRO symptoms specific to the mCRPC population. **Methods:** A PRO-CTCAE questionnaire, with 41 items corresponding to adverse events associated with the regimens commonly used for prostate cancer, was collected from patients with mCRPC, receiving treatment at the Department of Oncology, Rigshospitalet, Copenhagen between March 9 and June 8, 2015. Each patient contributed with multiple responses, collected every third week during treatment. The eleven questions regarding iADL were stratified by the antineoplastic agent administered (Docetaxel, Cabazitaxel, Radium-223 and Abiraterone) and grade of iADL (grade 0-4). **Results:** Fifty-four patients completed the questionnaires before each treatment cycle with a total of 168 responses (55 during Docetaxel, 66 during Cabazitaxel, 42 during Radium-223 and 5 during Abiraterone). 89% of responses from patients receiving Docetaxel reported that fatigue interfered with their ADL to some degree, and 22% to a high or very high degree. For Cabazitaxel, Radium-223 and Abiraterone the responses of fatigue iADL is 58%, 82% and 80% respectively. For general pain, the iADL is 52%, 25%, 58% and 80% respectively. For coughing responses with iADL is 29%, 8% 14%, and 20% respectively. **Conclusion:** It is noteworthy that all patient-reported symptoms contain a degree of iADL, ranging from 8 to 89% of responses across the different treatments and symptoms. Coughing is not reported to be a common adverse event for Docetaxel, Cabazitaxel, Radium-223 or Abiraterone. PRO may thus help inform caregivers about symptoms interfering with ADL and inform new patients that even though some of the known side-effects are relatively rare in registration studies,

patients having responded to PRO questionnaires have experienced iADL, previously not reported.

### (3049) A comparable pilot study of nurse-led consultations based on electronic patient-reported outcome (PRO) in a metastatic prostate cancer population during chemotherapy—an example of active PRO

Louise Bager, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark; Helle Pappot, MD DMSc, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark; Gry Assam Taarnhøj, MD, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark; Karin Piil, RN, PhD, Rigshospitalet, University Hospital of Copenhagen, Copenhagen, Denmark

**Aims:** This PRO study aims at developing a model of care for ePRO-based nurse-led consultations facilitating symptom management and shall investigate how these consultations can be introduced into a multidisciplinary treatment regimen for metastatic prostate cancer (mPC) patients receiving chemotherapy at an outpatient clinic with the goal of maintaining health-related quality of life (HRQL). **Methods:** This prospective study will apply a multi-method approach and consists of three parts. Part I; a data collection from standard care providing the study with a control group, physician reported toxicity by CTCAE, quality of life by EORTC QLQ-C30 and self-efficacy by SES6G will be collected. Part II; development of a model of care based on the PEPPA framework—A Participatory, Evidence-Based, Patient-Focused Process for Advanced Practice Nursing (APN) Role Development. Part II also comprises PRO item selection for use in part III. Part III; a feasibility and pilot study of pro-active PRO in nurse-led consultations, PRO by PRO-CTCAE will be added to the parameters collected in part I (see figure). Participants are Danish-speaking men with mPC receiving chemotherapy at the Department of Oncology, Rigshospitalet, Copenhagen. A consecutive sampling strategy will be used, and six nurses specialized in oncology with at least 2 years of experience in the department will conduct the nurse-led consultations. The software Kaiku Health will be used as PRO platform. HRQL will be compared between the standard care group ( $n = 60$ ) and the pro-active PRO nurse-led consultation group ( $n = 60$ ). **Results:** At ISO-QOL 2020 the study design and preliminary results from part I will be presented in detail. **Conclusion:** By presenting the study design and early results of this PRO-study with nurse-led consultations others can be inspired to work systematically with the pro-active use of PRO in cancer. We expect that nurse-led consultations based on ePRO in mPC can be implemented without impairing QoL.



### (3051) Health-related quality of life among family caregivers to patients treated for oesophageal cancer

Cecilia Haddad Ringborg, Karolinska Institutet, Department of Molecular Medicine and Surgery, Surgical Care Science, Stockholm, Sweden; Asif Johar, Karolinska Institutet, Department of Molecular Medicine and Surgery, Surgical Care Science, Stockholm, Sweden; Pernilla Lagergren, Karolinska Institutet, Department of Molecular Medicine and Surgery, Surgical Care Science and Department of Surgery and Cancer, Imperial College London, London, United Kingdom, Stockholm, Sweden

**Aims:** Little is known about factors influencing the health-related quality of life (HRQL) of family caregivers to oesophageal cancer patients. Therefore, the aim of this study was to elucidate how age, sex and educational level of a family caregiver to a patient treated for oesophageal cancer as well the tumor stage of the patient influences their HRQL. **Methods:** A prospective, nationwide and population-based cohort study entitled Oesophageal Surgery on Cancer patients—Adaptation and Recovery (OSCAR) was used for collection of family caregivers. The exposures were family caregivers' age, sex, education level and patients' tumor stage. Information about the family caregivers was self-reported except for educational level that was collected from a national register. Clinical information on patients' tumor stage was collected from medical records. The outcome was HRQL one year after the patient's operation measured by the RAND-36. Multivariable logistic regression analysis provided mean score differences (MSD) with 95% confidence intervals (CI). **Results:** In total 191 family caregivers were included. Family caregivers  $\geq 65$  years reported a clinically relevant and statistically significant lower physical function (MSD  $-10.4$ ; 95% CI  $-15.8$  to  $5.1$ ;  $p = 0.002$ ), a clinically relevant lower role functioning—physically (MSD  $-5.2$ ; 95% CI  $-16.0$  to  $5.6$ ) and a higher level of fatigue (MSD  $6.9$ ; 95% CI  $0.6$  to  $13.3$ ) compared to family caregivers  $< 65$  years. The female family caregivers scored a clinically relevant higher (i.e., better) level of functioning—physical (MSD  $8.7$ ; 95% CI  $-5.9$  to  $23.2$ ) and higher role functioning—emotional (MSD  $6.9$ ; 95% CI  $-9.0$  to  $22.7$ ) but a clinically relevant lower emotional well-being (MSD  $-6.3$ ; 95% CI  $-14.5$  to  $1.9$ ) compared to male family caregivers. Family caregivers with a higher educational level reported more pain (MSD  $9.6$ ; 95% CI  $-1.9$  to  $20.3$ ) compared to those with a lower education level. No clinically relevant differences for tumor stage were found. **Conclusion:** This study indicates that family caregivers' age, sex and educational-level clinically relevantly impact HRQL. More studies are needed to further demonstrate evidence for effects on quality of life in family caregivers to patients treated for oesophageal cancer.

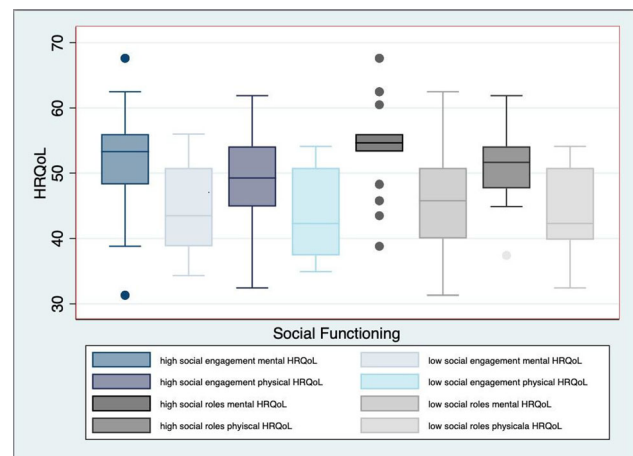
### Cancer Research in Older & Vulnerable Populations

#### (3052) Associations between pre-operative social support and social functioning and post-operative quality of life in patients undergoing radical cystectomy for bladder cancer

Claire Sokas, Brigham & Women's Hospital, Boston, Massachusetts, United States; Zara Cooper, M.D., M.Sc., Brigham & Women's Hospital, Boston, Massachusetts, United States; Andrea Pusic, M.D., Brigham & Women's Hospital, Boston, Massachusetts, United States; Matthew Mossanen, M.D., Brigham & Women's Hospital, Boston, Massachusetts, United States

**Aims:** Bladder cancer is the sixth most common cancer worldwide. Due to high treatment complication and recurrence rates, patients remain engaged with the healthcare system for cancer treatment for years following diagnosis. Patient-reported outcome measures (PROMs) are increasingly used to explore the long-term impacts of

treatment on patient health-related quality of life (HRQoL) and have revealed the negative impact of bladder cancer treatment on physical and social functioning. In this study, we aim to examine the association between pre-operative social support and social functioning on post-operative HRQoL. **Methods:** We identified patients who received radical cystectomy for bladder cancer at an urban tertiary cancer center between January 2018 and January 2020 and had documented pre- and post-operative PROMIS 10 scores. Presence of social support was defined as marital status (married vs. unmarried) and social functioning via patient response of "excellent/very good" to two social functioning questions on the PROMIS 10. We used a Wilcoxon rank-sum test to test associations between pre-operative social support and postoperative patient-reported physical and mental health-related quality of life (HRQoL) on PROMIS 10. **Results:** 45 patients met inclusion criteria. All were male and median age was 69 years (65–73); 84% were married. Marital status was not significantly associated with post-op HRQoL ( $p < 0.63$ ;  $p < 0.27$ ). Two-thirds of patients reported "excellent/very good" satisfaction with social engagement and half reported "excellent/very good" participation in social roles pre-op. Patients with high perceived pre-operative social functioning had better post-operative physical and mental HRQoL (see figure,  $p < 0.02$ ,  $p < 0.03$ ;  $p < 0.002$ ,  $p < 0.001$ ). **Conclusion:** Patient report of pre-operative social functioning, and not marital status, is associated with post-operative quality of life. Pre-treatment assessment of social function may be used to identify patients in need of additional support and at risk for poor HRQoL outcomes. Targeted psychosocial support interventions should be developed to improve outcomes in these patients.



#### (3053) Quality of life in postmenopausal breast cancer patients with localized disease after 5 years of endocrine treatment: a prospective study

Juan Ignacio Arraras, Complejo Hospitalario de Navarra, Pamplona, Spain; Jose Juan Illarramendi, Complejo Hospitalario de Navarra, Medical Oncology, Pamplona, Spain; Ana Manterola, Complejo Hospitalario de Navarra, Radiotherapeutic Oncology, Pamplona, Spain; Susana De la Cruz, Complejo Hospitalario de Navarra, Medical Oncology, Pamplona, Spain; Gemma Asin, Complejo Hospitalario de Navarra, Radiotherapeutic Oncology, Pamplona, Spain; Esteban Salgado, Complejo Hospitalario de Navarra, Medical Oncology, Pamplona, Spain; Marta Barrado, Complejo Hospitalario de Navarra, Radiotherapeutic Oncology, Pamplona, Spain; Lucia Teijeira, Complejo Hospitalario de Navarra, Medical Oncology,



Pamplona, Spain; Amaia Sola, Complejo Hospitalario de Navarra, Radiotherapeutic Oncology, Pamplona, Spain; Berta Ibañez, Navarrabiomed-CHN-UPNA. IdisNA. REDISSEC, Pamplona, Spain; Uxue Zarandona, Navarrabiomed-CHN-UPNA. IdisNA. REDISSEC, Pamplona, Spain; Uxue Zarandona, Complejo Hospitalario de Navarra, Pamplona, Spain; Mikel Rico, Complejo Hospitalario de Navarra, Radiotherapeutic Oncology, Pamplona, Spain; Maite Martinez, Complejo Hospitalario de Navarra, Medical Oncology, Pamplona, Spain; Fernando Arias, Complejo Hospitalario de Navarra, Radiotherapeutic Oncology, Pamplona, Spain; Idoia Morilla, Complejo Hospitalario de Navarra, Medical Oncology, Pamplona, Spain; Ignacio Visus, Complejo Hospitalario de Navarra, Radiotherapeutic Oncology, Pamplona, Spain; Antonio Viudez, Complejo Hospitalario de Navarra, Medical Oncology, Pamplona, Spain; Enrique Martinez, Complejo Hospitalario de Navarra, Radiotherapeutic Oncology, Pamplona, Spain; Ruth Vera, Complejo Hospitalario de Navarra, Medical Oncology, Pamplona, Spain

**Aims:** Quality of Life (QOL) is a key outcome of the attention that is offered to breast cancer patients survivors. More research on the effect of endocrine treatment (ET) on QOL is needed. The aims of the present work are to assess QOL in a sample of early-stage breast cancer survivors who had received ET, and study the changes in QL after ET cessation. **Methods:** A consecutive sample of stages I-III breast cancer patients treated at the Complejo Hospitalario de Navarra was invited to participate in the study. Patients were postmenopausal at diagnoses. Patients filled in the EORTC QLQ-C30 and QLQBR45 questionnaires, and those > 65 years old also the QLQ-ELD14 (elderly specific). Questionnaires were administered twice: at the end of ET (after five years of tamoxifen or aromatase inhibitor -AI), and six months after ET cessation. Differences in ET modality (tamoxifen-AI) in QOL (QLQ-C30, QLQ-BR45, QLQ-ELD14) (U Mann-Whitney), and changes between the two assessments in the three QL questionnaires were assessed (Wilcoxon test). **Results:** 156 patients filled in the first assessment (80 the QLQ-ELD14), 130 the second. Mean age 69 years; 21.6% had tamoxifen, 49% chemotherapy, 90% radiotherapy; 70% conservative surgery. QL scores in the first assessment were high in most areas (> 80/100 points functioning, < 20 points in symptoms areas) with moderate limitations (> 30 points) in sexual functioning and enjoyment; and joint stiffness (elderly specific); and light limitations (20-30 points) in emotional functioning, sleep disturbance, pain, global QL, ET Symptoms, future worries; and future perspective; and worries about others, maintaining purpose and family support (elderly specific) areas. Tamoxifen patients had better sexual functioning (24.2–13.9) and worse body image (89.2 94.6) ( $p < 0.05$ ) Changes between the two assessments ( $p < 0.01$ ): better QL in the second assessment in pain (23.5 first/19.8 s), sexual enjoyment (30.8–40.8), endocrine sexual symptoms (10.7–.6); worsening in future perspective (73.6–66.5) **Conclusion:** Postmenopausal early-stage breast cancer patients adapted well to five years of ET and to their disease. Few QOL differences were observed between ET groups. There was some QOL recovery after ET cessation. A worsening in future perspective could be related to feeling less supported by treatment.

**(3054) Satisfaction with communication in cancer patients: the role of patient evaluation of their competence in communicating with healthcare professionals. An international EORTC study**

Juan Ignacio Arraras, Complejo Hospitalario de Navarra, Pamplona, Spain; Johannes Giesinger, University Hospital of Psychiatry II, Medical University of Innsbruck., Innsbruck, Austria; Omar Shamieh, King Hussein Cancer Center, Amman, Jordan; Iqbal Bahar, Cachar Cancer hospital and Research Centre, Silchar, India; Eva Greimel, Medical University Graz, Graz, Austria; Michael Koller, Centre for

Clinical Studies, University Hospital Regensburg, Graz, Germany; Anne Bredart, Institut Curie and University Paris Descartes, Paris, France; Karin Kuljanic, University Hospital Center Rijeka, Rijeka, Croatia; Anna Costantini, S. Andrea Hospital, Sapienza University, Roma, Italy; Monika Sztankay, University Hospital of Psychiatry II, Medical University of Innsbruck, Innsbruck, Austria; Lisa M. Wintner, University Hospital of Psychiatry II, Medical University of Innsbruck, Innsbruck, Austria; Lisa M. Wintner, Medical University of Innsbruck, Innsbruck, Austria; Marina Carreiro de Sousa, Universidade dos Açores, Ponta Delgada, Portugal; Hiroto Ishiki, National Cancer Center, Tokyo, Japan; Meropi Kontogianni, Harokopio University, Athens, Greece; Maja Wolan, University of Rzeszow, Rzeszow, Poland; Yuichiro Kikawa, Kobe City Medical Center General Hospital, Kobe, Japan; Alicia Gainza, Complejo Hospitalario de Navarra, Pamplona, Spain; Dagmara Kullis, EORTC Headquarters, Brussels, Belgium

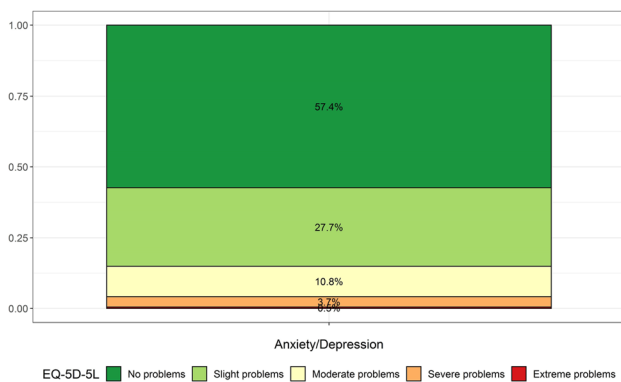
**Aims:** Communication is recognized as one of the main elements of supportive care in cancer patients. The Patient-Centered Communication Model, in which patients play a highly active role, is nowadays predominant. Satisfaction with care is a key Patient-Reported Experience. In this study we analyzed how patients' perceived competence in communicating with their healthcare professionals influenced their satisfaction with the communication. **Methods:** Our international study comprised cancer patients from 12 countries in different regions of Europe (Austria, Croatia, France, Germany, Greece, Italy, Poland, Portugal, Spain) or Asia (India, Japan, Jordan). The patients completed the four scales of the EORTC IN-PATSAT32 satisfaction questionnaire related to communication, and an additional question on their perception of their competence in communicating with their healthcare professionals (doctors and nurses) during consultation (the response categories ranged from 1 "not competent at all" to 10 "highly competent"). The patients were evaluated after a consultation with a doctor or nurse that was held at the start of a new treatment line. We conducted multivariate logistic regression analyses with the four EORTC IN-PATSAT32 scales (dichotomized 0–75/76–100) as dependent variables and the item on competence as the independent variable (adjusted for age, gender, disease stage, and treatment intention (curative/palliative)). **Results:** Data from 978 cancer patients were available for analysis (mean age 58.4 years; 50.2% women; 58.0% curative intention; and 52.7% previous treatment). The mean competence score was 8.1 (SD 1.8). Multivariate analysis showed that greater competence was statistically significantly ( $p < 0.01$ ) associated with greater satisfaction in all four IN-PATSAT32 scales: doctors' interpersonal skills (Odds ratio (OR) 1.41), doctors' availability (OR 1.32), nurses' interpersonal skills (OR 1.50), and nurses' availability (OR 1.29). Moreover, higher age was associated with greater satisfaction with doctors' availability and interpersonal skills, while being male was associated with greater satisfaction with doctors' availability. **Conclusion:** For the patients in our international sample, their evaluation of their competence in communicating with their healthcare professionals is associated with several dimensions of satisfaction with communication. Interventions to address patient competence may help improve communication between patients and healthcare professionals and increase patient satisfaction with communication.

**(3055) The real effects of prostate cancer treatment: EUPROMs prostate patient-driven quality of life study**

Lionne Venderbos, Department of Urology, Erasmus University Medical Center Rotterdam, Rotterdam, Netherlands; Andre Deschamps, Europa Uomo, Antwerp, Belgium; Ernst-Günther Carl, Europa Uomo, Antwerp, Belgium; John Dowling, Europa Uomo, Antwerp, Belgium; Sebastiaan Remmers, Department of Urology, Erasmus University Medical Center Rotterdam, Rotterdam,

Netherlands; Monique Roobol, Department of Urology, Erasmus University Medical Center Rotterdam, Rotterdam, Netherlands

**Aims:** A prostate cancer (PCa) diagnosis and its subsequent treatment will impact a man's quality of life (QoL). To what extent depends on the tumor stage and grade at diagnosis as well as the treatment that is chosen. To date, it is mainly investigator-initiated clinical studies that report on the QoL of men undergoing PCa treatment. Patient-reported outcomes (PROs) collected by a patient organization are, however, lacking. Therefore Europa Uomo, which is the voice of men with PCa in Europe, collected PRO data in a unique prostate patient-driven QoL study, with the aim to provide PCa patients with an idea of what really to expect after PCa treatment. **Methods:** PCa patients in Europe who received treatment were invited to complete a one-time, 20 min online survey, which included demographical questions and the items of the validated EQ-5D-5L measuring generic health, the EORTC-QLQ-C30 measuring QoL of cancer patients and the Expanded Prostate Cancer Index Composite short form 26 (EPIC-26) measuring prostate-specific function. The survey was available in 19 languages. **Results:** 2,943 men from 24 European countries completed the EUPROMs survey between August 21–November 19, 2019. Median age of men who completed the questionnaire was 71 (interquartile range (IQR) 65–75) and the majority was living with a spouse (81.9%). 1937 (65.8%) men indicated that they underwent a single treatment for their PCa, of which radical prostatectomy (RP) was the most reported. 636 (21.6%) men indicated that they underwent two treatments, 300 (10.2%) underwent three treatments and 70 (2.4%) underwent four or more treatments. 42% percent of men reported some anxiety or depression after treatment (Fig. 1), and fatigue and insomnia were affected too. Sexual function was impacted the most (EPIC-26 score of 28/100). Use of, and satisfaction with, medication and devices to help erections is very low. **Conclusion:** This one-of-a-kind patient-driven QoL study indicates the effects of PCa treatment on men via their self-reported QoL. The collected data provides a cross-sectional representation of the current PCa patient population and it shows that initial treatment is very often followed by subsequent treatment, which all have an impact on QoL over a long period.



### (3056) Quantifying the longitudinal association of daily functioning and quality of life in patients with lung cancer surgery during the first month after discharge

Yang Pu, School of Public Health and Management, Chongqing Medical University, Chongqing, China; Qiuling Shi, School of Public Health and Management, Chongqing Medical University, Chongqing, China; Wei Xu, School of Public Health and Management, Chongqing Medical University, Chongqing, China; Qingsong Yu, School of Public Health and Management, Chongqing Medical

University, Chongqing, China; Hongfan Yu, School of Public Health and Management, Chongqing Medical University, Chongqing, China

**Aims:** To explore how daily functioning changes of patients with lung cancer surgery after discharge and the impact of functional changes on quality of life. **Methods:** Patients with lung cancer surgery were longitudinally followed from 1 day before surgery to 1 month after discharge. The six symptom interference items of MD Anderson Symptom Inventory lung cancer module (MDASI-LC) were used to assess patients' daily functional status before surgery and weekly after discharge. The six symptom interference items were categorized as activity-related functioning (WAW, average score of work, activity, walking) and mood-related functioning (REM, average score of relations, enjoy life, mood). A single-item quality of life (SIQOL) measure with 0–10 scale (0 = couldn't be worse; 10 = couldn't be better) was used to assess patients' quality of life with the same schedule as MDASI-LC. **Results:** Among 512 patients recruited, 275 (53.7%) were males and 237 (46.3%) were females. The average age of patients was 55.07 years old (SD = 10.45). The WAW scores before surgery and in first to fourth week after discharge were  $1.69 \pm 3.79$ ,  $10.28 \pm 6.58$ ,  $9.00 \pm 6.15$ ,  $8.32 \pm 6.09$ ,  $7.31 \pm 5.69$ , respectively. The REM scores before surgery and in first to fourth week after discharge were  $2.53 \pm 4.35$ ,  $7.31 \pm 6.61$ ,  $5.99 \pm 5.75$ ,  $6.06 \pm 6.37$ ,  $5.49 \pm 5.56$ , respectively. The SIQOL scores before surgery and in first to fourth week after discharge were  $7.74 \pm 2.16$ ,  $6.05 \pm 2.21$ ,  $6.26 \pm 2.14$ ,  $6.09 \pm 2.29$ ,  $6.29 \pm 2.37$ , respectively. Mixed-effects model showed that the level of WAW ( $\beta_{waw} = -0.1249$ ,  $p < .0001$ ) and REM were the influencing factor of SIQOL ( $\beta_{rem} = -0.1780$ ,  $p < .0001$ ). **Conclusion:** The SIQOL of patients with lung cancer within 4 weeks after discharge was poorer pre-surgery level, and score of WAW and REM within 4 weeks after discharge was higher than pre-surgery level. After being discharged from the hospital, with the daily functions getting better, the quality of life also slowly improved. Strengthening the follow-up of patients' daily functional status after discharge could inform patient management to improve quality of life.

### (3057) Response shift induced by cognitive behavioral therapy targeting fatigue: results of three randomized controlled trials

Mathilde Verdam, PhD, Amsterdam UMC, Amsterdam, Netherlands; Fabiola Müller, Amsterdam UMC, Amsterdam, Netherlands; Frans Oort, Prof, University of Amsterdam, Amsterdam, Netherlands; Heleen Riper, Prof, VU University Amsterdam, Amsterdam, Netherlands; Annemieke van Straten, Prof, VU University Amsterdam, Amsterdam, Netherlands; Irma Verdonck-de Leeuw, Prof, Amsterdam UMC, Amsterdam, Netherlands; Mirjam Sprangers, Prof, Amsterdam UMC, Amsterdam, Netherlands; Hans Knoop, Prof, Amsterdam UMC, Amsterdam, Netherlands

**Aims:** Changes in patients' fatigue following cognitive behavioral therapy (CBT) might reflect the intended relief in fatigue severity as well as a change in the meaning of patients' self-evaluation of fatigue, i.e., response shift. As CBT targets patients' cognitions about fatigue, its effects are likely to induce response shift. However, response shift is rarely investigated within the context of CBT. Therefore, the current paper aims to (1) investigate the occurrence of response shift in patients undergoing CBT for fatigue; (2) estimate the impact of response shift on the intervention effect of CBT; (3) evaluate whether the occurrence of response shifts can be explained by changes in cognitions. **Methods:** We re-analyzed data of three randomized controlled trials (RCTs) on the efficacy of CBT to reduce fatigue in patients with chronic fatigue syndrome ( $n = 225$ ) or diabetes ( $n = 107$ ), and cancer survivors ( $n = 126$ ). Fatigue was assessed with the subscale fatigue of the checklist individual strength. Oort's

structural equation modeling method was applied to assess (1) the occurrence of recalibration, reprioritization and/or reconceptualization response shift; (2) the intervention effect while taking into account possible response shift; and (3) the explanatory role of fatigue- and activity-related cognitions on possible occurrences of response shift. Results: Reprioritization response shift was evidenced in the CBT and not the control groups of all three RCTs, where the fatigue aspect “exhaustion” became less important to patients with chronic fatigue syndrome (effect-size  $d = 0.93$ ), while the aspect “easily tired” became more important to diabetes patients ( $d = -0.89$ ) and cancer survivors ( $d = -0.85$ ). However, the detected response shifts did not affect the overall intervention effects. Changes in some cognitions were related to detected response shifts, but could not explain their occurrences. Conclusion: Change in patients’ fatigue following CBT partly reflects change in the meaning of patients’ self-evaluations, but these did not affect the overall CBT effect. Occurrence of response shift provides insight into differential CBT effects on specific fatigue aspects, and thus helps to understand how CBT reduces patients’ overall fatigue severity. Further research is needed to understand and explain the mechanisms of response shift and shed light on their possible clinical implications.

### **(3058) Dispositional optimism predicts development of anxiety and/or depression among survivors after esophageal cancer surgery**

Yangjun Liu, Karolinska Institute, Stockholm, Sweden; Erik Pettersson, Karolinska Institute, Stockholm, Sweden; Anna Schandl, Karolinska Institute, Stockholm, Sweden; Sheraz Markar, Imperial College London, London, Sweden; Asif Johar, Karolinska Institute, Stockholm, Sweden; Pernilla Lagergren, Karolinska Institute, Stockholm, Sweden

**Aims:** To assess the prevalence trajectory of anxiety and/or depression, and whether dispositional optimism can predict the development of anxiety and/or depression among postoperative esophageal cancer patients. **Methods:** This nationwide longitudinal study included 209 patients who survived for more than one year after esophageal cancer surgery, which were performed in Sweden between January 1, 2013 and December 31, 2017. The exposure was dispositional optimism assessed by the Life Orientation Test-Revised (LOT-R) at 1 year after surgery. The outcome was anxiety and/or depression, which were measured repeatedly at 1, 1.5 and 2 years after surgery by the Hospital Anxiety and Depression Scale. Latent growth curve model was used to assess the prevalence trajectory of anxiety and/or depression, and examine the predictive effect of dispositional optimism on the development of anxiety and/or depression over time after adjusting for demographic and clinical confounders. **Results:** The probability of having anxiety and/or depression continually increased from 1 year to 2 years after esophageal cancer surgery with odds ratio (OR) 2.81 and 95% confidence interval (CI) between 1.62 and 4.88. The odds of developing anxiety and/or depression decreased by 42% (OR 0.58, 95% CI 0.43 to 0.78) with 1 unit increase on the LOT-R sum score, and this protective effect was constant at 1, 1.5 and 2 years after esophageal cancer surgery. **Conclusion:** This study showed that the probability of having anxiety and/or depression increased from 1 year to 2 years after esophageal cancer surgery, and patients with higher dispositional optimism were constantly less likely to suffer from anxiety and/or depression. It is of great importance to evaluate patients’ dispositional optimism early to identify those with higher risk of developing anxiety and depression and to provide timely psychological interventions to them, which could be a potential way to prevent psychological distress and improve health-related quality of life after esophageal cancer surgery.

### **(3059) Long-term quality of life and sexual function of elderly people with endometrial or ovarian cancer**

Ariane Mamguem Kanga, Epidemiology and Quality of Life Research Unit, Inserm U1231, Georges Francois Leclerc Centre, Dijon, France; Leila Bengrine-Lefèvre, Medical Doctor, Medical Oncology Department, Georges François Leclerc Centre, Dijon, France; Valérie Quipourt, Medical Doctor, Department of Geriatrics and Internal Medicine, Hospital of Champmaillot, University Hospital, Dijon, France; Laure Favier, Medical Doctor, Medical Oncology Department, Georges François Leclerc Centre, Dijon, France; Ariane Darut-Jouve, Oncology Centre of park, Dijon, France; Sophie Marilier, Department of Geriatrics and Internal Medicine, Hospital of Champmaillot, University Hospital, Dijon, France; Patrick Arveux, Center for primary care and public health (Unisanté), University of Lausanne, Lausanne, Switzerland; Isabelle Desmoullins, MD, Medical Oncology Department, Georges François Leclerc Centre, Dijon, France; Thienhan Sandrine Dabakuyo-Yonli, Epidemiology and Quality of Life Research Unit, INSERM U1231, Georges Francois Leclerc Centre, Dijon, France

**Aims:** Describe and compare according to age the long-term Health-related quality of life (HRQoL), sexual function and living conditions of adults with either endometrial cancer (EC) or ovarian cancer (OC). **Methods:** A cross-sectional study was set up using data from the Côte d’Or gynecological cancer registry. A series of questionnaires assessing HRQoL (SF-12), sexual function (FSFI), anxiety/depression (HADS), social support (SSQ6) and precariousness (EPICES) were offered to women with EC or OC diagnosed between 2006 and 2013. HRQoL, sexual function, anxiety/depression, social support and precariousness scores were generated and compared according to age (< 70 years and  $\geq 70$  years). **Results:** The response rate was 46%. A total of 145 women with EC ( $n = 103$ ) and OC ( $n = 42$ ) participated in this study. Fifty-six percent and 38% of EC and OC survivors respectively were aged 70 and over. Age did not affect cancer treatment ( $p = 0.477$  for EC and  $p = 0.72$  for OC). The precariousness level did not differ between older and younger survivors with OC while older survivors with EC were more precarious. The physical component of HRQoL was more altered in older EC survivors. This deterioration concerned only the physical functioning dimension (MD = 24,  $p = 0.012$ ) for OC survivors while it concerned physical functioning (MD = 30,  $p < 0.0001$ ), role physical (MD = 22,  $p = 0.001$ ) and bodily pain (MD = 21,  $p = 0.001$ ) for EC survivors. Global health (MD = 11,  $p = 0.011$ ) and role emotional (MD = 12,  $p = 0.018$ ) were also deteriorated in elderly EC survivors. Sexual function was deteriorated regardless of age and cancer location with a more pronounced deterioration in elderly EC survivors for desire ( $p = 0.005$ ), arousal ( $p = 0.015$ ) and orgasm ( $p = 0.007$ ). Social support, anxiety and depression were not affected by age regardless of location. **Conclusion:** 6 years after diagnosis, the impact of cancer on HRQoL is greatest in elderly survivors with either EC or OC.

### **Care of Older & Vulnerable Populations**

#### **(3060) Use of quality of life assessments to support person-centered healthcare for older adults living with frailty and their family caregivers: Development of knowledge translation resources**

Kara Schick-Makaroff, PhD, University of Alberta, Edmonton, Alberta, Canada; Richard Sawatzky, PhD, Trinity Western University, Langley, British Columbia, Canada; Angela Wolff, PhD, Trinity Western University, Langley, British Columbia, Canada; Robin Cohen, PhD, McGill, Montreal, Quebec, Canada; Kelli Stajduhar, PhD, University of Victoria, Victoria, British Columbia,

Canada; Joakim Öhlen, PhD, University of Gothenburg, Victoria, Sweden; Danielle Judd, Trinity Western University, Langley, British Columbia, Canada; Autumn Beemer, University of Alberta, Edmonton, Alberta, Canada; Dominique Duquette, MN, University of Alberta, Victoria, British Columbia, Canada; Mehri Karimi-Dehkordi, PhD, University of Calgary, Edmonton, Alberta, Canada; Landa Terblanche, PhD, University of Calgary, Edmonton, Alberta, Canada; Landa Terblanche, PhD, Trinity Western University, Langley, British Columbia, Canada

**Aims:** This integrated knowledge translation (KT) project focused on developing KT resources to support use of quality of life (QOL) assessments with older adults living with frailty at home and their family caregivers. Resources were developed with knowledge user audiences at different levels of healthcare: patients and family caregivers (micro), healthcare professionals (micro), healthcare leaders (meso) and government decision-makers responsible for healthcare policy (macro). Our research objectives were to describe knowledge gaps, barriers and supportive resources regarding use of QOL assessment tools, including patient-reported outcome and experience measures (PROMs and PREMs), by knowledge users, and formatively evaluate KT resources. **Methods:** Guided by qualitative description, data were collected through 57 interviews and 3 focus groups. Fifty knowledge users participated; 23 of 50 took part in the formative evaluation. All data were recorded, transcribed, and thematically analyzed. NVivo 12 was used to organize and code data. Transcripts were re-read to generate an initial codebook, which was refined in an iterative fashion throughout the analysis. **Results:** There was a need to develop basic yet tailored resources about QOL assessments for each knowledge user audience. Older adults and family caregivers wanted understanding about different aspects of QOL assessments. Healthcare providers sought practical information on how to integrate QOL assessments to minimize burden and to inform person-centered care. Healthcare leaders desired information to understand how PROMs and PREMs could be used in programs, including quality improvement. Government decision-makers needed knowledge about how to access, use, and interpret PROM and PREM information in healthcare decision-making. **KT resources developed included 8 infographic brochures, 1 whiteboard animation, 1 live action video, and a webpage. Formative evaluation was used to refine resources and revealed an overall positive response by all four knowledge user audiences. The participants highlighted the importance and noted potential benefits in the QOL KT resources developed. Conclusion:** This project revealed the need to develop QOL assessment KT materials specific to the intended audiences. These resources provide older adults living with frailty and their family caregivers, clinicians, healthcare leaders, and decision-makers with practical supports to enable person-centered healthcare. Resources are available at [www.healthyqol.com](http://www.healthyqol.com)

### (3061) Provider perceptions of utility and impact of a same-day routine electronic patient-reported outcomes (PRO) assessment in clinical HIV care

Rob Fredericksen, PhD MPH, University of Washington, Seattle, Washington, United States; Duncan Short, ViiV Healthcare, London, United Kingdom; Emma Fitzsimmons, University of Washington, Seattle, Washington, United States; Alex Musten, Ontario HIV Treatment Network, Toronto, Ontario, Canada; Brenda Jacobs, Midway Specialty Clinics, Fort Pierce, Florida, United States; Shivali Suri, St. Michael's Hospital, Fort Pierce, Ontario, Canada; Divya Korlipara, Midway Specialty Clinics, Fort Pierce, Florida, United States; Vincent Hodge, Midway Specialty Care, Fort Pierce, Florida, United States; Kevin Gough, St. Michael's Hospital, Toronto, Ontario, Canada; Moti Ramgopal, Midway Specialty Clinics, Fort Pierce, Florida, United States; Jean Bacon, Midway Specialty Clinics, Fort Pierce, Florida, United States; Jean Bacon, Ontario HIV Treatment Network, Toronto, Ontario, Canada; Justin McReynolds, University of Washington, Seattle, Washington, United States; William Lober, University of Washington, Seattle, Washington, United States; Heidi Crane, University of Washington, Seattle, Washington, United States

**Aims:** We assessed provider and clinic staff perceptions of the utility and acceptability of tablet-based patient-reported outcomes (PRO) assessment integrated into routine HIV care in an academic ambulatory clinic and a community-based clinic in North America. **Methods:** Patients in HIV care self-administered a ~ 10 min PRO assessment of several clinical domains (e.g., antiretroviral adherence, substance use, depression/suicidal ideation, sexual risk behavior, intimate partner violence) on-site immediately prior to their routine care visit. Providers were furnished with succinct summary results before seeing the patient. We conducted (1) 1:1 semi-structured interviews, and (2) subsequent post-interview anonymous surveys with providers. We aggregated quantitative data; qualitative data were collected by digital recorder, transcribed by an independent agency, and coded using qualitative coding software. We coded within thematic areas, and identified key subthemes within each. **Results:** Provider survey data ( $n = 11$ ; 5 MDs, 1 nurse practitioner, 1 physician's assistant, 2 pharmacists, 2 RNs) showed strong agreement that PROs helped prioritize discussion topics with the patient, identified topics that would not otherwise have been addressed, led to more discussions on potentially sensitive topics, made the consultation easier, and added value to the visit overall (82% each); providers disagreed on whether PROs saved time during their consultation (50% agreed, 27% disagreed, 23% neither agreed or disagreed). In interviews, providers reported PROs facilitated identification and ability to address sensitive issues that would likely have been missed, particularly depression/suicidality, sexual behavior, and intimate partner violence. Several providers reported PROs allowed for more comprehensive identification of issues and concerns; this comprehensiveness led to an additional but manageable impact on workflow that was regarded as a valuable tradeoff. Providers reported PROs to be most useful with less well-known patients, with whom patient-provider communication was less established, and with patients not easily agitated or suspicious of questionnaires. **Conclusions:** Providers found PROs with results delivery prior to patient appointments both useful and acceptable for routine HIV care. The value added by PROs to patient care in terms of addressing topics not otherwise likely to have been identified, particularly depression and suicidal ideation, offset additional burden on clinic flow and provider workload.



**(3062) Patient perceptions of the utility and impact of a same-day self-administered routine electronic patient-reported outcomes (PRO) assessment in HIV care in two North American clinics**

Rob Fredericksen, PhD MPH, University of Washington, Seattle, Washington, United States; Duncan Short, ViiV Healthcare, London, United Kingdom; Emma Fitzsimmons, University of Washington, Seattle, Washington, United States; Brenda Jacobs, Midway Specialty Clinics, Ft. Pierce, Florida, United States; Alex Musten, Ontario HIV Treatment Network, Toronto, Ontario, Canada; Divya Korlipara, Midway Specialty Clinics, Toronto, Florida, United States; Shivali Suri, St. Michael's Hospital, Toronto, Ontario, Canada; Vincent Hodge, Midway Specialty Clinics, Ft. Pierce, Florida, United States; Moti Ramgopal, Midway Specialty Care, Ft. Pierce, Florida, United States; Kevin Gough, St. Michael's Hospital, Toronto, Ontario, Canada; Jean Bacon, St. Michael's Hospital, Toronto, Ontario, Canada; Jean Bacon, Ontario HIV Treatment Network, Toronto, Ontario, Canada; Justin McReynolds, University of Washington, Seattle, Washington, United States; William Lober, University of Washington, Seattle, Washington, United States; Heidi Crane, University of Washington, Seattle, Washington, United States

**Aims:** We assessed perceptions of patients living with HIV (PLWH) of the utility and impact of a same-day self-administered tablet-based patient-reported outcomes (PRO) assessment integrated into routine HIV care in two North American clinics. **Methods:** PLWH self-administered a PRO assessment of several clinical domains (e.g., antiretroviral adherence, substance use, depression/suicidal ideation, sexual risk behavior, partner violence) on-site immediately prior to their routine care visit. Providers were furnished with succinct summary results before seeing the patient. We (1) administered a post-appointment multiple choice patient survey querying utility of the PROs in the care visit, and (2) with a separate group of participants, conducted semi-structured 1:1 interview discussing utility and perceived impact of PROs in their appointment in more depth. We aggregated quantitative data; qualitative data were collected by digital recorder, transcribed by an independent agency, and analyzed using qualitative software. We coded within pre-established thematic areas, and identified key subthemes within each. **Results:** In the post-visit survey, PLWH ( $n = 200$ ; median age 50; 28% female; 41% black, 37% white) agreed PROs helped them consider their overall health (89%), helped remind them of health concerns to raise (81%), helped them decide what to talk about (67%), helped discuss topics that might not have otherwise arisen (76%), and helped discuss issues difficult to speak frankly about (71%); 82% reported PROs 'made the visit better overall.' In interviews ( $n = 30$ ; mean age 48; 30% female; 43% black, 36% white) PLWH concurred with and elaborated on these points, reporting PROs added value to their care by: disinhibiting honest responses to psychosocial and behavioral questions via remote response; improving recall and preparation for topic areas to address; enriching patient-provider communication and relationship by identifying mental health/quality of life-related needs; expanding comprehensiveness of care; and promoting self-evaluation in psychosocial and behavioral areas. Several PLWH reported feeling 'cared about' by being asked about mental health and quality of life-related needs. **Conclusion:** PLWH receiving care in two North American HIV clinics found PROs administered before the appointment to be useful for prioritizing discussion topics with their providers, helping initiate discussion on sensitive issues, and improving comprehensiveness of and satisfaction with care.

**(3063) Estimation of EQ-5D index responsiveness to change during an acute exacerbation in COPD patients. Analyses of the AERIS study**

Eliazar Sabater Cabrera, MSc, GSK, Wavre, Belgium; Daniel Molnar, GSK, Wavre, Belgium; Lucio Malvisi, GSK, Siena, Italy; Desmond Curran, GSK, Wavre, Belgium; Emmanuel Aris, GSK, Wavre, Belgium

**Aims:** The EuroQol-5D (EQ-5D) is the preferred method for estimations of health-state utility values used by Health Technology Agencies. The objective of the present analysis is to assess the responsiveness of the EQ-5D index scores during an acute exacerbation event in chronic obstructive pulmonary disease (COPD) patients. **Methods:** The AERIS study was a prospective, interventional, single-center, descriptive, hospital-based, cohort study conducted in the Southampton General Hospital, UK (NCT01360398). The cohort included COPD patients between 40 and 85 years of age with moderate, severe and very severe COPD, according to Global Initiative for Chronic Obstructive Lung Disease (GOLD), and a history of  $\geq 1$  acute exacerbation in the previous 12 months. Patient follow-up period was 2 years. EQ-5D was collected at baseline, every three months at scheduled visits and during exacerbation visits. Mean and standard deviations (SD) were reported for EQ-5D index at baseline visit and exacerbation visits. To assess the responsiveness to health-status change, the mean difference between EQ-5D index scores at baseline and exacerbation visits was estimated and 95% confidence interval (95%CI) was provided. **Results:** The recruited cohort comprised 127 COPD patients with a mean age of 66.8 years (SD 8.6) and a proportion of 53.5% males. The mean of exacerbations in the previous 12 months was 3.1 (SD 2.3) for all patients. During the two years of follow-up, 578 exacerbations were reported. Of those, 51 were mild, 495 moderate and 32 severe. The mean EQ-5D index score was 0.809 (SD 0.194) at baseline and 0.690 (SD 0.238) at the exacerbation visits. The mean EQ-5D score, by severity of exacerbation, was 0.798 (SD 0.192) for mild exacerbations, 0.687 (SD 0.237) for moderate exacerbations and 0.558 (SD 0.257) for severe exacerbations. The estimated overall change of EQ-5D index score from baseline to exacerbation was  $-0.100$  (95% CI  $-0.120, -0.079$ ). The estimated change was  $-0.020$  (95% CI  $-0.076, 0.037$ ) for mild exacerbation,  $-0.100$  (95% CI  $-0.122, -0.079$ ) for moderate exacerbation and  $-0.265$  (95% CI  $-0.419, -0.111$ ) for severe exacerbation. **Conclusion:** EQ-5D index is a sensitive tool able to capture the impairment in health-related quality of life in COPD patients during moderate and severe exacerbation.

**(3064) Measuring experienced quality in home care with patient-reported measures. Understanding the needs of key-stakeholders using the value-proposition canvas**

Roy Haex, MSC PDEng, Maastricht University, Maastricht, Netherlands; Theresa Thoma-Lürken, PhD, Maastricht University, Maastricht, Netherlands; Sandra Zwakhalen, PhD RN, Maastricht University, Maastricht, Netherlands; Anna Beurskens, PhD PT, Maastricht University, Maastricht, Netherlands

**Aims:** To optimize home care, it is essential to determine how care recipients experience quality of care. Traditionally, quality of care is

measured with normative quality indicators. The growing interest for qualitative patient-reported experience measures in home care requires insight into the needs of key-stakeholders. This study aims to understand the needs of clients, formal/informal caregivers, and managers/policy officers in measuring client's experienced quality of care in home care. **Methods:** Four focus group interviews and 25 semi-structured interviews with key-stakeholders were conducted and analyzed by means of content analysis. The value-proposition canvas was used as a thematic framework to explore the purpose of experience quality of care measures and related pains and gains. **Results:** There were two main needs for measuring experienced quality of care: first improving the primary care process of individual clients and second for learning and improving in home care team. Using experienced quality of care measures for external accountability on national level was considered less relevant. Participants described not having time and no clear procedure for conducting an evaluation as a pain of the current methods used. As gains they put forward the ability to informally evaluate experiences during care delivery and to openly discuss complaints with a familiar caregiver. It is important to define when to evaluate, by who, how to evaluate, what motivates and, what to do with outcomes. **Conclusion:** This study advocates that home care organizations should consider selecting methods that fit to clients' and caregivers' needs and prevent dilemmas in evaluating experienced quality of care. It is important to clearly define and communicate the needs for measuring experienced quality of care measures with all key-stakeholders and embed a feasible method in both the primary care process and care teams. The findings of this study can serve as a basis to develop or select methods, in co-creation with key-stakeholders, to assess the experienced quality in home care.

#### **(3065) What are the met and unmet needs of individuals with SCI from rehabilitation to community reintegration: A scoping review**

Rehab Alhasani, PT, BSc., MSc., PhD (candidate), McGill University, Montreal, Quebec, Canada; Diana Zidarov, PhD, Programme de physiothérapie, École de réadaptation, Faculté de Médecine, Université de Montréal, Montreal, Quebec, Canada; Amédé Gogovor, PhD, Institut universitaire sur la réadaptation en déficience physique de Montréal, Centre intégré universitaire de santé et de services sociaux du Centre-Sud-de-l'Île-de-Montréal, Montreal, Quebec, Canada; Melissa Manganaro, MSc, Constance Lethbridge Rehabilitation Center, CIUSSS Centre Ouest de l'île de Montreal, Montreal, Quebec, Canada; Mariana Neibert, MSc, McGill University, Montreal, Quebec, Canada; Matheus De Paiva Azevedo, BSc, McGill University, Montreal, Quebec, Canada; Claudine Auger, PhD, Programme de physiothérapie, École de réadaptation, Faculté de Médecine, Université de Montréal, Montreal, Quebec, Canada; Dany H. Gagnon, PhD, Programme de physiothérapie, École de réadaptation, Faculté de Médecine, Université de Montréal, Montreal, Quebec, Canada; Sara Ahmed, PhD, McGill University, Montreal, Quebec, Canada

**Aims:** To synthesize the met/unmet needs of individuals with SCI as perceived by themselves, their caregivers, and rehabilitation professionals across the continuum of care (i.e., acute, rehabilitation and community) and to provide recommendations to improve SCI care. To note, we are only presenting individuals with SCI perspective. **Methods:** The scoping review was conducted following the Arksey and O'Malley framework and reported according to the Preferred

Reporting Items for scoping review (PRISMA-ScR). An electronic search was performed in Ovid Medline, CINAHL, EMBASE, PsycINFO, Cochrane Library, Web of Science, ProQuest Dissertations and Theses from 1997 to 2017. Four independent reviewers screened all retrieved titles, abstracts, and full texts according to the inclusion criteria and extracted the data. A deductive thematic analysis was used to map each identified need to the International Classification of Functioning, Health, and Disability (ICF) framework. **Results:** A total of 99 articles met the criteria and were included. Of those, 45 articles reported the perspective of individuals with SCI. The SCI needs were classified, as met (24%), unmet (32%) and unspecified needs (51%), and according to ICF, into body function (10%), activities and participation (27%) and environmental factors (63%). At acute care, met needs reported by individuals with SCI included bladder and bowel management, self-care, support from health professionals; no unmet needs were reported. For rehabilitation care, information on existing services and education on consequences once discharged were well addressed; adapted/accessible transportation and access to health services were unmet needs. Finally, at community care, no met needs were reported; unmet needs included information for health services, educational services from health professionals, and services related to community re-integration. **Conclusion:** The results indicate that during rehabilitation, individuals require further education about resources in the community and follow-up of individuals with SCI to provide further intervention for improving their skills or for carrying out their activities of daily living. Furthermore, adults with SCI require opportunities to increase engagement in recreational activities, transportation, and access to other services such as work retraining. More resources should be invested by health authorities for the development and provision of community services focused on maintaining individuals' health and improving social participation.

#### **(3066) Patient-Reported Outcomes Research in Trauma: a Usability Study (PRIORiTy)**

Christel McMullan, PhD, Centre for Patient Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Sally Bradshaw, Centre for Patient Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Ameeta Retzer, PhD, Centre for Patient Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Anita Slade, Centre for Patient Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Tony Belli, Institute of Inflammation and Ageing, University of Birmingham, Birmingham, United Kingdom; Melanie Calvert, Centre for Patient Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Grace Turner, Centre for Patient Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom

**Aims:** Over 50 million people have a traumatic brain injury (TBI) each year and global incidence is rising. Improvements in clinical management of TBI have resulted in improved survival rates; subsequently, more people are living with life changing injuries and reduced quality of life. Electronic assessment of patient-reported outcomes (PROs) post-TBI may facilitate early identification of

ongoing issues, facilitate shared decision-making, and help improve long-term outcomes. The PRiORiTy study aims to explore the feasibility and acceptability of using an electronic Patient-Reported Outcome Measure (ePROM) system for patients with TBI. Methods: The study consists of three stages: (1) a qualitative study (semi-structured interviews with 28 participants) exploring patients' and clinicians' perceptions of an ePROM system; (2) the design of an ePROM system and a usability study through cognitive interviewing, to test this platform; (3) a feasibility and acceptability study in a clinical setting. Results: Findings from the qualitative study showed that all stakeholders were generally supportive of the development and use of an ePROM system as a flexible approach to identify, prioritize and evaluate ongoing issues and ensure that consultations focused on outcomes that matter to patients. Challenges included ensuring that patient issues are accurately captured, and difficulties in completion by patients due to cognitive impairment or lack of insight. Key features of an ePROM system identified by participants (simple layout, use of lay language, opportunity to send/receive feedback, and use of validated tools) were incorporated into the design of the ePROM system. The usability of this ePROM is currently being tested with a number of patients with TBI. Conclusion: Positive attitudes towards ePROMs demonstrate the potential to capture PROs electronically in routine clinical practice and research. The next steps are to refine the ePROM system based on the results of the usability study, and test the acceptability and feasibility of this platform in a clinical setting. It is anticipated that the PRiORiTy study will increase capacity for trauma-specific knowledge and expertise in relation to PROMs, as well as inform system development in other areas of trauma research.

### (3067) What matters to older persons recovering from fractures?

Ahmed Abou-Sharkh, PhD Candidate in Rehabilitation Science, McGill University, Montreal, Quebec, Canada; Nancy E. Mayo, PhD, McGill University Health Center Research Institute, Montreal, Quebec, Canada; Michelle Wall, MSc, McGill University Health Center Research Institute, Montreal, Quebec, Canada; Anthony Albers, MDCM FRCSC, McGill University Health Centre, St-Mary's Hospital, Montreal, Quebec, Canada; Stephane Bergeron, MD MPH FRCSC, McGill University Health Centre, Jewish General Hospital, Montreal, Quebec, Canada; Edward Harvey, HBSc MSc MDCM FRCSC, McGill University Health Centre, Montreal General Hospital, Montreal, Quebec, Canada; Suzanne N. Morin, MD MSc FRCP FACP, McGill University Health Centre, Montreal General Hospital, Montreal, Quebec, Canada

Aims: There is no fracture-specific measure of health-related quality of life (HRQL) suitable for the assessment of longer-term outcomes in this population. Most studies use impairment inventories or generic HRQL measures. As a basis for determining the need for a fracture-specific measure, this study aimed to identify areas spontaneously nominated by people recovering from fractures as affecting their quality of life. Methods: The data came from people enrolled in the HipMobile study, a randomized trial of a technology-facilitated fracture recovery intervention. At study entry and every month for 10 months, people were asked to nominate areas of their life affected by fracture, using an individualized measure, the Patient Generated Index (PGI). Text threads were mapped to the International Classification of Function, Disability and Health (ICF) creating a standard

nomenclature. Results: 62 participants (mean age 75 years; SD: 10) were recruited within a median of 78 (IQR 68-90) days post-fracture repair. At baseline, 212 text threads were processed and the most common areas of life impact were mobility expressed as stiffness and slowness, restricted participation in valued life areas including social and sports activities, autonomy/independence, walking particularly outdoors, endurance/stamina for activities including sitting and standing, pain, driving, activities of daily living, and climbing stairs. While the number of life areas nominated diminished over time, their content remained stable. Using direct questioning, pain was reported by 35% of the sample but nominated spontaneously as affecting quality of life by only 18%. Fatigue was identified in 45% of the sample but nominated by < 10%; depression and anxiety, identified in 10% and 20% of the sample respectively, were not nominated by anyone. Conclusion: This content suggests a complex measurement structure that combines the capacity to do the activity qualified by speed, sustainability, presence of pain and degree of autonomy. This structure would be best represented by latent classes. Domains of fatigue and mood which are common generic HRQL domains were present in this sample but not often or never nominated as impacting quality of life. The content of generic HRQL measures does not fully match what people with fractures say are important to them.

### (3068) Patient-reported outcome assessment in social care: Applications, benefits, and challenges

Sarah Hughes, MHSc, Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, United Kingdom; Olalekan Lee Aiyegbusi, PhD, Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, United Kingdom; Melanie Calvert, PhD, Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, United Kingdom; Phil Collis, Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, United Kingdom; Jon Glasby, PhD, School of Social Policy, University of Birmingham, B15 2TT UK, Birmingham, United Kingdom; Melanie Calvert, PhD, Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, United Kingdom

Aims: Globally, social care is facing unprecedented challenges. An aging population, a fragile workforce, a fragmented care system, and the Covid-19 crisis equates to a system under pressure. There are ongoing, urgent calls for reform. Unmet need is a significant concern and services are under strain to deliver high-quality, person-centered care to more people with limited resources. Patient-reported outcome measures (PROMs) are questionnaires that capture a person's views about their health, functioning, disease symptoms, and quality of life with well-established use in healthcare. To date, however, the use of PROMs' in social care is limited. This paper provides an overview of the applications and benefits that PROMs can bring to social care and discusses challenges to their implementation. Methods: Potential use for individual and aggregate PROM data in social care is presented, drawing on examples from healthcare. We outline applications and benefits of PROMs for social care service users, practitioners and other relevant stakeholders. Challenges to effective implementation are described and possible solutions proposed. Results: PROMs have a range of applications to offer social care. For the individual, this includes improved user-practitioner communication, medication optimization and adherence, and regular monitoring of symptoms and

function with real-time alerts for individuals who are at risk. In planning care, PROMs ensure an individual's health needs are prioritized appropriately. End-of-life measures help to provide supportive, responsive palliative care. At aggregate level, PROM data support provider comparisons, quality improvement, and better integration of health and social care. Effective deployment will depend on stakeholder engagement, selection and standardization of measures, and ensuring accessibility and access for vulnerable populations. Conclusion: PROMs are a proven means for enabling the delivery of high-quality, person-centered care. For these benefits to be realized in social care, further work clarifying current practices around the use of PROMs, barriers and facilitators to implementation and stakeholder engagement will be crucial. These initiatives will inform priority setting and the development of practical guidance on PROMs for the future benefit of those individuals, often vulnerable, who depend on social care.

### **(3069) Quality of life and urban cohabitation: social isolation of elderly people on urban sprawl**

Zeno Mutton, Master Degree in Psychology, University of Padua, Padua, Italy; Marta Casagrande, Psychologist, Associazione "Con Amore e con Rabbia," Padova, Italy; Cristian Bisato, Psychologist and Psychotherapist, Associazione "Con Amore e con Rabbia," Padova, Italy

**Aims:** The quality of life of elderly people is related to multiple aspects: from physical to mental health, from socio-economic to environmental conditions; therefore, to study this construct, it is necessary to take into consideration both the person as a whole, the interaction between the subject and the environment, and the quality of interpersonal relationships. If we consider these aspects as placed in the urban context where the urbanization rate is constantly increasing, the conformation of the city is following the urban sprawl model, and social diversification and cultural pluralism are growing, we have to face with the issue of the city as a place of cohabitation. An issue that highlights the link between cohabitation and quality of life. Psychological literature about quality of life during aging has been focusing on loneliness and its correlations with health outcomes favoring and individual perspective. The aim of the research is to deepen the meaning construction of loneliness by elderly people in their cultural context. **Methods:** We conducted the study in a territory of the north-east of Italy that has been engulfed in the last decades by the urban sprawl of the city of Padua. We conducted two focus groups with 19 elderly people. The focus groups were transcribed verbatim and we are conducting a thematic analysis with the support of Atlas.ti. **Results:** From the results we are expecting to individuate meanings and needs related to loneliness of older citizens in order to planning future researches and community interventions, and to inform administrations for enhance the quality of urban life of elderly people. **Conclusion:** We consider loneliness as a key construct for the analysis of quality of urban life, and therefore an essential point of view in order to investigate health of elderly people in relation to their social and built environment.

### **(3070) Determining the relationship among body mass index and psychosocial factors in knee Osteoarthritis**

Ezinne Ekediegwu, M.Sc, Astella physiotherapy clinics, Nnamdi Azikiwe University, Enugu, Nigeria; Adesola Odole, Ph.D, University of Ibadan, Ibadan, Nigeria; Ifeoma Nwosu, M.Sc, Nnamdi Azikiwe University, Enugu, Nigeria; Chigozie Uchenwoke, M.Sc, University of Nigeria, Enugu, Nigeria

**Aims:** The aim of this study is to determine the relationships of BMI and selected psychosocial factors (kinesiophobia, pain catastrophizing and self-efficacy) among individuals with knee OA in Nigeria. **Methods:** Seventy-seven consecutively sampled patients diagnosed of knee OA from three selected public hospitals in Enugu, South-East Nigeria, participated in this cross-sectional survey. Brief Fear of Movement Scale for Osteoarthritis (BFMSO), Pain Catastrophizing Scale (PCS) and Arthritis Self-Efficacy Scale-8 item (ASES-8 item) were used to assess K, PC, and SE respectively. Also, stadiometer and weighing scale were used to determine height and weight respectively. Data were analyzed using Pearson's correlation coefficient at  $p < 0.05$  and multiple linear regression. **Results:** Participants were aged  $58.04 \pm 12.46$  years. Female participants had higher BMI ( $31.51 \pm 6.82$ ) than the males ( $26.86 \pm 3.03$ ). The mean scores for BMI of the right knee, left knee and bilateral knee were  $29.00 \pm 5.35$ ,  $24.78 \pm 3.74$ ,  $33.02 \pm 6.80$  respectively. Significant positive correlations were found between BMI and PC ( $r = 0.35$ ) whereas significant negative correlations existed between BMI and SE ( $r = -0.30$ ). Significant predictive markers of BMI were PC ( $\beta = 0.21$ ) and SE ( $\beta = -0.89$ ). **Conclusion:** Body mass index, PC and SE correlate significantly in individuals with knee OA. The results call for the routine integration of psychologically informed physiotherapy practice in the management of knee OA.

### **Clinical Care Research I**

#### **(3071) Engaging with parents, caregivers, and clinicians to capture the health-related quality of life of children living with AADC deficiency for a vignette and discrete choice experiment study**

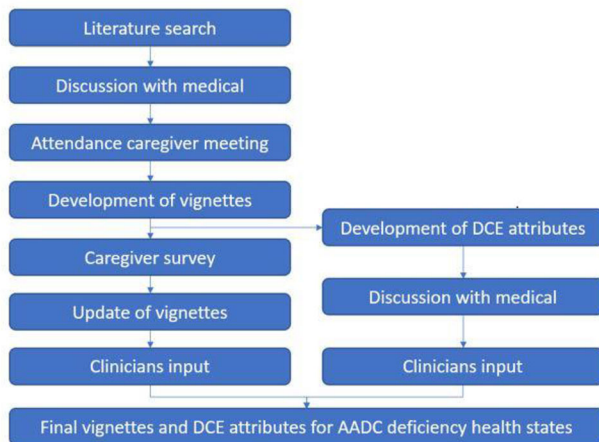
Adam Smith, York Health Economics Consortium, York, United Kingdom; Andria Hanbury, PhD, York Health Economics Consortium, York, United Kingdom; Katharina Buesch, PhD, PTC Therapeutics, Steinhausen, Switzerland

**Aims:** Health-related quality of life (HRQoL) is difficult to measure in rare diseases, especially in pediatric populations. However, capturing HRQoL is critical to evaluating treatment benefit and cost-effectiveness. Given the ultra-rare nature of AADC deficiency (AADCd) rigorous assessment of HRQoL data through proxy caregiver/parent self-report is challenging. Alternatively, HRQoL impact may be ascertained through vignette studies and discrete choice experiments (DCE) using the general public. To maintain face and content validity, caregivers/parents and clinicians treating these patients should be involved in the design of the vignettes and the DCE experiment. The study objective was to develop vignettes and to identify key DCE attributes to estimate patient and caregiver HRQoL. **Methods:** Following a literature review, further insight into the HRQoL impact of AADCd was obtained via discussion with clinicians, as well as from a caregiver/parent advisory board. To ensure the relevance of the subsequently developed vignettes, caregivers/parents were also asked to review and provide input into the descriptions via an anonymized survey. Additional input was obtained during an advisory board with clinicians currently caring for children with AADC-d. As vignettes focus on current state, and ignore improvements on therapy, input into key DCE attributes were obtained including ranking of attributes by clinical, patient and caregiver importance. All input was taken into consideration when finalizing the vignettes and DCE attributes. **Results:** The caregiver/parent and clinician input were used to develop vignettes describing 5 health states: bedridden, head control, sitting unsupported, standing with assistance, and walking with assistance. Six attributes were identified for the DCE: mobility, muscle weakness, oculogyric crises, feeding, cognitive impairment, and crying.



**Conclusion:** This study employed expert opinion to produce vignettes and attributes of AADC-d. These will be used in subsequent studies with the general population to derive HRQoL (utilities) for the health states described in the vignettes, as well as the disutilities associated with attributes of AADC-d. The data will be used to inform a model evaluating the cost-effectiveness of AADC-d treatment.

**Figure 1.** Process outline



### (3072) Quality of Life (QOL) differences between Vietnamese workers and students in Japan

Atsuko Yamaguchi, PhD, Japanese Society of Quality of Life Research, Kobe, Japan; Hideyuki Miyauchi, Japanese Society of Quality of Life Research, Kobe, Japan; Dan T.T. Vu, VNU University of Languages and International Studies, Vietnam National University, Hanoi, Hanoi, Vietnam; Vinh T.H. Nguyen, MD, PhD, Hanoi Medical University, Hanoi, Vietnam; Rika Hayashida, University of Nagasaki, Siebold, Nishisonogi-gun, Nagasaki, Japan; Yoshimi Endo, Graduate School of Medicine, Osaka University, Nishisonogi-gun, Nagasaki, Japan; Shinichi Suzuki, Kashima Hospital, Iwaki, Fukushima, Japan; Taiji Omura, Japanese Society of Quality of Life Research, Kobe, Japan; Takuya Itou, Japanese Society of Quality of Life Research, Iwaki, Fukushima, Japan; Kouki Kaketaka, Japanese Society of Quality of Life Research, Kobe, Japan; Ayako Tsukihashi, Japanese Society of Quality of Life Research, Kobe, Japan; Ayako Tsukihashi, Japanese Society of Quality of Life Research, Kobe, Japan; Hiroshi Shimagami, Japanese Society of Quality of Life Research, Kobe, Japan; Akihiro Yamamoto, Japanese Society of Quality of Life Research, Kobe, Japan; Li-sa Chang, Japanese Society of Quality of Life Research, Kobe, Japan; Shuuto Yamashita, Japanese Society of Quality of Life Research, Kobe, Japan; Masaya Kusuda, Japanese Society of Quality of Life Research, Kobe, Japan; Yoshihiro Tairako, Tounou Sangyou Shinkou Jigyuu Cooperation, Iwaki, Fukushima, Japan; Ichirou Nagano, Japanese Society of Quality of Life Research, Kobe, Japan; Yoshie Tada, Narahamachi Iwaki Office, Iwaki, Fukushima, Japan; Michiko Kobayashi, MD, Japanese Society of Quality of Life Research, Kobe, Japan; Haruyasu Fujita, MD, PhD, Japanese Society of Quality of Life Research, Kobe, Japan; Kouzaburo Adachi, MD, PhD, Japanese Society of Quality of Life Research, Kobe, Japan; Atsuhiko Fukuoka, MD, PhD, Japanese Society of Quality of Life Research, Kobe, Japan; Tomotaka Sobue, MD, PhD, Graduate School of Medicine, Osaka University, Osaka, Japan; Yoshifumi Noda, Japanese Society of Quality of Life Research, Kobe, Japan; Atsuo Okada, Japanese Society of Quality of

Life Research, Kobe, Japan; Takashi Mandai, MD, PhD, Japanese Society of Quality of Life Research, Kobe, Japan

**Aims:** In Japan, the number of foreign workers is increasing and their poor working condition is an urgent issue under discussion. Especially, the system of technical intern trainees (TI trainees) has many problems because they are called “TI trainee” in status but expected to perform hard and low wages work with time limit. The purpose of this study was to investigate the differences in the QOL changes between TI trainees and students from Vietnam, and to examine the support that foreign workers like TI trainees would need. **Methods:** The target populations were 21 TI trainees (average age: 24.4) and 36 students (average age: 28.8) selected from 68 Vietnamese residing in Kanazawa, Japan. We used original self-administered QOL questionnaire comprising 40 questions of 13 categories and 26 Vietnamese specific questions. Data were collected from Dec 2019 to Jan 2020 and analyzed by t-test using SPSS. **Results:** In the TI trainee group, compared with before coming to Japan, there were significant QOL deteriorations in “sleep” ( $p < 0.01$ ), “mental problems” ( $p < 0.01$ ), “physical problems” ( $p < 0.01$ ), “well-being” ( $p < 0.05$ ), and “work performance” ( $p < 0.05$ ) after coming to Japan. On the other hand, in the student group, there were significant QOL improvements in “environmental problems” ( $p < 0.01$ ), “medical problems” ( $p < 0.05$ ), and “passion for life” ( $p < 0.05$ ), but also significant QOL deteriorations in “sleep” ( $p < 0.01$ ), “mental problems” ( $p < 0.01$ ), “dietary problems” ( $p < 0.01$ ), “physical problems” ( $p < 0.05$ ), “work performance” ( $p < 0.05$ ), and “sexual life” ( $p < 0.05$ ). However, compared with before coming to Japan, there were no significant total QOL changes in both groups after coming to Japan. **Conclusion:** There were significant QOL deteriorated categories but no significant QOL improved categories in the TI trainees. However, there were both significant QOL improved and QOL deteriorated categories in the students. These results showed that foreign workers like TI trainees may need more support to improve their QOL in Japan. Moving forward, we must collect more foreign workers’ data focused on difference of position to refine the findings, and provide the evidence to justify how we should accept not only their labor input but also be responsible for their holistic living from the stand point of the Japanese Society of Quality of Life Research.

### (3073) Clinicians’ perspectives on the integration of patient-reported outcomes in a dermatology clinic

Vanina Taliercio, MD, MS, University of Utah, Salt Lake City, Utah, United States; Ashley Snyder, MPH, University of Utah, Salt Lake City, Utah, United States; Kristina Duffin, MD, MS, University of Utah, Salt Lake City, Utah, United States; Allison Biggs, University of Utah, Salt Lake City, Utah, United States; Jacob Kean, PhD, University of Utah, Salt Lake City, Utah, United States; Rachel Hess, MD, MS, University of Utah, Salt Lake City, Utah, United States; Aaron Secrest, MD, PhD, University of Utah, Salt Lake City, Utah, United States

**Aims:** Chronic skin conditions have profound negative effects on psychological and physical functioning. Failure to acknowledge or address these effects can lead to poor treatment adherence and/or patient dissatisfaction. University of Utah’s Dermatology Clinics routinely use the Skindex-16 questionnaire, a patient-reported outcome (PRO) measure, to capture the burden of skin disease. Despite clinical PRO use being highly recommended, real-world adoption has been tremendously slow. We interviewed dermatology clinicians to understand their opinions and perceptions about the facilitators and barriers to PRO use in their daily practice. **Methods:** We conducted in-person semi-structured interviews with 19 clinicians using an interview guide developed based on a literature review and expert

opinion. Interviews were audio-recorded and transcribed verbatim. Two researchers coded the narratives and conducted a thematic analysis using grounded-theory approach. NVivo 12 software was used. Results: We grouped clinicians' beliefs about Skindex-16 into four categories. Most clinicians (15/19) recognized potential benefits to using Skindex-16 including revealing patients' hidden concerns, opening a deeper conversation, spotlighting discrepancies in severity assessments, improving shared decision-making for individualized care, promoting sensitivity to patients' concerns, refining feedback for patients, and providing data for research. Conversely, most clinicians (16/19) also recognized disadvantages to using Skindex-16, such as irrelevance for some diseases, lack of effectiveness in capturing important details, sharing Skindex-16 scores may worry or confuse some patients, and Skindex-16 use might increase liability. Some clinicians (8/19) also recognized the struggles and complaints of patients about Skindex-16, which then might dissuade the clinicians' use of Skindex-16. Finally, most clinicians (14/19) do not believe they have all the elements needed to use PROs successfully in the clinic based on their beliefs in their own ability to interpret Skindex-16 scores and environmental barriers such as time pressure and impact on clinical flow. Conclusion: Most clinicians were open to using Skindex-16 and believed it could be useful for their practice. However, the barriers they listed are very real and need to be addressed to make this PRO more practical for routine clinical care. This study provides a better understanding of the preferences, concerns, and expectations of clinicians regarding PRO implementation in daily practice.

#### **(3074) The role of psychosocial issues on the interpretation of quality of life (QoL) data in Inflammatory Bowel Disease (IBD)**

Monica Hadi, phd, Evidera, London, United Kingdom; Debbie Cooke, phd, University of Surrey, Guildford, United Kingdom

Aims: Inflammatory Bowels Disease (IBD) is a chronic condition which may substantially impair patients' health-related quality of life (HRQOL); however, there is some debate regarding the role of psychosocial factors (such as stress, anxiety, and depression) in triggering or exacerbating the course of IBD. Little is known about how people with IBD perceive this phenomenon and how effectively existing IBD quality of life (QoL) instruments capture it. The main aim of this research was to explore the psychosocial issues that affect the course of IBD and to consider the implications of these findings for the collection and interpretation of QoL data. Methods: Semi-structured qualitative interviews were conducted with adults with IBD recruited through the UK charity Crohn's and Colitis UK. Transcripts were analyzed using thematic analysis to identify, interpret, and link concepts. Findings from this research were examined against several existing IBD QoL measures. Results: A total of 22 individuals with IBD were interviewed. The mean age was 46 years, 45% were male, 60% had CD, and 40% had UC. Participants frequently indicated that stress and anxiety, whether caused by general everyday life (e.g., employment, family, travel) or the disease itself (e.g., toilet habits, discomfort, physical functioning) affected their condition and were related to further relapse of IBD. Coping habits included problem, emotion, and avoidance-based coping and these were particularly important in terms of coping with toilet habits. A review of existing IBD QoL measures highlighted that several measures capture emotional and social impairment as a result of IBD; however, there is little consideration for the impact of psychosocial factors on further disease impairment. Conclusion: Individuals with IBD reported a number of common psychosocial issues which they perceived to have an impact on triggering and exacerbating their IBD. Findings from this research can add further insight to the quality and interpretation of QoL data collected through existing measures, and the kind of support that can

be provided to people with IBD by their carers and health care professionals involved in their care.

#### **(3076) Gamification—the future of real-world patient-reported studies? Analysis of the ethical and legal challenges of applying gamification in observational studies using smartphone apps: a review of the UK and France**

Anna Richards, MA, Vitaccess, Oxford, United Kingdom; Catherine Bottomley, MPharm PhD, Vitaccess, Oxford, United Kingdom

Aims: Gamification within e-health, is a technique that aims to increase patient engagement, overall user experience, and retention through the implementation of game-play elements within a digital study. There has been an increase in its use in real-world studies involving smartphone apps. The aim of this work was to identify the types of incentives that may pose country-specific legal and ethical challenges for application within a real-world study and propose solutions to create a harmonized version for multi-country studies, with a focus on the UK and France. Methods: A targeted review was carried out on the core types of gamification, covering both intrinsic and extrinsic motivation-inducing features such as; points (leading to monetary or non-monetary rewards), performance graphs, badges (collectables), leader boards, avatars, team-mates, meaningful stories and monetary (value) gain. These types of gamification were then analyzed in the context of ethical and legal compliance for real-world studies in the UK and France. Results: We found that some types of gamification could not be associated with the completion of Patient-Reported Outcome Measures (PROMs) within a real-world study, without the risk of compromising instrument validity. Features that were found to be associated with no ethical or legal implication were: performance graphs, badges, meaningful stories or points of no-monetary real-world value. Some exceptions included monetary donations made to charity on behalf of participant engagement, France and the UK differed in permission for this on varying ethical, legal requirements and context grounds. For example, points were allowed in France, but cannot be monetary, or indirectly providing monetary incentive to 3rd party (in accordance with French law). GDPR considerations were highlighted for features involving personally identifiable data such as leader boards, avatars, or team-mates. Conclusion: A one-size-fits-all approach to gamification cannot be used globally in patient-reported app-based studies. Ethical and legal considerations should be investigated for each individual country during the design stage of a study, seeking input from in-country legal and ethical experts before establishing the most suitable incentive framework for gamification for the study.

#### **(3077) Perceived value and utility of the PROMIS Family Relationships measure among providers caring for chronically ill children: A qualitative interview study**

Harald Kliems, MA, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, United States; Kathryn E. Flynn, PhD, Medical College of Wisconsin, Milwaukee, Wisconsin, United States; Elizabeth A. Mann, MD, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, United States; Maria Schletzbaum, BA, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, United States; Elizabeth D. Cox, MD PhD, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, United States

Aims: The Patient-Reported Outcomes Measurement Information System (PROMIS) Family Relationships Short Form measure provides a validated, brief way to assess child patients' perceptions of several facets of family relationships. Our study assessed how healthcare providers perceive the value they may get from using the

family relationships measure in clinical practice, how they would use the information, and what barriers they anticipate. **Methods:** We conducted semi-structured in-person interviews with 20 healthcare providers who care for children with asthma, type 1 diabetes, and sickle cell disease at two academic medical centers. Interviewees included physicians, nurse practitioners, social workers, and health psychologists. None of the interviewees had used the Family Relationships measure at the time of the interview. Interviews were audio-recorded and fully transcribed. Two trained coders analyzed the interview transcripts, using a content analysis approach. **Results:** Healthcare providers widely acknowledged the important role that families and family relationships play in the health of children with chronic conditions. In their current practice, providers already have and use a variety of information about the families of their patients from different sources such as flow sheets, nursing histories, or informal conversations. Interviewees expressed different perspectives on the additional value they may gain from the Family Relationships measure. Some providers thought it would add information that they currently do not have, provide a more efficient and systematic way to get information, or serve as a conversation aid for the clinical encounter. Other providers contrasted potential advantages with concerns about perceived overlap with existing information (e.g., from other measures), the additional burden on providers and patients in time-constrained clinical encounters, and what they would do with the information from the measure. Of note, a majority of interviewees expressed interest in seeing answers to each individual item rather than a summary score alone. **Conclusion:** Implementing the PROMIS Pediatric Family Relationships measure into clinical practice may be welcomed by some providers but face skepticism or indifference from others. Addressing providers' concerns about the measure's additional value should be part of any implementation, for example by integrating the measure into existing workflows and avoiding duplication of already existing information about family relationships.

### **(3078) Quality of Life (QOL) changes of Vietnamese from the view point of the length of stay in Japan**

Atsuko Yamaguchi, PhD, Japanese Society of Quality of Life Research, Kobe, Japan; Hideyuki Miyauchi, Japanese Society of Quality of Life Research, Kobe, Japan; Dan T.T. Vu, VNU University of Languages and International Studies, Vietnam National University, Hanoi, Hanoi, Vietnam; Vinh T.H. Nguyen, MD, PhD, Hanoi Medical University, Hanoi, Vietnam; Rika Hayashida, University of Nagasaki, Siebold, Nishisonogi-gun, Nagasaki, Japan; Yoshimi Endo, Graduate School of Medicine, Osaka University, Nishisonogi-gun, Nagasaki, Japan; Shinichi Suzuki, Kashima Hospital, Iwaki, Fukushima, Japan; Taiji Omura, Japanese Society of Quality of Life Research, Kobe, Japan; Takuya Itou, Japanese Society of Quality of Life Research, Iwaki, Fukushima, Japan; Kouki Kaketaka, Japanese Society of Quality of Life Research, Kobe, Japan; Ayako Tsukihashi, Japanese Society of Quality of Life Research, Kobe, Japan; Ayako Tsukihashi, Japanese Society of Quality of Life Research, Kobe, Japan; Hiroshi Shimagami, Japanese Society of Quality of Life Research, Kobe, Japan; Akihiro Yamamoto, Japanese Society of Quality of Life Research, Kobe, Japan; Li-sa Chang, Japanese Society of Quality of Life Research, Kobe, Japan; Shuuto Yamashita, Japanese Society of Quality of Life Research, Kobe, Japan; Masaya Kusuda, Japanese Society of Quality of Life Research, Kobe, Japan; Yoshihiro Tairako, Tounou Sangyou Shinkou Jigyuu Cooperation, Iwaki, Fukushima, Japan; Ichirou Nagano, Japanese Society of Quality of Life Research, Kobe, Japan; Yoshie Tada, Narahamachi Iwaki Office, Iwaki, Fukushima, Japan; Michiko Kobayashi, MD Japanese Society of Quality of Life Research, Kobe, Japan; Haruyasu Fujita, MD, PhD, Japanese Society of Quality of Life Research, Kobe,

Japan; Kouzaburo Adachi, MD, PhD, Japanese Society of Quality of Life Research, Kobe, Japan; Atsuhiko Fukuoka, MD, PhD, Japanese Society of Quality of Life Research, Kobe, Japan; Tomotaka Sobue, MD, PhD, Graduate School of Medicine, Osaka University, Osaka, Japan; Yoshifumi Noda, Japanese Society of Quality of Life Research, Kobe, Japan; Atsuo Okada, Japanese Society of Quality of Life Research, Kobe, Japan; Takashi Mandai, MD, PhD, Japanese Society of Quality of Life Research, Kobe, Japan

**Aims:** The purpose of this study is to investigate the QOL changes of foreigners coming to Japan from the view point of the length of their stay in Japan. We must support the foreigner's QOL after coming to Japan. **Methods:** 68 Vietnamese divided into two groups participated in this study: The first group is the short period staying Japan (under 2 years) 26 Vietnamese group. The second one is the long period staying Japan (over 2 years) 42 Vietnamese one. We used our new original self-administered QOL questionnaire including 40 questions divided into 13 categories and 26 Vietnamese specific questions. **Results:** Our new original QOL questionnaire had the excellent enough reliability and validity for clinical use. In the short period staying in Japan group, compared with before coming to Japan, there was the significant total QOL deterioration ( $p < 0.05$ ) after coming to Japan. But in the long period staying in Japan group, compared with before coming to Japan, there was no significant total QOL one after coming to Japan. In the short period staying group in Japan, compared with before coming to Japan there were the significant QOL deteriorations in dietary problem ( $p < 0.01$ ), in sleep ( $p < 0.01$ ), in mental problem ( $p < 0.01$ ), in physical problems ( $p < 0.01$ ), in work performance ( $p < 0.05$ ), and in sexual life ( $p < 0.05$ ) after coming to Japan. On the other hand, in the long period staying group in Japan, compared with before coming to Japan there were the significant QOL deteriorations in sleep ( $p < 0.05$ ), in mental problems ( $p < 0.05$ ), in physical problems ( $p < 0.05$ ), in work performance ( $p < 0.05$ ). And there were the significant QOL improvements in medical problems ( $p < 0.01$ ), environmental problems ( $p < 0.01$ ), in economic problems ( $p < 0.01$ ), and in passion for life ( $p < 0.01$ ) after coming to Japan respectively. **Conclusion:** Living years in foreign country is one of the most important factors for the QOL levels for foreigners in Japan. We must help and support the improvement of foreigner's QOL after coming to Japan by annual evolved methodology. So, we must continue our QOL study continuously in future for the happiness of foreigners coming to Japan from the stand point of Japanese Society of Quality of Life Research.

### **(3079) The type of psychiatrists that patients want; a study in secondary care**

Richard Laugharne, Cornwall Partnership NHS Foundation Trust, Liskeard, United Kingdom; Stefan Priebe, WHO Collaborating Centre for Mental Health Service Development, London, United Kingdom; Agnes Chevalier, WHO Collaborating Centre for Mental Health Service Development, London, United Kingdom; Catherine Paton, Devon Partnership NHS Trust, Exeter, United Kingdom; Rajaei Sharma, University of Exeter, Truro, United Kingdom; Alison O'Kelly, Cornwall Partnership NHS Foundation Trust, Truro, United Kingdom; Giles Richards, Cornwall Partnership NHS Foundation Trust, Liskeard, United Kingdom

**Aims:** Patient involvement and considering patient preferences are central principles in healthcare. There appears to be no research to date investigating patients' preferences for socio-cultural characteristics or behavioral qualities of psychiatrists. In addition, there is a dearth of literature examining patient involvement for improving professional performance in medicine. It can take up to 17 years for research to translate into practice in the UK, but this could be decreased if we maximize the role of patients in professional

development. We aimed to assess which characteristics of psychiatrists are most important to patients. This examined socio-cultural characteristics, behaviors and gender bias. Methods: We conducted a survey of patients (132) in community mental health teams across two sites (East Cornwall, East London). Patients completed a brief questionnaire ranking the importance of different socio-cultural characteristics and behaviors of psychiatrists. Results: Patients cared more about age and gender than religion, social background or marital status, but the majority were not concerned with any of these factors. Four clear preferences (from a choice of ten) regarding behavioral qualities were identified as important: explaining things clearly, dedication to personal treatment, being friendly and polite and up to date with medical knowledge. Optimism and recommendation by patients or general practitioners were not as important. Conclusion: Patients are fairly unconcerned about age, gender, religion and social background of psychiatrists. Characteristics they care about most include communication skills, competence, dedication to personal treatment and friendliness. Explaining things clearly is particularly important. This indicates specific areas of improvement for training and further research.

### **(3080) What are the barriers and facilitators to integration of PROMs in routine pediatric clinical care for asthma?**

Sumedh Bele, University of Calgary, Calgary, Alberta, Canada; Maria J. Santana, University of Calgary, Calgary, Alberta, Canada

Aims: Integrating Patient-Reported Outcomes Measures (PROMs) in routine clinical care has shown to have positive impact on health and healthcare, however, there is a scarcity of evidence on the implementation of PROMs in routine pediatric chronic clinical care. Asthma is the most common chronic condition among children, therefore, this qualitative study aims to identify barriers and facilitators to the integration of PROMs in the routine pediatric clinical care at the Alberta Children's Hospital (ACH) Asthma Clinic. Methods: This study is guided by the Theoretical Domains Framework (TDF). The interview guide for data collection includes two to four questions for each of the 14 domains of the TDF. Using stratified purposive sampling strategy, we are recruiting a diverse sample of 20 healthcare providers, pediatric patients receiving care at the asthma clinic and their family caregivers to conduct 14 semi-structured individual interviews and one focus group. Interview and focus group recordings are being transcribed verbatim. Qualitative data analysis software NVivo 12 (QSR, Australia) is being used to code, organize, and manage the data to facilitate data interpretation and analysis. Results: Data collection and analyses is currently underway, and the results will be available at the time of the conference. Results: of this study will be shared with the staff at the ACH asthma clinic to enhance their understanding of the barriers and facilitators to implementation of PROMs in within their own clinic. Facilitators identified through this study will be utilized to facilitate uptake of PROMs, while the barriers will be mitigated using various behavioral change techniques that are likely to change behavior among potential users of PROMs. These results will be crucial to inform the next phase of the study, i.e., piloting the implementation of PROMs using an electronic platform (KidsPRO) at the ACH asthma clinic. Conclusion: While some evidence exists for the enablers and barriers to integration of PROMs in adult care, a comprehensive, systematic, and theory-informed exploration of barriers to the integration of PROMs in routine asthma clinical care is lacking. This study attempts to fill this knowledge gap.

### **(3081) Preferences of women for water immersion during labor and birth**

Thomas Poder, Ph.D., Université de Montréal, Montréal, Quebec, Canada; Nathalie Carrier, M.Sc., CRCHUS, Sherbrooke, Quebec, Canada; Chantal Camden, Ph.D., Université de Sherbrooke, Sherbrooke, Quebec, Canada; Mathieu Roy, Ph.D., CRCHUS, Sherbrooke, Quebec, Canada

Aims: To evaluate women's preferences for water immersion during labor and birth. Methods: An online discrete choice experiment (DCE) was conducted between August 28 and September 9, 2019 to evaluate women's preferences. The DCE included 12 choice cards with 6 attributes (i.e., birth mode, duration of the labor phase, pain sensation, risk of severe perineal tears, risk of death of the newborn, and newborn general condition). Utilities were estimated using logit, latent class, and hierarchical Bayesian (HB) analyses. Results: A total of 1088 subjects completed the survey and at least part of the DCE. The risk of death of the newborn was given high priority by women in all analyses, except in one case, while the risk of severe perineal tears was always considered the least important attribute. The birth mode had moderate importance in the logit model but greater importance in the HB model. The latent class analysis clearly revealed three sub-groups of women. The largest group included 52.9% of women, who were interested in water birth if it could reduce pain and would be risk free for the newborn. The second group included 30.8% of women, who were interested in water birth but only during the labor phase. Finally, the third group (16.2%) did not want to consider water birth, regardless of its risks and benefits. Follow-up questions revealed that many women were interested in water birth only if they could be assured that there would be no risk for the newborn. Additionally, being away from the hospital in case of complications seemed to be a barrier, and women preferred a water birth at the hospital rather than at a birthing center. Conclusion: This study provided insights in favor of water immersion during labor and birth contingent upon the safety of the procedure for the newborn.

## **Clinical Care Research II**

### **(3082) Patient Experience with Patient-Entered Data: "It's valuable to me as a patient, and it's valuable to my healthcare providers"**

Andrew Schuster, BA, Cleveland Clinic, Cleveland, Ohio, United States; Mary Beth Mercer, MPH, Cleveland Clinic, Cleveland, Ohio, United States; Brittany Lapin, PhD, Cleveland Clinic, Cleveland, Ohio, United States; Irene Katzan, MD, Cleveland Clinic, Cleveland, Ohio, United States

Aims: Health status surveys are completed by patients in our healthcare system prior to outpatient appointments. Knowing patients' experiences with patient-entered data (PED) is important to assess and improve the quality of the process. The aim of this study was to explore the patient experience of completing PED across clinical areas and to co-develop a patient experience question to include upon survey completion. Methods: Two focus groups were conducted with patients who have completed PED surveys at our healthcare system. Participants were provided sample patient experience questions to generate feedback and ideas. Qualitative analysis using the framework method was used to identify key topics and themes in the focus group transcripts. Results: Eighteen participants attended two focus groups (56% female; mean age 68 ± 7). Participants supported asking patients a single question about their PED experience. Two themes emerged from the data: usefulness and relevance. Patient perception on whether their answers will be useful to their healthcare provider is a key factor in patients' viewpoints about

completing patient surveys. Additional factors affecting patient's viewpoints include whether they feel the questions improve their own self-awareness and are relevant to their specific health concern(s). Patients who do not find the survey items to be relevant for their upcoming appointment may have a negative experience with PED completion. Focus group participants chose the following survey item as the most appropriate single question to assess patient experience: These questions will help my provider understand my health (strongly agree, agree, disagree, strongly disagree). Conclusion: Findings from our study underscore the perceived value of PED for patients and demonstrate their support of a brief assessment of patient experience upon survey completion. Patients associate a positive experience with perceptions that the information they provide in PED surveys will be relevant to their care and will be useful, particularly to their health-care provider. The patient experience question selected by participants assesses what patients consider to be the most valuable aspect of PED collection—to help their provider better understand their health.

### **(3083) Quality of Life (QOL) changes for Vietnamese people coming to Japan in order to work and study**

Rika Hayashida, University of Nagasaki, Siebold, Nishisonogi-gun, Nagasaki, Japan; Hideyuki Miyauchi, Japanese Society of Quality of Life Research, Kobe, Japan; Atsuko Yamaguchi, Japanese Society of Quality of Life Research, Kobe, Japan; Dan T.T. Vu, VNU University of Languages and International Studies, Vietnam National University, Hanoi, Vietnam; Vinh T.H. Nguyen, Hanoi Medical University, Hanoi, Vietnam; Yoshimi Endo, Graduate School of Medicine, Osaka University, Hanoi, Japan; Shinichi Suzuki, Kashima Hospital, Iwaki, Fukushima, Japan; Taiji Omura, Japanese Society of Quality of Life Research, Kobe, Japan; Takuya Itou, Japanese Society of Quality of Life Research, Kobe, Japan; Kouki Kaketaka, Japanese Society of Quality of Life Research, Kobe, Japan; Ayako Tsukihashi, Japanese Society of Quality of Life Research, Kobe, Japan; Ayako Tsukihashi, Japanese Society of Quality of Life Research, Kobe, Japan; Hiroshi Shimagami, Japanese Society of Quality of Life Research, Kobe, Japan; Akihiro Yamamoto, Japanese Society of Quality of Life Research, Kobe, Japan; Li-sa Chang, Japanese Society of Quality of Life Research, Kobe, Japan; Shuuto Yamashita, Japanese Society of Quality of Life Research, Kobe, Japan; Masaya Kusuda, Japanese Society of Quality of Life Research, Kobe, Japan; Yoshihiro Tairako, Tounou Sangyou Shinkou Jigyuu Cooperation, Iwaki, Fukushima, Japan; Ichirou Nagano, Japanese Society of Quality of Life Research, Kobe, Japan; Yoshie Tada, Narahamachi Iwaki Office, Iwaki, Fukushima, Japan; Michiko Kobayashi, Japanese Society of Quality of Life Research, Kobe, Japan; Haruyasu Fujita, Japanese Society of Quality of Life Research, Kobe, Japan; Kouzaburo Adachi, MD, PhD, Japanese Society of Quality of Life Research, Kobe, Japan; Atsuhiko Fukuoka, MD, PhD, Japanese Society of Quality of Life Research, Kobe, Japan; Tomotaka Sobue, MD, PhD, Graduate School of Medicine, Osaka University, Osaka, Japan; Yoshifumi Noda, Japanese Society of Quality of Life Research, Kobe, Japan; Atsuo Okada, Japanese Society of Quality of Life Research, Kobe, Japan; Takashi Mandai, MD, PhD, Japanese Society of Quality of Life Research, Kobe, Japan

**Aims:** The purpose of this study was to investigate the QOL changes of foreigners coming to Japan in order to work and study. We must

help and support the improvement of foreigner's QOL after coming to Japan. **Methods:** 68 Vietnamese participated in this study. We used our new original self-administered QOL questionnaire including 40 questions divided into 13 categories and 26 Vietnamese specific questions. **Results:** Cronbach's alpha coefficients of our new original QOL questionnaire were excellent enough to accept for clinical use: 0.84 in environmental problems, 0.80 in social participation, 0.73 medical problems, etc. before coming to Japan, and 0.85 in environmental problems, 0.81 in dietary problems, 0.77 in well-being, etc., respectively after coming to Japan. Our new original QOL questionnaire contained 14 main factors and cumulative contribution was 0.71 before coming to Japan. And the one contained 12 main factors and cumulative contribution was 0.70 after coming to Japan, too. Compared with before coming to Japan, there were significant QOL deteriorations in well-being ( $p < 0.01$ ), in dietary problems ( $p < 0.05$ ), in sleep ( $p < 0.05$ ), in mental problems ( $p < 0.05$ ), in physical problems ( $p < 0.05$ ), and in work performance ( $p < 0.05$ ) after coming to Japan. On the other hand, compared with before coming to Japan, there were significant QOL improvements in medical problems ( $p < 0.01$ ), in economical problems ( $p < 0.01$ ), and passion for life ( $p < 0.05$ ) after coming to Japan, too. But compared with before coming to Japan, there was no significant total QOL changes after coming to Japan. **Conclusion:** In many kinds of categories, there were both significant QOL improved and significant deteriorated categories after coming to Japan. We must help and support the improvements of foreigner's QOL after coming to Japan. So, we must continue our QOL study continuously in future for the happiness of foreigners coming to Japan in order to work and study, from the stand point of Japanese Society of Quality of Life Research.

### **(3084) Utility value set for the SF-6Dv2 in Quebec using a hybrid approach**

Thomas Poder, Ph.D., Université de Montréal, Montréal, Quebec, Canada

**Aims:** Cost-utility analysis is increasingly used by decision-makers, but no utility value set is available for Quebec. The aim of this study is to produce a utility value set for Quebec representative of the health preferences of the general population. **Methods:** Between March 2020 and April 2020, an online survey was conducted. The generic preference-based measure used was the new version of the Short Form Six Dimensions (SF-6Dv2). A method combining time trade-off (TTO) and discrete choice experiment (DCE) was used. 216 health states from the SF-6Dv2 were selected using an orthogonal main effects design. Each respondent completed 9 TTO and 7 DCE. In each block of TTO to complete, 7 health states were randomly selected from 76 of the 216 health states, the remaining 2 health states corresponded to the worst health state (i.e., pits) and the health state of the respondent (i.e., SF-6Dv2 previously completed). The DCE section consisted of 7 pairs of health states which were randomly allocated in 10 blocks; each respondent completed one block. The survey also included sociodemographic and debriefing questions. **Results:** 2087 subjects started the survey and 1196 completed it. The mean duration to complete the survey was 59 min. About 35.8% of the respondents found the TTO and DCE sections difficult or very difficult to complete, while 37.6% found it easy or very easy. About 0.7% considered their answers to be of low or very low quality and 78.7% of good or very good quality. Using the hyreg function in Stata 14, the estimate

indicated consistent decrements in all dimensions of the SF-6Dv2. The Pain dimension showed the highest disutility coefficients. Conclusion: Preliminary results indicated that to conduct an online survey to estimate a utility value set for the SF-6Dv2 in Quebec is feasible. Also, the disutilities associated with the worst levels in each dimension were the highest, especially in the Pain dimension.

**(3085) Healthcare providers' implementation of patient-report outcome and experience measures in clinical practice: a mixed-method systematic review using an implementation science framework**

Angela Wolff, RN PhD, Trinity Western University, Langley, British Columbia, Canada; Andrea Dresselhuis, RN MSN, Trinity Western University, Langley, British Columbia, Canada; Duncan Dixon, MEd, MLIS, Trinity Western University, Langley, British Columbia, Canada; Samar Hejazi, PhD, Fraser Health Authority, Surrey, British Columbia, Canada; Deborah Gibson, RN MSN, Trinity Western University, Langley, British Columbia, Canada; Barbara Astle, RN PhD, Trinity Western University, Langley, British Columbia, Canada; Sarah Liva, RN PhD, Trinity Western University, Langley, British Columbia, Canada; Fuchsia Howard, RN PhD, University of British Columbia, Vancouver, British Columbia, Canada

Aims: Substantial literature has highlighted the importance of patient-reported outcome and experience measures (PROMs and PREMs, respectively) to collect clinically relevant information from patients to better understand and address what matters to them. Data from PROMs and PREMs (PROM/EMs) is critical to support clinical decision-making in a person-centered approach. Although many structures and processes exist to support the use of aggregated data, the implementation of patient-reported measures by healthcare providers (HCPs) in clinical practice can be a struggle. This project meets a clinically driven need to synthesize the abundant evidence about how HCPs implement PROM/EMs (and resultant individual-level data) as a routine part of their everyday practice. Methods: A mixed-method systematic review was undertaken to synthesize the grey literature as well as peer-reviewed research evidence (qualitative and quantitative) and quality improvement/implementation studies from eight databases (2010–2020). Combining the 169 keywords synonymous to PROM/EMs with the 41 keywords for implementation yielded 26,134 citations. After applying screening criteria to determine relevance to the extraction questions, 155 sources of evidence that met the criteria were critically appraised using validated tools and data were extracted using the data management software NVivoTM. An integrative synthesis approach using an implementation science framework guided analysis of the extracted data. Results: This review identified (a) providers' experiences in applying PROM/EMs in clinical practice, (b) ways providers integrate PROM/EMs to interpret individual-level data and inform clinical decision-making, and (c) factors that influence implementation and seamless integration of PROM/EMs in everyday practice. The results will exemplify the use of salient strategies for the integration of PROM/EMs by HCPs to develop plans of care, make day-to-day clinical decisions, determine results of care, and ensure continuity of care. Conclusion: The implementation of PROM/EMs into practice requires the uptake of the rich, existing evidence about healthcare providers to support clinical decision-making in a person-centered approach. Furthermore, incorporating implementation theory is crucial to address current barriers encountered. The end-of-grant knowledge translation is a guideline about effective knowledge translation strategies for decision-makers and HCPs on how to use PROM/EMs data for clinical practice decisions at the point of care.

**(3086) Perceived benefits and limitations of using patient-reported outcome measures in clinical practice with individual patients: a systematic review of qualitative studies**

Rachel Campbell, PhD, University of Sydney, Faculty of Science, School of Psychology, Sydney, NSW, Australia; Angela Ju, PhD, University of Sydney, Faculty of Science, School of Psychology, Sydney, Australia; Madeleine King, PhD, University of Sydney, Faculty of Science, School of Psychology, Sydney, Australia; Claudia Rutherford, PhD, University of Sydney, Faculty of Science, School of Psychology, Sydney, Australia

Aims: Patient-reported outcome measures (PROMs) provide important information about the impact of disease and treatment from patients' perspectives. There is increasing interest in using PROMs in clinical settings to inform the care of individual patients. However, evidence regarding whether use of PROMs in clinical settings improves patient outcomes is equivocal. Given this ambiguity, we aimed to determine the benefits and limitations of using PROMs in clinical practice from patient and clinician perspectives. Methods: Systematic review searching Medline, Embase and PsychINFO from inception to January 2020. Qualitative studies examining patients' and/or clinicians' experiences of using PROMs in clinical practice were included. Study screening and data extraction were performed by two independent reviewers. Qualitative data from included studies were analyzed by thematic synthesis. Results: Of 2217 abstracts retrieved, 47 articles reporting 46 studies met eligibility. Seven themes were identified: (1) Active patient involvement (enables awareness and reflection, goal setting, discussion of sensitive topics and influences honesty); (2) Focus of consultation (shifts focus, prioritizes patient needs); (3) Quality of care (prompts action, enables holistic, tailored care, can inaccurately estimate problems); (4) Standardized monitoring of patient outcomes (useful for monitoring treatment effectiveness and PRO changes); (5) Patient-clinician relationship (provides reassurance, inhibits interaction and rapport); (6) Lacks valuable information (PROM not clinically meaningful, provides redundant information); and (7) Not suitable for all patients (e.g., low literacy & cognitively impaired). Conclusion: Both patients and clinicians reported many benefits from using PROMs in clinical practice but also highlighted several limitations. These limitations shed some light on why PROM interventions may not always lead to improved patient outcomes and provide important considerations for the design and implementation of future PROM interventions.

**(3087) Medical staff translation resources and levers to increase migrants' acceptability of rapid tests during the medical consultation at the French migration Office**

Guillaume Roucoux, Master, PRO Team—Inserm UMR 1123, Paris, France; Saleh Ali Fahmi, MD, PRO Team—Inserm UMR 1123, Paris, France; Frédérique Thonon, PhD, PRO Team—Inserm UMR 1123, Paris, France; Olivier Chassany, MD PhD, PRO Team—Inserm UMR 1123, Paris, France; Olivia Rousset-Torrente, Master, PRO Team—Inserm UMR 1123, Paris, France; Martin Duracinsky, MD PhD, PRO Team—Inserm UMR 1123, Paris, France

Aims: High prevalence of HIV, HVB and HCV among migrants justifies targeted screening recommendations from health authorities. Therefore, the STRADA study implemented the use of rapid tests, offered to migrants during the medical consultation at the French migration Office (OFII). Since a previous study showed that language was the major obstacle for the medical staff to offer screening, we focused on the resources doctors and nurses have to communicate with non-French speakers to act for a better acceptance rate. Methods: Individual semi-structured interviews were conducted with 10 doctors

and 10 nurses, in 4 centers of the OFII. Different categories were analyzed through investigator triangulation, then with Sonal software. Results: Doctors and nurses have access to several translation resources: translated documentation and professional interpreters either in attendance or over the phone. But experience shows that they also use their own approximate linguistic skills, the migrant's relation, Google Translate, a health coworker or even another migrant to facilitate communication. The majority of doctors and nurses favor the presence of an interpreter—some prefer a professional one, others a migrant's relative, over documentation or a translation tool. Telephone service offering professional translation is the least used and satisfactory of all. However, almost all caregiver staff members would accept to use an interactive application facilitating screening. The caregiver staff brought out arguments explained to migrants, such as the screening is threefold (HIV, HBV, HCV), anonymous, non-mandatory, quick, painless, reducing risk and good to know one's status and health state. An application should not replace the human presence and connection necessary to announce a positive result. Conclusion: To overcome language barriers in the more effective way when offering a screening, the suggestion of an application received a positive answer from the medical staff. In case of positive results, care is ensured rapidly and for free.

### (3088) Self-reported acne severity in adolescents aligns well with Skindex-Mini, a new three-question quality of life measure

Holly Lind, BS, University of Utah School of Medicine, Salt Lake City, Utah, United States; Vanina L Taliercio, MD, MS, University of Utah Department of Dermatology, Salt Lake City, Utah, United States; Allison M Biggs, University of Utah Department of Dermatology, Salt Lake City, Utah, United States; Zachary H Hopkins, MD, Broward Health Department of Dermatology, Fort Lauderdale, Florida, United States; Leyda Cordova, MBA, University of Utah School of Business, Salt Lake City, Utah, United States; Cameron Jensen, MBA, University of Utah School of Business, Salt Lake City, Utah, United States; Preston Murray, MBA, University of Utah School of Business, Salt Lake City, Utah, United States; Fereshtha Paghmani, MBA, University of Utah School of Business, Salt Lake City, Utah, United States; Adam Sessions, MBA, University of Utah School of Business, Salt Lake City, Utah, United States; Alex Villeda, MBA, University of Utah School of Business, Salt Lake City, Utah, United States; Benjamin West, MBA, University of Utah School of Business, Salt Lake City, Utah, United States; Benjamin West, MBA, University of Utah School of Business, Salt Lake City, Utah, United States; Aaron M Secrest, MD, PhD, University of Utah Department of Dermatology, Salt Lake City, Utah, United States

**Aims:** Acne affects up to 75% of adolescents and often affects visible areas of the body (face, neck, upper chest), negatively impacting quality of life (QOL). Acne patients who report a high impact on QOL were more likely to seek care from a dermatologist. Recently, the Skindex-Mini was adapted and validated from the Skindex-16, a legacy dermatology-specific QOL measure. Skindex-Mini consists of three questions assessing the impact of skin conditions on patients' symptoms, emotions, and functional ability, and was developed as an efficient tool for use in routine clinical care. We aimed to determine how well Skindex-Mini scores aligned with self-reported acne severity in adolescents with acne. **Methods:** As part of an online marketing survey about topical acne treatments, we asked participants to assess their acne severity using validated grading criteria. Participants also completed the Skindex-Mini (range 0–100, 100 = highest QOL impact) as well as demographic and socioeconomic questions. Categorical variables were compared using Chi square; continuous variables with Kruskal–Wallis. SPSS v26 was used for all analyses.

**Results:** Of 141 adolescents who reported they had acne, 137 (97%) completed the Skindex-Mini. Mean age was 15.9 ( $\pm$  2.0) years, and 93 (69%) were female. Females were more likely to report more severe acne ( $p = 0.03$ ). Adolescent acne impacts emotions more than symptoms or functional ability at all severity levels. Median overall Skindex-Mini scores and scores for each domain increased significantly as self-reported acne severity increased (Table 1,  $p < 0.02$  for all). **Conclusion:** Skindex-Mini is a new three-question QOL assessment designed for routine clinical use in dermatology. These data are limited in that most participants reported only mild acne severity. We found that Skindex-Mini aligned well with self-reported acne severity in adolescents and could serve as a simple, objective tool to track QOL impact while treating adolescent acne patients. Future work should assess Skindex-Mini's performance across all acne severities, as well as correlations with clinician-reported assessments of acne severity.

**Table 1** Skindex-Mini scores by self-reported severity of teenage acne (median, IQR,  $n=137$ ). Kruskal-Wallis test used to explore differences.

Skindex-Mini Scores	Range	Self-reported Acne Severity			p-value
		Mild $n=115$	Moderate $n=19$	Severe $n=3$	
Skindex-Mini Overall	0–94	27.8 (12.5–50.0)	61.1 (38.9–77.8)	77.8 (66.7–83.3)	0.001
Skindex-Mini Symptoms	0–100	33.3 (16.7–50.0)	50.0 (33.3–83.3)	100.0 (66.7–100.0)	0.001
Skindex-Mini Emotions	0–100	50.0 (16.7–66.7)	66.7 (33.3–100.0)	100.0 (66.7–100.0)	0.004
Skindex-Mini Functioning	0–100	16.7 (0.0–50.0)	50.0 (16.7–50.0)	50.0 (33.3–66.7)	0.02

### (3089) Quality of life (QOL) for fathers of infants and children

Rika Hayashida, University of Nagasaki, Siebold, Nishisonogi-gun, Nagasaki, Japan; Michiko Kobayashi, Japanese Society of Quality of Life Research, Kobe, Japan; Takashi Mandai, Japanese Society of Quality of Life Research, Kobe, Japan

**Aims:** The purpose of this study was to develop and to investigate the original QOL questionnaire for fathers taking care of infants and children. **Methods:** Data collection was conducted from January to October 2019 with fathers of infants and children six years of age and under in Japan. Sixty-two fathers participated in this study. The original self-administered QOL questionnaire for fathers consists of 23 questions divided into 8 categories. **Results:** The Cronbach's alpha coefficients of father's questionnaire were high enough to accept for clinical use for fathers: 0.93 in eating habits, 0.90 in well-being, 0.90 in financial circumstances, 0.89 in living environment, 0.81 in working environment, respectively. The original QOL questionnaire contained 5 main factors which matched the 8 categories. There were significant positive correlations between well-being and living environment ( $r = 0.68$ ,  $p < 0.01$ ), well-being and eating habits ( $r = 0.59$ ,  $p < 0.01$ ), living environment and eating habits ( $r = 0.53$ ,  $p < 0.01$ ), well-being and sleeping habits ( $r = 0.64$ ,  $p < 0.01$ ), living environment and sleeping habits ( $r = 0.63$ ,  $p < 0.01$ ), eating habits and sleeping habits ( $r = 0.61$ ,  $p < 0.01$ ), well-being and working environment ( $r = 0.60$ ,  $p < 0.01$ ), living environment and working environment ( $r = 0.71$ ,  $p < 0.01$ ), working environment and financial circumstances ( $r = 0.60$ ,  $p < 0.01$ ), respectively. Fathers taking care of children (aged 3 to 6), as compared to those taking care of infants (aged 0 to 2), showed a significantly higher levels of QOL in financial circumstances ( $z = -3.61$ ,  $p < 0.01$ ) and working environment ( $z = -3.03$ ,  $p < 0.01$ ). There were also significantly higher levels of QOL in financial circumstances ( $z = -2.30$ ,  $p < 0.05$ ) and working environment ( $z = -2.73$ ,  $p < 0.01$ ) of fathers who wanted their siblings to guide them, as compared with fathers who didn't want guidance. **Conclusion:** These findings indicate that the original QOL questionnaire had sufficient reliability and potency of validity to use for fathers of infants and children. Living environment, enriched by security, and the existence of a supporter, are most important to improving the QOL of fathers taking care of infants and children.

### **(3090) Multi-stakeholder collaboration to overcome the challenge of recruitment and data collection in hospitalized patients with influenza**

Ewelina Rzepa, PhD, Parexel, London, United Kingdom; Andrea Ireland, PhD, Janssen Inc., New York, New York, United States; Katja Rüdell, PhD, Parexel, London, United Kingdom; Eva Brown-Hajdukova, PhD, Parexel, Prague, Czech Republic; Elliote Hirshberg, MD, Intermountain Healthcare, Salt Lake City, Utah, United States; Jorie Butler, PhD, University of Utah School of Medicine, Salt Lake City, Utah, United States; Valerie Aston, MBA, Intermountain Healthcare, Salt Lake City, Utah, United States; Samuel Brown, MD, University of Utah School of Medicine, Salt Lake City, Utah, United States; Ithan Peltan, MD, Intermountain Healthcare, Salt Lake City, Utah, United States

**Aims:** Validation of a novel Patient-Reported Outcomes (PRO) measure proves challenging when significant time constraints exist for identifying and enrolling required study patient population in hospital settings. We describe a program of collaborative work between different stakeholders to address this challenge. The collaboration aimed to assure an efficient study initiation and an effective patient identification plan for successful completion of a fast-track study of hospitalized patients with seasonal influenza to validate a novel influenza symptom diary. Validation of this diary for use in clinical trials is important to regulators and other key stakeholders to confirm data from the diary are reliable and valid. **Methods:** Parexel engaged a research team from an integrated 24-hospital healthcare system in the United States (Intermountain Healthcare). To initiate the study before the US influenza season, the stakeholders obtained IRB approval for a protocol outlining the process of identifying, enrolling and interviewing up to 25 patients with positive molecular test for influenza performed within 48 h before or 72 h after hospital arrival. Identification of potentially eligible subjects used custom queries of the electronic health records (EHR) based on inclusion criteria. Socio-demographic and clinical variables including NEWS score, comorbidities and medication use were also collected via EHR. Patient eligibility was further confirmed by a manual chart review and only eligible patients were approached about the study. Informed consent was obtained prior to hospital discharge. Patients who consented to the study were scheduled for 90 min, audio-recorded interview 1-6 weeks (target 3 weeks) after hospital discharge. The interviews were conducted in person or via videoconference and followed a semi-structured discussion guide that covered concept elicitation (experience of symptoms, impact on daily activities, relationship of symptoms and comorbidities) and cognitive debriefing of the influenza symptom diary. **Results:** A high-performing multi-stakeholder team and an efficient patient recruitment plan involving a hospital with EHR allowed rapid protocol development, identification and enrollment of challenging patients into the study (within 5 days) followed up with interviews shortly after hospital discharge. **Conclusion:** Collaborative efforts between different stakeholders are an efficient method to conduct a validation study of novel PRO measures with a challenging hospitalized population.

### **Innovative Technology**

#### **(3092) Ethical considerations when accessing, analyzing, and reporting social media data**

Chloe Johnson, MSc, DRG (part of Clarivate), Manchester, United Kingdom; Helen Kitchen, MSc, DRG (part of Clarivate), Manchester, United Kingdom; Chris Marshall, MSc, DRG (part of Clarivate), Manchester, United Kingdom; Jake Macey, MSc, DRG (part of Clarivate), Bicester, United Kingdom; Natalie V J Aldhouse, MSc,

DRG (part of Clarivate), Manchester, United Kingdom; Tamara Al-Zubeidi, BSc, DRG (part of Clarivate), Manchester, United Kingdom; Hannah C Pegram, BSc, DRG (part of Clarivate), Manchester, United Kingdom; Maile Hunter, MBA, DRG (part of Clarivate), Nashville, Tennessee, United States; Sarah Knight, MSc, DRG (part of Clarivate), Bicester, United Kingdom

**Aims:** To present ethical considerations and recommendations for accessing, analyzing, and reporting social media data in the context of patient-reported outcome (PRO) development. **Methods:** The FDA Patient-Focused Drug Development (PFDD) draft guidance suggests analysis of social media data (e.g., social networking sites, blogs, forums) as an initial, supplementary data source to inform development of research tools (e.g., interview guides). The data obtained can be a useful source of preliminary information where there is little published qualitative data. We reviewed guidance from the FDA, European Medicines Agency, and British Psychological Society on the conduct of social media research to provide an overview of ethical considerations and recommendations for PRO researchers. **Results:** There are several ethical principles to consider when accessing, analyzing and reporting social media data. The principles of general observational research should be applied to data obtained via social media, in that individuals should only be publicly observed in circumstances where they would expect to be observed by strangers. Websites that require a log-in or password should not be used as this represents a private environment; therefore individuals would not expect to be observed. Researchers should not attempt to influence the content posted on social media (e.g., posting targeted questions) without consent. It is important to acknowledge the lack of informed consent or confirmation of diagnosis when interpreting the findings from social media data. When reporting the findings from a social media review, the original 'poster/blogger' should remain anonymous. Verbatim quotes should not be included in reports/publications. It is the responsibility of the researcher to ensure that any paraphrased quotes included in publications cannot be 'reverse' searched to identify the original source. In addition, any individual demographic data obtained should only be reported when necessary to address the research aims. These considerations are particularly important when reporting sensitive topics or in rare/vulnerable populations. **Conclusion:** Social media data can provide rich, preliminary insights regarding the symptoms and impacts of a condition. It is important that researchers access, analyze, and report social media data in an ethical manner to protect individuals' privacy, despite the perception that social media posts are 'public.'

#### **(3093) Complexities in localizing participant profile survey questions in real-world studies for smartphone apps**

Sujyotee Kretz, MSc in Technical, medical and scientific translation with translation technology, Vitaccess, London, United Kingdom; Anna Richards, MA in Translation Studies, Vitaccess, Oxford, United Kingdom

**Aims:** Real-world evidence (RWE) studies administered on smartphone apps can include several types of data, including Patient-Reported Outcome Measures (PROMs) and symptom trackers. A vital part of the app is the participant profile which seeks to obtain background information about participants, such as demographics, medical history, diagnosis and treatment information to allow data analysis by specific patient groups. The aim of this research was to identify potential challenges in the localization of participant profile survey questions using an observational study in neuromuscular disease as a case study. **Methods:** The participant profile survey was localized from UK English for United States, Canada, Japan, France, Belgium, Germany, Italy and Spain. The localization process of the survey



followed the ISPOR Principles of Good Practice and involved collaboration between PRO experts, app user experience and user interface developers, linguists and localization project managers to develop the source content and translations. The reports produced from each localization step were used to identify any potential challenges. Results: The types of content that were complex to localize were categorized as follows: (1) Cultural: elements such as date format, measurement units, post code format vary across different countries. In some countries there is no equivalent of NHS number. (2) Legal: the translation had to differ from the source for legal reasons, for example, providing NHS equivalent number and full address is prohibited in some countries. For these countries, the content had to be adapted to satisfy legal requirements. (3) Medical: treatment may vary from one country to another. For example, the localization of generic and brand names of treatments involved clinician input to ensure the content was accurate. Conclusion: Localizing the participant profile survey requires a flexible approach to app content design. Identifying challenges during source content development is an essential process for determining how it will translate into the app and what impact it will have on data analysis. A translatability assessment on the content is also recommended. These steps allow content to be optimized for each country and for the study to be understood and viewed as relevant by the participants, thereby encouraging participant engagement.

**(3094) Using consumer-friendly wearables to associate patient-reported quality of life and tech-reported physical activity and sleep in healthy seniors**

Vlad Manea, M.Sc., Quality of Life Technologies Lab, University of Copenhagen, Copenhagen, Denmark; Katarzyna Wac, Ph.D., Quality of Life Technologies Lab, University of Copenhagen, Copenhagen, Denmark

Aims: Behaviors account for 50% health risk and affect life quality later in life. Numerous studies quantified the relationships between isolated behaviors and life quality in clinical cases, short-term, using momentary-reported outcomes or expensive wearables. However, little research studied relations across multiple behaviors in healthy seniors wearing their own devices long-term (7-120 days). Methods: 42 seniors in Spain and Hungary (aged  $68.78 \pm 6.30$ ) patient-reported Quality of Life (EQ-5D-3L) and tech-reported daily life behaviors (Fitbit Charge 2). We align answers to intervals (7–120 days) by administration date and end date, within a leeway proportional to the interval. duration We derive patient-reported variables and tech-reported variables (energy, steps, distance, duration of sedentary, activity, sleep, and resting heart rate) in absolute and, where relevant, relative (compositional) quantities. We quantify Spearman associations at  $\alpha = 0.05$ . Results:  $n = 31$  participants (aged  $70.66 \pm 3.15$ ; 21 in Spain and 10 in Hungary) provided 54 EQ-5D-3L answers ( $1.72 \pm 1.12$ /person) and 9.150 Fitbit days ( $295.16 \pm 247.25$ /person). 10 participants reported mild disease. In all participants, distance and steps associated with mobility ( $r = 0.71$ ),  $p < 0.005$ . Sleep duration inversely associated with anxiety ( $-0.57$ ) and pain ( $-0.52$ ); vigorous duration associated with health state (0.70); relative light activity associated with health state (0.63),  $p < 0.005$ . In healthy participants, absolute sedentary duration associated with a lack of mobility (0.57) and pain (0.69), and a high resting heart rate associated with poor health (0.56),  $p < 0.005$ . Relative sedentary duration associated with pain (0.62) and lack of anxiety (0.54) while light relative duration associated with health state (0.64). Relative sleep duration associated with health state (0.65). In sick participants, distance and steps associated with mobility (0.71) and lack of anxiety ( $-0.57$ ), stronger, less significant, and only over longer periods than absolute quantities. Conclusion:

Our method is feasible in associating behaviors and mobility, pain, and health status for short periods (7-21 days) in a small sample of healthy participants. Monitoring physical activity log-term (90-120 days) helped better assess mobility, anxiety, and health state in sick seniors. Our results provide insights for designs targeting interventions for seniors.

Table 1: Participant characteristics

	Spain		Hungary		Count	Smoking	Spain	Hungary
	21	10	21	10				
Age	71.00	70.50	Yes	5	1			
Median	70.67	71.10	No	16	9			
Mean	70.67	71.10	Alcohol					
Stdev	3.08	4.69	Never	7	2			
Gender			Monthly	3	5			
Female	13	5	Weekly	8	0			
Male	8	5	Few days	1	2			
Education			Daily	2	1			
Primary	6	0	Systolic blood pressure					
Secondary	4	2	9					
High school	4	1	Median	125.5	120.00			
University	7	7	Mean	124.4	129.45			
Living			Stdev	15.31	22.07			
Alone	9	1	Body mass index					
Partner	10	9	Median	24.72	26.83			
Children	2	0	Mean	26.01	28.37			
Grandchildren	0	0	Stdev	4.71	4.10			

Table 2: Spearman associations of wearable behaviours (rows) and EQ-5D-3L Quality of Life (columns), all participants

	Mobility	Self-care	Usual activities	Pain / discomfort	Anxiety / depression	Health state
Energy out	0.21,	0.17	0.17	0.14,	0.02	0.03
Distance	0.45 <sup>0.21</sup> , 0.28,	0.28,	0.05	0.26,	0.00,	0.00,
Step count	0.47 <sup>0.14</sup> , 0.27,	0.27,	0.01,	0.21,	0.02,	0.02,
Sedentary duration	0.56 <sup>0.31</sup> , 0.18	0.18	0.18	0.50 <sup>0.17</sup> ,	0.35 <sup>0.33</sup> ,	0.34 <sup>0.35</sup> ,
Light physical activity duration	0.49 <sup>0.20</sup> ,	0.14,	0.14,	0.25,	0.41 <sup>0.27</sup> ,	0.14
Fair physical activity duration	0.45,	0.28,	0.28,	0.03	0.03,	0.03,
Vigorous physical activity duration	0.22,	0.00	0.00	0.11	0.19,	0.70 <sup>0.13</sup> ,
Daily physical activity duration	0.32,	0.00	0.00	0.00	0.30,	0.54 <sup>0.13</sup> ,
Sleep duration	0.11,	0.30,	0.30,	0.30 <sup>0.15</sup> ,	0.57 <sup>0.16</sup> ,	0.44 <sup>0.18</sup> ,
Resting heart rate	0.16	0.14	0.14	0.38 <sup>0.15</sup> ,	0.19,	0.20,
PAC CLR Sedentary duration	0.28	0.00	0.00	0.08	0.59 <sup>0.15</sup> ,	0.64 <sup>0.13</sup> ,
PAC CLR Light physical activity duration	0.18,	0.00	0.00	0.26,	0.15,	0.63 <sup>0.11</sup> ,
PAC CLR Fair physical activity duration	0.00	0.00	0.00	0.08 <sup>0.18</sup> ,	0.23	0.50 <sup>0.13</sup> ,
PAC CLR Vigorous physical activity duration	0.00	0.00	0.00	0.26,	0.07	0.50 <sup>0.13</sup> ,
PASC CLR Sedentary duration	0.00	0.00	0.00	0.13,	0.30	0.62 <sup>0.13</sup> ,
PASC CLR Light physical activity duration	0.00	0.00	0.00	0.30,	0.03,	0.14
PASC CLR Fair physical activity duration	0.00	0.00	0.00	0.02,	0.42	0.49 <sup>0.17</sup> ,
PASC CLR Vigorous physical activity duration	0.00	0.00	0.00	0.41,	0.25,	0.54 <sup>0.16</sup> ,
PASC CLR Sleep duration	0.00	0.00	0.00	0.58 <sup>0.15</sup> ,	0.30	0.73 <sup>0.12</sup> ,

CLR: association strength (normal script), duration, sample size, and significance (superscript) and direction (subscript). Example: 0.46<sup>0.11</sup> depicts a (negative) association of -0.46 and  $p < 0.005$  for an interval with 28 days yielding 11 alignments. Significance: \* for  $p < 0.05$ ; † for  $p < 0.005$ . Colors: gray (significant associations) or white (not significant associations). PAC: physical activity compositional; PASC: physical activity and sleep compositional; CLR: centered log ratio of the duration.

Table 3: Spearman associations of wearable behaviours (rows) and EQ-5D-3L Quality of Life (columns), healthy participants

	Mobility	Self-care	Usual activities	Pain / discomfort	Anxiety / depression	Health state
Energy out	0.25,	0.25	0.25	0.13,	0.15	0.11
Distance	0.39 <sup>0.15</sup> ,	0.30,	0.30,	0.05	0.18,	0.14,
Step count	0.46 <sup>0.11</sup> ,	0.35,	0.35,	0.02	0.12,	0.09,
Sedentary duration	0.57 <sup>0.11</sup> ,	0.24	0.24	0.69 <sup>0.13</sup> ,	0.53 <sup>0.17</sup> ,	0.37 <sup>0.18</sup> ,
Light physical activity duration	0.46 <sup>0.15</sup> ,	0.12,	0.12,	0.27,	0.47 <sup>0.17</sup> ,	0.44 <sup>0.18</sup> ,
Fair physical activity duration	0.30,	0.30,	0.30,	0.55 <sup>0.14</sup> ,	0.52 <sup>0.15</sup> ,	0.46 <sup>0.17</sup> ,
Vigorous physical activity duration	0.00	0.00	0.00	0.04	0.15	0.50 <sup>0.14</sup> ,
Daily physical activity duration	0.00	0.00	0.00	0.04	0.04	0.21,
Sleep duration	0.38,	0.38,	0.38,	0.45,	0.47 <sup>0.11</sup> ,	0.40
Resting heart rate	0.35	0.16	0.16	0.00	0.08,	0.56 <sup>0.16</sup> ,
PAC CLR Sedentary duration	0.00	0.00	0.00	0.62 <sup>0.13</sup> ,	0.44 <sup>0.18</sup> ,	0.32
PAC CLR Light physical activity duration	0.00	0.00	0.00	0.18	0.30,	0.64 <sup>0.15</sup> ,
PAC CLR Fair physical activity duration	0.00	0.00	0.00	0.12	0.02,	0.15,
PAC CLR Vigorous physical activity duration	0.00	0.00	0.00	0.00	0.19,	0.52
PASC CLR Sedentary duration	0.00	0.00	0.00	0.56 <sup>0.13</sup> ,	0.01,	0.22
PASC CLR Light physical activity duration	0.00	0.00	0.00	0.22	0.00	0.33
PASC CLR Fair physical activity duration	0.00	0.00	0.00	0.00	0.13,	0.00,
PASC CLR Vigorous physical activity duration	0.00	0.00	0.00	0.19,	0.25,	0.53
PASC CLR Sleep duration	0.00	0.00	0.00	0.22,	0.38,	0.65 <sup>0.17</sup> ,

CLR: association strength (normal script), duration, sample size, and significance (superscript) and direction (subscript). Example: 0.46<sup>0.11</sup> depicts a (negative) association of -0.46 and  $p < 0.005$  for an interval with 28 days yielding 11 alignments. Significance: \* for  $p < 0.05$ ; † for  $p < 0.005$ . Colors: gray (significant associations) or white (not significant associations). PAC: physical activity compositional; PASC: physical activity and sleep compositional; CLR: centered log ratio of the duration.

Table 4: Spearman associations of wearable behaviours (rows) and EQ-5D-3L Quality of Life (columns), sick participants

	Mobility	Pain / discomfort	Anxiety / depression	Health state
Energy out	0.25,	0.45,	0.64 <sup>0.11</sup> ,	0.16
Distance	0.71 <sup>0.11</sup> ,	0.19,	0.62 <sup>0.13</sup> ,	0.15
Step count	0.71 <sup>0.11</sup> ,	0.23,	0.57 <sup>0.13</sup> ,	0.41
Sedentary duration	0.67 <sup>0.11</sup> ,	0.03	0.07	0.23
Light physical activity duration	0.22,	0.20	0.12,	0.12
Sleep duration	0.26,	0.12	0.03,	0.06
Resting heart rate	0.06	0.10,	0.39,	0.30

CLR: association strength (normal script), duration, sample size, and significance (superscript) and direction (subscript). Example: 0.46<sup>0.11</sup> depicts a (negative) association of -0.46 and  $p < 0.005$  for an interval with 28 days yielding 11 alignments. Significance: \* for  $p < 0.05$ ; † for  $p < 0.005$ . Colors: gray (significant associations) or white (not significant associations). Alignment sample size showed only a selection of aligned variables.

**(3095) Association of physical activity intensity with scores on PROMIS® Pediatric measures in children and adolescents with asthma**

Nicole Lucas, Duke University, Durham, North Carolina, United States; Courtney Mann, Duke University, Durham, North Carolina, United States; Michelle Hernandez, University of North Carolina - Chapel Hill, Chapel Hill, North Carolina, United States; Wanda Phipatanakul, Boston Children’s Hospital, Boston, Massachusetts, United States; Li Lin, Duke University, Durham, North Carolina, United States; Antonia V. Bennett, University of North Carolina—Chapel Hill, Durham, North Carolina, United States; Allison Burbank, University of North Carolina—Chapel Hill, Chapel Hill, North Carolina, United States; Steven J. Lippmann, Duke University,

Durham, North Carolina, United States; Amparito Cunningham, Boston Children's Hospital, Boston, Massachusetts, United States; Michelle Maciag, Boston Children's Hospital, Boston, Massachusetts, United States; Nicole Akar-Ghibril, Boston Children's Hospital, Boston, Massachusetts, United States; Nicole Akar-Ghibril, Boston Children's Hospital, Boston, Massachusetts, United States; Bryce B. Reeve, Duke University, Durham, North Carolina, United States

**Aims:** The Garmin VivoFit3 is an accelerometry-based pedometer that records the number of steps walked and can distinguish between different types of physical activity intensity levels (i.e., walking, running or sedentary). In children and adolescents with asthma, one important mechanism linking asthma and obesity involves sedentary lifestyle and physical activity avoidance. However, little data is available regarding the association of sedentary lifestyle with health-related quality of life (HRQOL) outcomes like anxiety, fatigue or depression. This study explores the association of physical activity intensity and HRQOL scores in children and adolescents with asthma. **Methods:** Participants between the ages of 8–17 years with uncontrolled or partly controlled asthma were recruited from two academic medical centers. Participants completed lung function tests (spirometry) and other asthma control tests at two clinic visits. Between visits, participants wore the Garmin VivoFit3 monitor for 28 days and completed PROMIS® Pediatric measures (Asthma Impact, Anxiety, Depressive Symptoms, Fatigue, Mobility and Peer Relationships) on Days 7, 14, 21 and 28. Pearson's correlation coefficient was used to assess association between PROMIS measures and physical activity. **Results:** Out of 105 children and adolescents enrolled, 90 (44% female) who had activity data and completed PROMIS measures for at least one of the weekly time points were included. Higher levels of sedentary minutes were associated with worse PROMIS Pediatric Asthma Impact (Week 1:  $r = 0.37$ ,  $p < 0.01$ ; Week 2:  $r = 0.3$ ,  $p < 0.01$ ; Week 4:  $r = 0.37$ ,  $p < 0.01$ ), Depressive Symptoms (Week 1:  $r = 0.25$ ; Week 4:  $r = 0.25$ ), Fatigue (Week 1:  $r = -0.34$ ; Week 4:  $r = 0.27$ ), and Mobility scores (Week 1:  $r = -0.34$ ,  $p < 0.01$ ; Week 2:  $r = -0.38$ ,  $p < 0.01$ ; Week 3:  $r = -0.24$ ). Higher levels of walking or running minutes were associated with better scores on PROMIS Pediatric measures of Asthma Impact and Mobility at all four weekly time points. **Conclusion:** This study provides an exploratory evaluation of the association between physical activity intensity and HRQOL data among children with asthma. **Results:** indicate that children and adolescents with asthma who are more physically active have better HRQOL outcomes than those who spend more time in sedentary activities.

### **(3096) Challenges with using paper format legacy Patient-Reported Outcome Measures (PROMs) in digital real-world evidence (RWE) studies for apps**

Sujyotee Kretz, MSc in Technical, medical and scientific translation with translation technology, Vitaccess, London, United Kingdom; Cathy Spinage, Vitaccess, Oxford, United Kingdom; Catherine Bottomley, Master of Pharmacy, Vitaccess, Oxford, United Kingdom

**Aims:** Patient-Reported Outcome Measures are commonly designed to meet development, psychometric and scaling standards. Translations of these instruments undergo rigorous methodologies to ascertain equivalence. Once finalized, such versions are under the jurisdiction of the authors or license holders and known as legacy texts. The objective of this research is to explore the challenges and solutions associated with applying validated paper PROMs and their translated equivalents to digital RWE studies. **Methods:** Three separate RWE studies on a blood disorder, an inherited nerve disorder and a neuromuscular disease involving PROMS were used for this

research. Content analysis was done through the linguistic validation process, involving adaptation of the wording and structure of PROMs to fit the context, configuration and logic of the electronic format of study. This step is known as ePRO update. Quality Assurance reports, part of this step, ensured that only the required edits are made. **Results:** Using legacy PROMs requires a number of upfront considerations. Agreement had to be reached on changes in the source through discussion with authors and app user interface and user experience developers. ePRO edits included changing wordings such as “tick” to “select,” reconfiguring the PRO structure and reformulating the placement of wordings to suit the app capabilities and logic. More complexities arose when updates on legacy fell outside of the realm of the agreed changes for reasons such as translations not corresponding to the source, typographical, grammatical errors and form of address not consistent with other parts of the app. Finally, the requirement to maintain instrument validity by using the same instrument wording when moving to an electronic version potentially meant that the app content was not optimum from a user experience perspective. **Conclusion:** Collaboration with PROM authors plays a crucial part in reaching an agreement for the ePRO content. If the translated instrument is inconsistent with the paper to electronic changes in source, this can have an impact on content validity, and subsequently data pooling and analysis. A detailed rationale for changes needs to be provided to the author to enable them to make an informed decision on legacy updates.

### **(3097) COVID-19: facilitating a paradigm shift in patient disease experience research**

Jane Wells, Adelphi Values Ltd, Bollington, United Kingdom; Dale Chandler, Adelphi Values Ltd, Bollington, United Kingdom; Sophie Tatlock, Adelphi Values Ltd, Bollington, United Kingdom

**Rationale:** Patient experience research methods across numerous therapeutic areas are rapidly evolving due to growing recognition of the importance of the patient voice. In parallel, global healthcare systems have been evolving, to improve efficiencies using distance or e-medicine. The global COVID-19 pandemic has provided an immediate and further need to adapt patient experience research methods. **Objective:** We examine the paradigm shift in healthcare provision, expedited by COVID-19, to discuss how patient experience research can incorporate largely existing technologies to adapt to change, while maintaining safety and methodological robustness. **Discussion:** is informed by FDA COVID-19 guidance, patient-focused drug development (PFDD) guidance and learnings from case studies. **Discussion:** The need to minimize healthcare provider (HCP) workload in patient experience research (e.g., identifying patients, providing confirmed diagnosis) is heightened due to COVID-19. Time restrictions already exist for in-person consultations, restricting opportunities for participant recruitment. Furthermore, a paucity of patients presenting to HCPs due to COVID-19 highlights the need for HCPs to actively identify patients for recruitment. Alternative consultation mediums taking place due to COVID-19, including telephone and web-based platforms, can limit HCP workload while eliminating physical contact. Such platforms may be leveraged to identify patients for research. Well managed patient self-referral may be suitable and in-line with regulatory guidance to encourage active patient engagement. Patient consenting and the handling of study materials can be conducted remotely. FDA and ethical review board guidance recommend use of e-consent and secure electronic transfer of materials where possible; this has proven to be efficient. Evidence shows that telephone/teleconference interviews/focus groups, typically conducted in patient experience research, sacrifice nothing or

little in data quality. App-based data collection provides opportunity for patients to provide multi-media and event driven data, while in their own environment. A shift to electronic mediums may in some populations make participation more accessible, mitigating time and monetary costs of travel and possibly reduce participation biases. Conclusion: COVID-19 has required researchers to adapt and to use technology to conduct safe, efficient and robust research. This necessary use of technology may lead to a paradigm shift and increase the use of e-recruitment, e-consent and electronic data collection.

### **(3098) Developing patient-reported outcome measures of ‘medical technology experiences’ for children with medical complexity**

Nora Fayed, Queen’s University, Kingston, Ontario, Canada; Samantha Noyek, Queen’s University, Kingston, Ontario, Canada; Nathalie Major, The Children’s Hospital of Eastern Ontario, Ottawa, Ontario, Canada; Julia Orkin, SickKids Hospital, Toronto, Ontario, Canada; Michael Sidra, Alberta Health Services, Edmonton, Alberta, Canada; Eyal Cohen, SickKids Hospital, Edmonton, Ontario, Canada

**Aims:** Our objective was to develop a tool to assess this outcome in the form of a patient-reported outcome measure (PROM). **Methods:** A mixed-method design was used to develop the PROM. In Phase I, a qualitative study generated candidate items for quantitative validation. Primary caregivers of CMC, refined the language and appropriateness of the items based on criteria of i) importance, ii) clarity/comprehension and iii) ethical acceptability. In Phase II, pilot validation of the factor structure of the measures was tested with a convenience sample of self-referred caregivers of CMC across Canada. The factor structure of the tools was determined using exploratory factor analysis and tested using internal consistency and correlations to related measures. **Results:** Phase I, caregivers ( $n = 32$ ) of CMCs aged between 18 months to 18 years old (mean age = 6.26) were interviewed. All children were enteral fed, G-tube ( $n = 25$ ), with pump ( $n = 12$ ), GJ-tube ( $n = 5$ ), J-tube ( $n = 1$ ), NG-tube ( $n = 1$ ). Other medical technology included suction ( $n = 15$ ), oxygen ( $n = 8$ ), oximetre ( $n = 6$ ), cough assist ( $n = 6$ ), nebulizer ( $n = 6$ ), ventilator ( $n = 3$ ). The subscales were: general experiences, child perspectives, impact to child, family impact, sleep, caregiver’s technical confidence, cost, supplies, daycare/school support, health care provider support, community mobility, peers involvement/acceptance, public acceptance. In phase II, the factor structure was initially validated with a convenience sample distinct from Phase I ( $n = 39$ ), with Cronbach’s  $\alpha$  ranging 0.76–0.89 on all subscales. **Conclusion:** Development of PROMs for CMC and initial pilot testing are complete. The factor structure indicates that they can be used to promote provider to caregiver communication through measurement in their current form. Use in tertiary care and home care settings is needed to validate the tools for quality improvement, benchmarking, clinical relevance, and testing intervention effectiveness.

### **(3099) Addressing future implementation challenges of electronic clinical outcome assessment (eCOA) libraries**

Denise Dauria, Medidata Solutions, New York, New York, United States; Paul O’Donohoe, Medidata Solutions, London, United Kingdom; Megan Turner, GlaxoSmithKline, Collegeville, Pennsylvania, United States; Caroline Anfray, Mapi Research Trust, Lyon, France; Marie-Pierre Emery, Mapi Research Trust, Lyon, France

**Aims:** Those who have implemented COAs electronically have long imagined a repository of assessments that is stored and seamlessly reused from study to study, commonly referred to as an “eCOA Library.” As eCOA libraries shift from concept to reality, it is important to review the current major implementation challenges facing eCOA studies and determine the impact libraries will have on these hurdles. **Methods:** Analysis of feedback from COA developers and sponsors on eCOA development as well as the authors’ experience in implementing eCOAs using various modalities. **Results:** Reviewing industry feedback, today’s major eCOA implementation challenges are categorized as: high start-up costs, lengthy implementation timelines, varying copyright requirements, and limited stakeholder experience. Using an eCOA library may directly result in reductions to both study-specific costs and timelines. However, new costs for maintenance of the library and the learning curve of technical specialists, depending on implementation complexity, may be introduced as a result. While some copyright holders are receptive of eCOA libraries as a way to legally and technically streamline processes, others will still require independent licensing and screenshot review processes. This will most likely result in some studies using an eCOA library and traditional implementation methods in parallel. Developing specific processes to implement COAs into eCOA libraries will ensure that the COAs’ integrity is maintained and avoid redundancy in screenshot review and usability testing. Finally, the efficiencies gained from eCOA libraries seem to suggest that less eCOA-specific knowledge will be needed during the implementation cycle. Yet, it appears unclear if additional knowledge transfers and training would be needed to account for eCOA set-up becoming more automated. **Conclusion:** It seems fair to hypothesize that eCOA libraries will have obvious implementation benefits and that these will mostly outweigh new challenges introduced. Once the industry begins to more regularly deploy studies from eCOA libraries, it is recommended that further research take place to measure impact on each of the current major challenges. However, acknowledging eCOA implementation challenges are reduced, but not eliminated, when using an eCOA library ensures that realistic expectations are set as vendors and sponsors begin using this new technology.

### **Instrument Development & Evaluation I**

#### **(3100) The evaluation of individual patient-level change in rheumatoid arthritis: do current estimates reflect a meaningful improvement for patients?**

Kate Williams, PhD, Acaster Lloyd Consulting, London, United Kingdom; Peter C Taylor, MA, PhD, FRCP, FRCPE, Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, United Kingdom; Robin K Dore, MD, Robin K. Dore Inc, private practice, Tustin, California, United States; Jenya Antonova, M.S., Ph.D., Gilead Sciences Inc, Foster City, California, United States; Mark Genovese, MD, Gilead Sciences Inc, Foster City, California, United States; Sarah Acaster, MSc, Acaster Lloyd Consulting, Foster City, United Kingdom

**Aims:** Rheumatoid arthritis (RA) is a chronic inflammatory disease, characterized by pain and fatigue. Janus kinase inhibitors (JAKis) are treatments for RA that can alleviate symptoms and improve health-related quality of life (HRQoL). Patient-reported outcomes (PROs) are widely used to evaluate outcomes in RA clinical trials. Determining the level of change on a PRO that is meaningful to patients enables the interpretation of whether a treatment is effective from a patient perspective. Although the minimum (clinically) important

difference (MID/MCID) offers meaningful interpretation of group differences on a PRO, the same MID/MCID values may not be sufficient to interpret individual patient-level change. This study aimed to determine how individual patient-level change has been evaluated in RA clinical trials of JAKis and to evaluate the methods used to derive these change thresholds. Methods: A targeted literature search was conducted to identify clinical trials of JAKis reporting PROs in RA. Relevant papers were reviewed to identify values used to evaluate individual patient-level change on PROs. Follow-up literature searches were conducted to identify papers reporting the methodology used to derive these thresholds, as well as alternative thresholds. A critical appraisal of the methodology used to derive these thresholds was conducted. Results: Among 102 papers meeting the inclusion criteria, 35 reported the proportion of patients achieving an MID/MCID on PROs (Health Assessment Questionnaire Disability Index (HAQ-DI), Pain Visual Analogue Scale (VAS), Patient Global Assessment (PtGA), Short-Form 36 (SF-36), Functional Assessment of Chronic Illness–Fatigue (FACIT–Fatigue), Insomnia Severity Index (ISI), morning stiffness severity and duration). Several papers applied MID/MCID thresholds to individual patient-level change, but we found no evidence to support their application in this way (Table 1). We identified alternative thresholds for individual patient-level change for some PROs, but the methods used to derive these were not in line with established standards. Conclusion: The MID/MCID is commonly used as a threshold for meaningful individual patient-level change in PRO's in JAKi clinical trials, but there is no evidence that these thresholds represent a meaningful improvement for patients. Further research is needed to identify the level of change on PROs that is meaningful to patients, to support future patient-level analyses in RA.

### **(3101) Modification of an existing patient-reported outcome measure (PROM) for a new context of use in a seasonal infectious disease**

Todd Saretsky, Merck & Co., Inc., Kenilworth, New Jersey, United States; Carla DeMuro, RTI Health Solutions, Research Triangle Park, North Carolina, United States; Lyn Finelli, Merck & Co., Inc., Kenilworth, New Jersey, United States; Matthew Phillips, Merck & Co., Inc., Kenilworth, New Jersey, United States; Sandy Lewis, RTI Health Solutions, Research Triangle Park, North Carolina, United States; Emily Martin, University of Michigan School of Public Health, Research Triangle Park, Michigan, United States; Richard Osborne, Measured Solutions for Health Pty Ltd, Alphington, VIC, Australia; Josephine Norquist, Merck & Co., Inc., Kenilworth, New Jersey, United States

Aims: FDA's patient-focused drug development draft guidance suggests a decision process to determine the need for de-novo instrument development and proposes methods to support modification of existing instruments that conform to good measurement principles. Papadopoulos et al. (2020) also described guidelines for evaluating content validity of an existing clinical outcome assessment for a new context/target patient population; including conceptual match, input from target population, instrument content, and modifications. These methodologies can be especially beneficial when developing instruments in areas where there may be a scarcity of patients available for instrument development activities, such as in rare disease populations or diseases impacted by seasonality with limited patient access. Here we describe steps involved in modifying an existing influenza symptom assessment instrument (Osborne, 2011), for use in individuals with RSV infection, a seasonal infectious disease. Methods: As a core constellation of symptoms is common to both RSV and influenza, the Influenza Intensity and Impact Questionnaire (Flu-iiQ),

a well-validated 25-item PROM, was selected for modification; informed by conceptual mapping of instrument content to input from three clinicians and available literature. Combined concept elicitation (CE)/cognitive debriefing (CD) interviews with laboratory-confirmed RSV patients were conducted to confirm content validity and optimize instrument wording. Results: Concepts clinically important in describing RSV infection but not included in the Flu-iiQ were added (e.g., wheezing, cough with phlegm, shortness of breath) and irrelevant items were removed (e.g., neck pain and impact on others). Findings from 20 RSV patient interviews (aged: 26-78; 70% female) confirmed the draft RSV Infection, Intensity and Impact Questionnaire (RSV-iiiQ) items accurately reflected the patient experience of RSV infection symptom severity and impact. Only minor revisions to question wording were suggested. The RSV-iiiQ consists of 29 items with four hypothesized domains: respiratory and systemic symptoms, and functional and emotional impact. Conclusion: Creation of the RSV-iiiQ is aligned with recent recommendations for modifying existing PROMs. The RSV-iiiQ is fit for purpose in an adult RSV patient population and incorporates relevant items from the Flu-iiQ. Modification offered efficiencies through conduct of split CE/CD interviews reducing sample size, limited required new item generation, and reduced timelines. Preliminary psychometric evaluation is underway.

### **(3102) Design of a treatment satisfaction measure for patients undergoing varicose vein treatment: the Venous Treatment Satisfaction Questionnaire (VenousTSQ)**

Charlie Gilbride, PhD, Health Psychology Research Ltd, Egham, United Kingdom; Janet Bayfield, Royal Holloway University of London, Egham, United Kingdom; Alison Wilson, Health Psychology Research Ltd, Egham, United Kingdom; Anita Bradley-Gilbride, PhD, Health Psychology Research Ltd, Egham, United Kingdom; Kathleen Gibson, MD, Lake Washington Vascular, Bellevue, Washington, United States; Manjit Gohel, MD, Addenbrooke's Hospital, Bellevue, United Kingdom; Clare Bradley, PhD, Royal Holloway University of London, Egham, United Kingdom

Aims: Varicose veins are common in Western Europe and the USA. Often painful and causing cosmetic concerns, the underlying condition of venous insufficiency, if untreated, may have serious complications. As options for treatment have increased, the need for a patient satisfaction measure has also grown. We describe the design of a varicose vein treatment satisfaction measure based on existing templates for other conditions including the widely used Diabetes Treatment Satisfaction Questionnaire (DTSQ © Bradley), our existing -TSQ Item Library, and input from patients in the UK and USA with experience of various varicose vein procedures. Methods: Collaborating vascular surgeons—from the UK (MG) and USA (KG)—identified relevant items from our -TSQ Item Library and suggested new items they felt were important for patients with varicose veins. Following clinician feedback and our literature review, we prepared a draft VenousTSQ using the existing AneurysmTSQ (© Bradley) as a template. Ten patients were recruited from Addenbrooke's Vascular Unit (Cambridge, UK) and 4 patients from Lake Washington Vascular (Washington, USA). Interviews, conducted between 4 days and approximately 18 months post procedure, elicited experiences of treatment and sources of satisfaction/dissatisfaction prior to VenousTSQ completion. The VenousTSQ was modified between sets of interviews until no further changes were required. Results: The VenousTSQ consists of two parts: The VenousTSQ–early (VenousTSQe) asks patients about experiences around the time of the treatment procedure and is intended to be administered only once up to one month post-procedure. The VenousTSQ–status (VenousTSQs)

asks patients about potentially ongoing aspects of treatment and is designed for use at multiple time points. Of the 16 unique items forming the VenousTSQ, 12 were from our Item Library. Only one required significant modification. Conclusion: The Item Library facilitated the design of the VenousTSQ with novel items ensuring that all important aspects of (dis)satisfaction with varicose vein treatments are covered. The early version of the questionnaire for single use together with a status measure for repeated use is an innovation in -TSQ design. Large-scale data collection is underway to allow for psychometric analyses to determine optimal scoring and establish validity and reliability of the VenousTSQ.

**(3103) Devising and testing a scoring algorithm for the Food Allergy Quality of Life Questionnaire (FAQLQ) across multiple forms**

Helen Doll, MSc, DPhil, Clinical Outcomes Solutions, Folkestone, Arizona, United Kingdom; Cristian Gugu, Clinical Outcomes Solutions, Chicago, Illinois, United States; Libby Floden, PhD, MPH, Clinical Outcomes Solutions, Tucson, Arizona, United States; David Norval, Aimmune Therapeutics, London, United Kingdom; Britanni Wayne, Clinical Outcomes Solutions, Chicago, Illinois, United States; Shengsheng Yu, Aimmune Therapeutics, Chicago, California, United States; Audrey Dunn-Galvin, University College, Cork, United Kingdom; Sarah Donelson, Aimmune Therapeutics, Brisbane, California, United States

**Aims:** To develop and examine the performance of a revised scoring algorithm for the multi-domain and multi-form Food Allergy Quality of Life Questionnaire (FAQLQ) administered in three randomized controlled trials (PALISADE, ARC003,  $n = 545$ ; RAMSES, ARC007,  $n = 493$ ; ARTEMIS, ARC010,  $n = 168$ ) aimed at reducing allergic reaction in Peanut Allergy (PA). **Methods:** The FAQLQ includes specific forms for children (8–12 years; FAQLQ-CF), teenagers (13–17 years; FAQLQ-TF), adults (18 + years; FAQLQ-AF), and parent reports for children (FAQLQ-PF) and teenagers (FAQLQ-PFT). The standard scoring algorithms are inconsistent across forms. Domain scores are calculated by either dividing the total item score by the number of items with non-missing responses or, for FAQLQ-PF, by the total number of items. The Allergen Avoidance (AA) and Dietary Restrictions (DR) domains are scored separately in FAQLQ-CF but combined in FAQLQ-AF and FAQLQ-TF; the DR and Social Restrictions (SR) domains are scored separately in FAQLQ-PFT but combined in FAQLQ-PF. Total scores are also calculated differently. In the revised algorithm, FAQLQ-CF, FAQLQ-TF, FAQLQ-AF were combined into a FAQLQ-CTAF form after common items showed no or minimal differential item functioning; all domain scores were calculated with reference to non-missing items; the AA and DR domains in FAQLQ-CF and SR and DR domains in FAQLQ-PFT were combined; the Total score was calculated as the mean domain score. Spearman correlation coefficients assessed association between and within standard and revised domain scores. **Results:** Correlations between domain scores from the two scoring algorithms ranged from 0.879 to 1.00 (Table 1). This was supported by domain correlations within the standard scoring algorithm in which correlations between domains within the same form were higher than those between forms. Strong ( $r = 0.804$ ) and adequate ( $r = 0.549$ ) support was found for combining AA and DR in FAQLQ-CF, and SR and DR in FAQLQ-PFT domains, respectively (Table 2). **Conclusion:** Findings from this study suggest the revised scoring algorithm is appropriate for analyzing FAQLQ data in PA. The standard and revised domain scores were consistent, supporting the combination of domains and the use of the number of completed items to calculate domain scores, a less-biased approach when there are missing items in a data set.

**Table 1: Correlations<sup>[1]</sup> between the Standard<sup>[2]</sup> and Revised<sup>[3]</sup> FAQLQ<sup>[4]</sup> Domain Scores at Baseline**  
Population: Pooled PRP<sup>[5]</sup>

Domain	Correlation
<b>Combined Child-Teenager-Adult Forms (CTAF)</b>	
Allergen Avoidance / Dietary Restrictions (AADR-CTAF)	1.000
Allergen Avoidance (AA-CTAF)	0.957 with AADR-CTAF
Dietary Restrictions (DR-CTAF)	0.930 with AADR-CTAF
Emotional Impact (EI-CTAF)	1.000
Food Allergy-related Health (FAH-CTAF)	1.000
Risk of Accidental Exposure (RAE-CTAF)	1.000
<b>Total Score</b>	1.000
<b>Parent Form (PF)</b>	
Emotional Impact (EI-PF)	0.966
Food Anxiety (FA-PF)	0.988
Social and Dietary Limitations (SDL-PF)	0.999
<b>Total Score</b>	1.000
<b>Parent Form – Teenager (PFT)</b>	
Emotional Impact (EI-PFT)	1.000
Food Anxiety (FA-PFT)	1.000
Social Restrictions (SR-PFT)	0.893 with SDL-PF
Dietary Restrictions (DR-PFT)	0.879 with SDL-PF
<b>Total Score</b>	1.000

Abbreviations: PRP = Patient-reported outcome population, ITT = Intent-to-treat population, PRO = Patient-reported outcome, CTAF = Unified Children-Teenager-Adult Form, PF = Parent Form, PFT = Parent Form – Teenagers

[1] A Spearman rank correlation was calculated for each domain. Since the revised scoring algorithm combined two sets of domains (AA-CTAF and DR-CTAF into AADR-CTAF; SR-PFT and DR-PFT into SR-PFT), correlations for each set are presented for the separate domains (standard) with the combined (revised) domains. Hence, these correlations only capture the impact of how missing values were scored. Based on Cohen's recommendations, the following values represent the magnitude of association: small = 0.1, moderate = 0.3 and large = 0.5.

[2] For the standard scoring, domain score = sum of item scores / number of completed items, except for the FAQLQ-PF in which domain score = sum of item scores / number of items. The Allergen Avoidance and Dietary Restrictions domains are scored separately for the CF form but together in the AF and TF forms; the Dietary Restrictions and Social Restrictions domains are scored separately in the PFT but together in the PF. The Total score = the sum of all items / number of completed items, except for the FAQLQ-PF in which Total score = sum of domain scores / number of domains.

[3] For the revised scoring, domain score = sum of item scores / number of completed items. The Allergen Avoidance and Dietary Restrictions domains are scored together in the CF, AF, and TF forms; the Dietary Restrictions and Social Restrictions domains are scored together in the PFT and PF. The Total score = sum of the domain scores / number of domains.

[4] The FAQLQ measures an individual's HRQL using data from the patient or by the child's parent. Statements are rated using a 7-point scale, from 1 = 'not' to 7 = 'extremely', where a higher score indicates a lower quality of life due to the impact of food allergy.

[5] The PRP (N=1206) is a subset of the FAP (ITT) and included subjects who completed at least one PRO measure at Baseline.

**Table 2: Convergent and Divergent Validity<sup>[1]</sup> Correlations<sup>[2]</sup> for the FAQLQ<sup>[3]</sup> Domain Scores at Baseline**  
Population: Pooled PRP<sup>[4]</sup>, Scoring<sup>[5]</sup>: Standard Scoring

Domain	AADR-CTAF	AA-CTAF	DR-CTAF	EI-CTAF	FAH-CTAF	RAE-CTAF	EI-PF	FA-PF	SDL-PF	EI-PFT	FA-PFT	DR-PFT	SR-PFT
AADR-CTAF	1.000												
AA-CTAF	0.804	1.000											
DR-CTAF	0.701	0.576	0.637	1.000									
EI-CTAF	0.573	0.704	0.738	0.843	1.000								
FAH-CTAF	0.755	0.704	0.738	0.695	0.766	1.000							
RAE-CTAF	0.412	0.447	0.404	0.447	0.404	0.447	1.000						
EI-PF	0.404	0.385	0.385	0.404	0.385	0.404	0.872	1.000					
FA-PF	0.372	0.385	0.372	0.385	0.372	0.385	0.815	0.759	1.000				
SDL-PF	0.583	0.592	0.583	0.483	0.495	0.495	0.790	0.790	0.790	1.000			
EI-PFT	0.592	0.592	0.592	0.504	0.492	0.492	0.679	0.679	0.679	0.691	1.000		
FA-PFT	0.491	0.491	0.491	0.319	0.371	0.371	0.679	0.691	0.679	0.691	0.679	1.000	
DR-PFT													1.000

**Table 2: Convergent and Divergent Validity<sup>[1]</sup> Correlations<sup>[2]</sup> for the FAQLQ<sup>[3]</sup> Domain Scores at Baseline**  
Population: Pooled PRP<sup>[4]</sup>, Scoring<sup>[5]</sup>: Standard Scoring

SR-PFT	0.524	0.443	0.508	0.763	0.691	0.549	1.000
SR-PFT	0.524	0.443	0.508	0.763	0.691	0.549	1.000

Abbreviations: PRP = Patient-reported outcome population, FAS = Full analysis set population, ITT = Intent-to-treat population, PRO = Patient-reported outcome, CTAF = Unified Children-Teenager-Adult Form, PF = Parent Form, PFT = Parent Form – Teenagers, AA = Allergen Avoidance, DR = Dietary Restrictions, AADR = Allergen Avoidance / Dietary Restrictions, EI = Emotional Impact, FAH = Food Allergy-related Health, RAE = Risk of Accidental Exposure, FA = Food Anxiety, FA = Social and Dietary Limitations.

Note: Sources of convergent validity are highlighted in yellow. Non-highlighted areas are sources of divergent validity. Dashes (–) denote correlations that could not be estimated because of how the domains did not co-occur due to differences between the domain structure across the five FAQLQ forms.

[1] Convergent validity is the degree to which an instrument is associated with other instruments that measure the same construct whereas divergent validity is the degree to which an instrument is less associated with other instruments that do not measure the same construct.

[2] A Spearman rank correlation was calculated for each domain. Based on Cohen's recommendations (1988), the following values represent the magnitude of association: small = 0.1, moderate = 0.3 and large = 0.5.

[3] The FAQLQ measures an individual's HRQL using data from the patient or by the child's parent. Statements are rated using a 7-point scale, from 1 = 'not' to 7 = 'extremely', where a higher score indicates a lower quality of life due to the impact of food allergy.

[4] The PRP (n=1206) is a subset of the FAP (ITT) and included subjects who completed at least one PRO measure at Baseline.

[5] For the standard scoring, domain = total / number of completed questions, except for the FAQLQ-PF which = total / number of questions. The HRQL score = the sum of the domains / the number of domains.

**(3104) Scoping methods for developing cross-cutting core physical function outcome sets for sarcopenia and rare disorders**

Sara Shaunfield, PhD, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; John Devin Peipert, PhD, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; George J. Greene, PhD, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Cindy Nowinski, MD, PhD, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Jin-Shei Lai, PhD, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Jack Guralnik, PhD, MD, University of Maryland School of Medicine, Chicago, Maryland, United States; Ryne Carney, BA, Alliance for Aging Research, Washington, D.C., United States; Vanessa Boulanger, MSc, National Organization for Rare Disorders, Danbury, Connecticut, United States; Robert Chapman, BA, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Chelsea Perschon, BA, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Allison Seebald, BA, Northwestern University

Feinberg School of Medicine, Chicago, Illinois, United States; Allison Seebald, National Organization for Rare Disorders, Danbury, Connecticut, United States; David Cella, PhD, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States

**Aims:** To employ novel scoping methods for efficient data collection and identification of six (6) conditions for development of a patient-reported and performance-based physical function (PF) core outcome set for regulatory use in sarcopenia and rare disorder clinical trials. **Methods:** As part of a 2-year planning UG3 award from the US Food and Drug Administration (FDA), we are employing scoping literature reviews and patient interviews to select conditions for cross-cutting PF clinical outcome assessment (COA) development. First, drawing from stakeholder input, we identified 11 candidate conditions that have a measurable PF impact and address a perceived regulatory gap. Second, 11 scoping reviews, one per condition, and a measure scan are being conducted to explore PF impacts and assessments. Third, scoping interviews are being conducted with  $\geq 3$  patients representing each condition ( $N \geq 33$ ) to conduct targeted systematic exploration of PF limitations, severity, and quality of life importance. Interview data will be analyzed to identify PF similarities across conditions to support the selection of 6 model conditions. **Results:** Preliminary activities yielded selection of 6 sarcopenia-related conditions (heart failure, chronic obstructive pulmonary disease, advanced cancer, hip fracture, Parkinson's disease, osteoarthritis) and 5 rare disorders (facioscapulohumeral muscular dystrophy, idiopathic pulmonary fibrosis, systemic sclerosis, myositis, hepatocellular carcinoma). Scoping reviews and interviews are underway. Findings of these activities are being compiled and presented to stakeholders (patients, caregivers, industry, clinicians, FDA), whose input will be synthesized in a gap analysis summarizing needed development and validation work. The gap analysis will be presented to stakeholders to inform collaborative selection of 6 model conditions (3 sarcopenia-related, 3 rare disorders) to carry forward in the 3-year UH3 phase for refining and testing: (1) PF patient-reported outcome measures based on the PROMIS PF Item Bank and (2) PF performance outcome measures based on the NIH Toolbox and the Short Physical Performance Battery. **Conclusion:** In the UG3 phase, we are employing novel scoping methods to select conditions appropriate for cross-cutting PF COA development. Scoping methods are optimal for efficient data collection and analysis representing diverse contexts aiming to identify cross-cutting themes or outcomes. Limitations and future directions will be discussed.

### **(3105) Validation of the PROMIS Sleep Disturbance and Sleep-related Impairment item banks in the Dutch general population**

Caroline Terwee, PhD, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, Netherlands; Marianna Avetisyan, PhD, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, Netherlands; Martine Crins, PhD, Amsterdam Rehabilitation Research Center | Reade, Amsterdam, Netherlands; Leo Roorda, PhD, Amsterdam Rehabilitation Research Center | Reade, Amsterdam, Netherlands

**Aims:** The Patient-Reported Outcomes Measurement Information System (PROMIS®) v1.0 item banks 'Sleep Disturbance' (SD, 27 items) and 'Sleep-Related Impairment' (SRI, 16 items) were developed to measure self-reported aspects of sleepiness, sleep quality, and functional impact of sleep problems more efficiently and precisely than current instruments, by using Computerized Adaptive Testing (CAT). We validated these item banks in a Dutch general population. **Methods:** Participants in an internet panel completed both item banks. Unidimensionality, local dependence, monotonicity, Graded

Response Model (GRM) item fit, Differential Item Functioning (DIF) for age, gender, education, region, ethnicity, and language (Dutch compared to US Wave 1 data), and reliability were assessed. **Results:** A representative Dutch sample of 1006 people participated. We found sufficient unidimensionality for the both item banks (SD: CFI = 0.93, TLI = 0.92, RMSEA = 0.13, Omega H = 0.85, ECV = 0.76; SRI: CFI = 0.96, TLI = 0.95, RMSEA = 0.17, Omega H = 0.85, ECV = 0.76). Some local dependence was found (SD: 4.8%; SRI: 0.8% item-pair correlations  $> 0.20$ ), sufficient monotonicity (SD: H = 0.60; SRI: H = 0.65), and good IRT item fit (SD: zero out of 27 items with S-X2 p value  $< 0.001$ ; SRI: two out of 16 items). For SD DIF for age was found for four items; younger persons report more sleep problems compared to older persons with similar levels of sleep disturbances. No DIF was found for SRI. We found a reliability of at least 0.90 with simulated CATs (based on US item parameters) in 96% of the participants with on average 4.5 (range 2–12) items for SD and 75% of the participants with on average 6.2 (range 3–12) items for SRI. **Conclusion:** The PROMIS sleep item banks showed sufficient psychometric properties in a general Dutch population and can be used as CAT. PROMIS CATs allow reliable and valid measurement in an efficient and user-friendly way with limited administration time.

### **(3106) Development of a short version of a weight-specific quality of life instrument (YQOL-W-SF)**

Li Zhang, Zhejiang University, Hangzhou, China; Yingjing Zheng, Zhejiang University, Hangzhou, China; Ying Chen, Zhejiang University, Hangzhou, China; Boyan Chen, Zhejiang University, Hangzhou, China; Rui Zhou, Zhejiang University, Hangzhou, China; Hao Chen, Zhejiang University, Hangzhou, China; Todd C Edwards, University of Washington, Seattle, Washington, United States; Donald L Patrick, University of Washington, Seattle, Washington, United States; Hongmei Wang, Zhejiang University, Hangzhou, China

**Aims:** To develop a short form of the Youth Quality of Life Instrument-Weight Module, the YQOL-W-SF, with the aim of preserving the conceptual model and, as far as possible, the content and the psychometric properties of the original instrument. **Methods:** A two-phase developing procedure was used: (1) Shortening Phase. An expert panel with 10 professionals and a youth group with 30 teenagers were organized to collect relevant opinions about item screening. Based on the field survey data of the original instrument, each item was analyzed using both item response theory (IRT) and classic test theory (CTT) methods. (2) Validation Phase. The psychometric properties of the YQOL-W-SF were assessed in an independent sample of 630 teenagers with different weight states (normal weight, overweight, and obesity). **Results:** The shortened YQOL-W scale retained 8 items of the original instrument with at least two or more items per dimension. The CFA analysis showed preservation of the original three-dimensional structure. The YQOL-W-SF had acceptable internal consistency reliability with Cronbach's  $\alpha$  ranging from 0.71 to 0.92 and robust test-retest reliability with the ICCs all higher than 0.9. The content validity, convergent and discriminant validity, known groups validity were all demonstrated comparable to those of the original YQOL-W. The targeting analysis with Rasch model showed that the average location in each dimension were greater than 0, which indicated different level of ceiling effect. Subjects had the most difficulty obtaining high quality of life scores on item 9 (Worry what people saying) as it had the highest average threshold. In addition, there was an out-of-order threshold for each item of the YQOL-W-SF, suggesting that the option settings for the items need to be improved. **Conclusion:** The 8-item YQOL-W-SF, which can be improved in terms of item option setting for better, has

good psychometric properties and can be used for the measurement of weight related quality of life for children and adolescents in China.

**(3107) Design of an individualized questionnaire to measure the impact of cancer on quality of life: the CancerDQoL**

Michelle D. Taylor, PhD, DCLinPsy, Psychology Department, Royal Holloway, University of London, Egham, United Kingdom; Amy Garden, MSc, Psychology Department, Royal Holloway, University of London, Egham, United Kingdom; Jonathan Davidson, MSc, Psychology Department, Royal Holloway, University of London, Egham, United Kingdom; Charlie J. Gilbride, PhD, Health Psychology Research Limited, Egham, United Kingdom; Clare Bradley, PhD, Health Psychology Research Unit, Royal Holloway, University of London, Egham, United Kingdom

**Aims:** To develop a questionnaire to measure the impact of cancer on quality of life (QoL), which is needed to capture the effects of cancer and its treatments on QoL as opposed to health status and functioning, measured by existing commonly used questionnaires. **Methods:** Design of the Cancer-Dependent Quality of Life (CancerDQoL) questionnaire used the template of the Audit of Diabetes Dependent QoL (ADDQoL© Bradley) questionnaire and library of items from -DQoLs for all conditions. Sharing a common template, -DQoL questionnaires specify life domains and respondents rate applicable items for the impact of their condition and importance for their QoL. Item selection, face validity and content validity were established through clinician and patient ratings of the importance and relevance of 60 life domains from the -DQoL Item Library. Semi-structured interviews were conducted with 25 English-speaking participants with a range of different cancers attending a cancer center in Zimbabwe (age range: 25–78 years; 16 women, 9 men). Ten interviews were subsequently conducted with UK English-speaking participants with a range of cancer diagnoses attending Maggie's Centres in London and Dundee (age range 40–76; 5 women, 5 men) to adapt the instrument for UK use. **Results:** The first draft of the CancerDQoL contained 26 domain-specific items from the -DQoL Item Library. Participants required no new items. Mean domain scores from 25 Zimbabwean patients indicated that cancer negatively impacted on all domains of life included, with the exception of 'fertility concerns.' Mean weighted impact scores were greatest for physical capability ('do physically'), 'enjoyment of food and drink' and 'dependency.' 'Fertility concerns' and 'feelings about past medical care/self-care' had the least mean weighted impact. UK interviews led to further refinement of the questionnaire and confirmed that no new items were required. One item, 'weight,' was removed because this domain was considered a physical symptom more appropriate for a symptom questionnaire. **Conclusion:** Face and content validity of the individualized CancerDQoL is established for an adult sample of English-speaking cancer patients in Zimbabwe and confirmed in an adaptation following UK interviews. Large-scale data collection with the CancerDQoL is now underway in the UK to enable psychometric evaluation.

**(3108) The SF-6D as a parsimonious method of operationalizing successful aging: proof-of-concept among people aging with HIV**

Nancy Mayo, McGill University, Montreal, Quebec, Canada; Marie-José Brouillette, MD, Department of Psychiatry, McGill University, Montreal, Quebec, Canada; Lesley Fellows, MD PhD, Montreal Neurological Institute, Montreal, Quebec, Canada; submitted on behalf of Brain Health Now Investigators

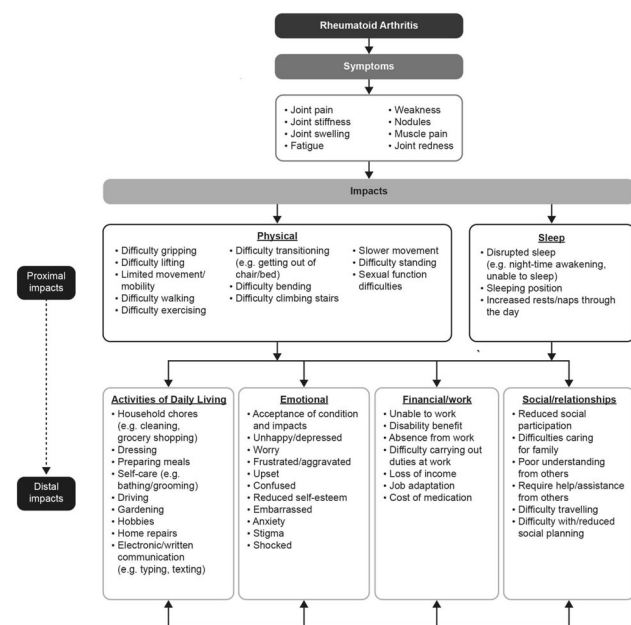
**Aims:** There is no standard way of operationalizing successful aging as attention is focused more often on the negative aspects of aging such as frailty, multi-morbidity, and cognitive impairment. Among people with chronic health conditions, successful aging can still be a goal but no measure exists for this important life outcome. The purpose of this study is to identify a way of operationalizing successful aging in people with HIV. A second objective was to identify factors that placed people with HIV at promise for successful aging **Methods:** Participants ( $\geq 50$  years) were from the Positive Brain Health Now (BHN) cohort which recruited from five Canadian sites (2014–2016) with ongoing follow-up. The first operational definition of successful aging was having 7 or 8 of 8 health-related quality of life subscales from the RAND-36 at norm or above. Logistic regression screened personal, biological, life-style, resilience, and environmental factors for inclusion in a regression tree model to identify promise factors. As the criterion approach (7 + subscales of SF-36 at norm of above) was too cumbersome for clinical use, regression tree analysis was applied to identify the fewest number of items that could be used. **Results:** Of the 536 people over the age of 50 at study entry, 77 (14.4%) met the 7 + criterion for successful aging at entry and over time. Self-reported cognitive ability was strongly associated with successful aging as were the variables related to the environment, resilience, social network, and motivation. The most important items of the RAND-36 for identifying successful aging were the 6 dimensions of the SF-6D. To match the 7 + criterion would require no more than 4 of the 6 domains with mild impairment or only 1 impaired domain. **Conclusion:** Successful aging could be identified directly using the SF-6D. To optimize successful aging, brain health, resilience, and a supportive environment contribute and are amenable to intervention.

**(3109) Development of the Rheumatoid Arthritis Symptom and Impact Questionnaire (RASIQ): a novel patient-reported outcome instrument for use in rheumatoid arthritis clinical trials**

Brandon Becker, GlaxoSmithKline, Upper Providence, Pennsylvania, United States; Marguerite Bracher, GlaxoSmithKline, Stevenage, United Kingdom; Deven Chauhan, GlaxoSmithKline, Uxbridge, United Kingdom; Regina Rendas-Baum, Optum, Johnston, Rhode Island, United States; Xiaochen Lin, Optum, Johnston, Rhode Island, United States; Kimberly Raymond, Optum, Johnston, Rhode Island, United States; Meaghan O'Connor, Optum, Johnston, Rhode Island, United States; Mark Kosinski, Optum, Johnston, Rhode Island, United States

**Aims:** Despite the availability of disease-modifying treatments, many patients with rheumatoid arthritis (RA) experience severe impacts, particularly from pain. Patient-reported outcome (PRO) measures provide insights into patient disease experience; however, instruments are needed that better reflect the aspects of RA that are most important to patients. We describe the development of the RA Symptom and Impact Questionnaire (RASIQ), a novel RA-specific PRO instrument developed in accordance with US FDA PRO guidance. **Methods:** A literature review identified key concepts relating to RA symptoms and impacts. Initial items were developed and reviewed by two key opinion leaders (KOLs) to evaluate relevance and wording and identify missing concepts. Concept elicitation interviews with patients with RA were conducted to confirm the concepts. Three rounds of cognitive debriefing (CD) interviews assessed the relevance, comprehensibility and recall period of the RASIQ. The measurement model was evaluated with extensive psychometric analyses using data from two phase II trials, followed by a final round of CD interviews. **Results:** The literature review identified

key RA symptoms and impacts, including pain, fatigue, and sleep disturbances, which formed the basis of the initial item pool. Following KOL review, the item pool was refined to 29 items (21 symptoms, 8 impacts); concept elicitation interviews ( $n = 30$ ) confirmed the relevance and importance of these concepts (Fig. 1). Thirteen items deemed redundant or irrelevant to the RA experience during initial CD interviews ( $n = 15$ ) were removed. Non-parametric item response theory and factor analyses showed a 3-factor model (Fig. 2) was optimal: Joint Pain (7 items), Joint Stiffness (4 items) and Impacts (5 items, capturing energy/tiredness, rest and sleep impacts). Further CD interviews ( $n = 12$ ) demonstrated that the final 16-item instrument contained relevant and understandable items and a suitable 24-h recall period. Conclusion: The RASIQ is a novel, patient-centric instrument with appropriate psychometric properties that evaluates key concepts important to patients with RA, particularly those related to pain, which many PRO instruments currently used in RA capture inadequately. The RASIQ may allow better evaluation of RA treatment needs. Studies funded by GSK (206981; 206577; HO-16-16897). Medical writing support provided by Liam Campbell, Fishawack Indicia Ltd, UK, funded by GSK.



**(3110) Health Personnel Perceived Quality of Life Scale alongside work-related quality of life scale and professional quality of life scale**

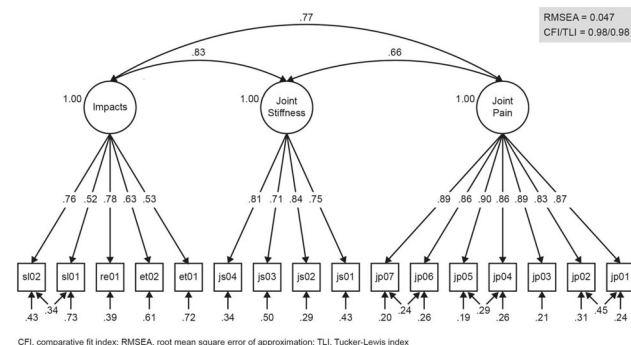
Joshua Chiroma Gandhi, Ph.D., University of Jos, Jos, Nigeria; Paul Samani Wai, PhD, Plateau State University, Bokkos, Nigeria; Dauda Akwai Saleh, MSc, Plateau State University, Bokkos, Nigeria; Beben Wokji Wukatda, MSc MHM, Plateau AIDS Control Agency, Jos, Nigeria

Aims: Health personnel are integral in determining patient-reported outcomes, therefore there’s need to determine health personnel perceived quality of life. Despite the cogent need, none of the extant instruments seem suitable enough to assess health personnel perceived quality of life in Nigeria. The study developed a Health Personnel Perceived Quality of Life Scale based on vitamin model analogy and psychometric theory. Methods: It adopted multifaceted item generation and pretesting techniques, anchored on modified Delphi process and cognitive testing interviews. A predictive correlational design, which provides empirical evidence about relationships among variables, was adopted. The study sample, who were selected by multi-stage sampling at 24 tertiary hospitals across Nigeria, consisted 3698 health personnel [1999 male (54.1%) and 1699 female (45.9%)]. Among these, 50.9% ( $x = 241.67$ ;  $SD = 22.725$ ) were nurses, 35.0% ( $x = 241.66$ ;  $SD = 22.586$ ) were medical doctors, and the remaining 14.81% consisted of other health personnel. Data were collected using the health personnel perceived quality of life scale alongside professional quality of life scale and work-related quality of life scale. Statistical analysis used item-total statistics, factor analysis, Pearson correlations and multiple regressions. Results: Results revealed significant alpha coefficients (raw = .964; standardized = .962). The reliability preceding validity was .964 while after validity was .971. Convergence test of correlation with professional quality of life scale [ $r = .33, p < .05$ ] and work-related quality of life scale [ $r = .98, p < .05$ ] established significant concurrent validity. Results also showed significant quality of life differences based on age [ $F = 4207.28, p < .05$ ]; marital status [ $F = 19, 681, p < .05$ ]; and ethnicity [ $F = 2, 326, p < .05$ ]. Principal component analysis extracted 2 factors from the 65 items, based on eigenvalue greater than 1 factor variance. From these items, 49.66% (factor 1) and 30.92% (factor 2) were retained while the remaining 18.42% (not loaded) were dropped. Conclusion: The finding implies that age, marital status and ethnicity are significant predictors of perceived quality of life among health personnel. It goes to suggest variation between health personnel perceived quality of life scale on one hand and (on the other hand) professional quality of life scale and work-related quality of life scale.

**Instrument Development & Evaluation II**

**(3111) Development and psychometric evaluation of the US English Wound-QoL questionnaire to assess health-related quality of life in people with chronic wounds**

Rachel Sommer, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Catharina von Stülpnagel, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Caroline Fife, The U.S. Wound Registry, The Woodlands, Texas, United States; Mandy Blasingame, The U.S. Wound Registry, The Woodlands, Texas, United States; Melissa Anders, CHI St. Luke’s Wound Clinic, The Woodlands, Texas, United States; Debi Thompson, CHI St. Luke’s Wound Clinic, The Woodlands, Texas, United States; Matthias Augustin, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Neuza da Silva, University Medical Center Hamburg-



CFI, comparative fit index; RMSEA, root mean square error of approximation; TLI, Tucker-Lewis index



Eppendorf, Hamburg, Germany; Christine Blome, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

**Aims:** People with chronic wounds perceive an impaired health-related quality of life (HRQoL). For the assessment of HRQoL, the Wound-QoL was originally developed and psychometrically validated for use in Germany. As the Wound-QoL is to be used across different countries, international validations are required. Therefore this study aimed to psychometrically validate the US English Wound-QoL version. **Methods:** Psychometric testing was performed based on cross-sectional data deriving from the US Wound Registry. The following clinical data were used as convergent criteria: pain intensity; total number of active wounds; surface area of the largest wound [cm<sup>2</sup>]; surface area of all active wounds [cm<sup>2</sup>]; malodorous exudate (yes/no). Therefore, the Wound-QoL global score was correlated with the criteria using Spearman's correlation. To evaluate internal consistency, Cronbach's alphas were determined. In addition, item response theory (IRT) analysis was used. A confirmatory factor analysis (CFA) was performed. The comparative fit index (CFI) and the root mean square error of approximation (RMSEA) were investigated as fit statistics for the CFA. **Results:** In total, 599 patients were included. Participants were aged between 18 and 95 years (M 63.7; SD 15.9) years, 47.4% female. The majority had venous leg ulcers (21.5%). Excellent internal consistency ( $\alpha > 0.9$  for the global score), low floor and ceiling effects of the subscales could be largely confirmed. Significant correlations were found between Wound-QoL global score and pain intensity ( $r = 0.40$ ), surface area of the largest wound ( $r = 0.25$ ), total surface area ( $r = 0.27$ ), number of active wounds ( $r = 0.14$ ) and malodorous exudate ( $r = 0.15$ ), indicating satisfactory convergent validity. The CFA produced a CFI value of 0.900 and a RMSEA of 0.092 and was considered as acceptable. IRT analyses could not fully confirm the factorial model underlying the Wound-QoL subscales. **Conclusion:** The Wound-QoL showed good psychometric properties, demonstrating reliability and validity. IRT analysis suggests that the instrument may benefit from future adaptations the factor structure developed for the original German version could not be fully confirmed. The Wound-QoL is a feasible tool for HRQoL measurement in people with chronic wounds in the US.

### (3112) Where do we need better assessment to evaluate treatment for hereditary angioedema (HAE)?

Milenka Jean-Baptiste, MPH, Evidera, Bethesda, Maryland, United States; Dylan Supina, PhD, CSL Behring, King of Prussia, Pennsylvania, United States; Robbin Itzler, PhD, CSL Behring, King of Prussia, Pennsylvania, United States; Subhransu Prusty, MD, CSL Behring, Marburg, Germany; Mona L. Martin, RN, MPA, Evidera, Seattle, Washington, United States

**Aims:** The patient experience with symptoms and the overall burden of their condition is changing with the development of new prophylactic treatment for Hereditary Angioedema (HAE), a genetic disorder characterized by episodic painful swelling attacks throughout the body. In recent years, new instruments have been developed to assess HRQoL in this patient population. While quality of life is an important construct, it is broad, and does not specifically address symptom frequency, severity, duration, type, and timing related to HAE attacks in a way that will support the assessment of patient status in response to new treatments. **Methods:** HAE specific HRQoL assessment tools currently in use (The 29- Angioedema Quality of Life Score (AE-QoL), the 25-item Hereditary Angioedema Quality of Life Score (HAE-QoL), the recently developed 25 item Hereditary Angioedema Association Quality of Life Score (HAEA-QoL), and the 12-item Hereditary Angioedema Activity Score (HAE-AS) were reviewed. The concepts covered were compared to the spectrum of

symptoms reported in the published literature and in 11 expert clinician interviews as those patients report to their doctors. **Results:** Published literature reports at least 21 different types of symptoms being reported to clinicians. The clinician interviews yielded 39 different symptoms patients report to their doctors, and suggested the new prophylactic treatments are changing the severity, duration and type of symptoms patients experience. The current assessment tools evaluate disease burden and HRQoL by assessing the impact of HAE on daily functioning, performance and feelings. None of the current measures address symptom severity, duration, frequency, or timing in relation to the HAE swelling episodes. **Conclusion:** Changes in function, feeling and quality of life are important aspects of patient status. However, the symptom characteristics that patients are reporting to their doctors as important and sometimes dramatic areas of change are quite possibly more sensitive constructs for assessing patient status. Without coverage of key symptom-specific changes in the patient experience of attack and non-attack HAE symptoms, the ability to adequately assess the true effect of new HAE treatments will be limited.

### (3113) Validity of the EQ-5D-Y in patients with asthma. ARCA study (Asthma Research in Children and Adolescents)

Karina Mayoral, Health Services Research Group, Hospital del Mar Medical Research Institute, Barcelona, Spain; Catalina Lizano, Health Services Research Group, Hospital del Mar Medical Research Institute, Barcelona, Spain; Angels Pont, Health Services Research Group, Hospital del Mar Medical Research Institute, Barcelona, Spain; Manuel Praena, Centro de Salud La Candelaria, Sevilla, Spain; Inés de Mir, Hospital Vall d'Hebron, Barcelona, Spain; Laura Valdesoiro, Hospital Parc Taulí, Barcelona, Spain; José Antonio Castillo, Hospital Miguel Servet, Zaragoza, Spain; Gimena Hernández, CAP Vila Olímpica, Parc Sanitari Pere Virgili, Barcelona, Spain; Eva Tato, Hospital Universitario Araba, Álava, Spain; Olga Cortés, Centro de Salud Canillejas, Madrid, Spain; Montse Ferrer, Centro de Salud Canillejas, Madrid, Spain; Montse Ferrer, Health Services Research Group, Hospital del Mar Medical Research Institute, Barcelona, Spain

**Aims:** The EuroQol Group developed the EQ-5D-Youth (EQ-5D-Y) for children aged 8 years or older as a derived version of the EQ-5D for adults. An EQ-5D-Y proxy version was also developed for children below 8 years of age. The aim of this study is to evaluate the validity of the EQ-5D-Y in children with asthma. **Methods:** Children from the ARCA study (an observational, longitudinal prospective multicenter study), aged 6-11 with a clinical diagnosis of persistent asthma. Patient-reported outcomes are collected principally through a mobile application including the EQ-5D-Y, the Pediatric Asthma Impact Scale (PAIS-PROMIS) and the Asthma Control Questionnaire (ACQ). The EQ-5D-Y measures 5 dimensions ("mobility," "looking after myself," "doing usual activities," "having pain/discomfort," and "feeling worried/sad/unhappy") with three-level Likert scale responses and a visual analogue scale to assess the general health (EQ-VAS). An equally weighted summary score was constructed with the 5 dimensions (range 0–100). The PAIS includes 8 items, providing a raw score, converted into a standardized score with a mean of 50 and a standard deviation of 10. The ACQ is composed of 5 items. Construct validity was evaluated through: 1) A multi-trait multi-method matrix between EQ-5D-Y and PAIS, constructed with Spearman correlations; and 2) Known groups comparisons based on the ACQ (well-controlled, intermediate, and not well-controlled). Construct validity hypotheses were stated a priori. **Results:** EQ-5D-Y was completed by 89 children, 61 were self-reported and 28 answered by proxy, 64% were male, with a mean age of 9.1 (1.8) years. The dimensions that showed higher percentages of participants with

problems were “doing usual activities” (30%) and “having pain/discomfort” (18%). Mean (SD) of the EQ-5D-Y summary score was 93.3 (9.5), with ceiling and floor effects of 57.3% and 0%, respectively. Results: of the multi-trait multi-method matrix for convergent validity confirmed six of the 9 relationships previously hypothesized as moderate-substantial (0.37, 0.39, 0.41, 0.44, 0.45, and 0.53). Statistically significant differences ( $p < 0.05$ ) between groups defined by the ACQ were found in the EQ-5D-Y summary score and EQ-VAS with large effect sizes (0.79, 1.04 respectively). Conclusion: These results support the EQ-5D-Y as a valid instrument for evaluating Health-Related Quality of Life in children with asthma.

**(3114) Construct validity of a new Japanese multiplicative, multi-attribute utility function and eight single-attribute utility functions for the Health Utilities Index Mark 3**

Shinichi Noto, Professor, Niigata University of Health and Welfare, Niigata, Japan; Takeru Shiroiwa, Center for Outcomes Research and Economic Evaluation for Health (C2H), Wako, Japan; Makoto Kobayashi, CRECON Medical Assessment Inc., Tokyo, Japan; Tatsunori Murata, CRECON Medical Assessment Inc., Tokyo, Japan; Ryota Izumi, Seirei Christopher University, Hamamatsu, Japan; Takashi Fukuda, Center for Outcomes Research and Economic Evaluation for Health (C2H), Hamamatsu, Japan

**Aims:** The Health Utilities Index Mark 3 (HUI3) is a generic multi-attribute, preference-based system for assessing health-related quality of life. It is widely used overseas as an outcome measure and for estimating quality-adjusted life years. We have published a new Japanese multiplicative, multi-attribute utility function and eight single-attribute utility functions for the Health Utilities Index Mark 3 in JPRO on 2020. Our aims are to test the construct validity of the new Japanese multiplicative, multi-attribute utility function and eight single-attribute utility functions for the HUI3. **Methods:** Data for this analysis are from the stroke rehabilitation outcome (SRO) study, multicenter study of subacute phase stroke rehabilitation ( $n = 526$ ) in Japan. Patient reported outcomes (PROs) in the SRO study were assessed using HUI3 and EQ-5D-5L. Concurrent validity and convergent validity were assessed by calculating the correlation between the HUI3 and the EQ-5D-5L and the Barthel Index (BI). Mean overall HUI3 scores and single scores were estimated and compared for groups classified according to the modified Rankin scale (MRS). **Results:** The average overall HUI3 score was  $0.31 \pm 0.25$ . A new Japanese multiplicative, multi-attribute utility function for the HUI3 showed good concurrent validity and convergent validity, with correlations between the EQ-5D-5L and BI values were 0.79 and 0.76, respectively. Mean overall HUI3 scores were as follows: MRS 1 = 0.69 (SD = 0.19), MRS 2 = 0.51 (SD = 0.20), MRS 3 = 0.32 (SD = 0.15), MRS 4 = 0.16 (SD = 0.11), MRS 5 = 0.03 (SD = 0.03). As a result, Japanese utility function for the HUI3 showed statistically significant differences between changes for patients categorized on MRS. Worst mean single score was ambulation (0.50) and, subsequently cognition (0.59), and dexterity (0.63). **Conclusion:** Findings support the good construct validity of new Japanese multiplicative, multi-attribute utility function and eight single-attribute utility functions for the HUI3.

**(3115) The EORTC QLU-C10D—development and investigation of general population utility norms for Canada, France, Germany, Italy, Poland and the UK**

Micha Johannes Pilz, Department of Psychiatry, Psychotherapy and Psychosomatics, University Hospital of Psychiatry II, Innsbruck,

Austria, Innsbruck, Austria; Sandra Nolte, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Medical Department, Division of Psychosomatic Medicine, Berlin, Germany, Berlin, Germany; Gregor Liegl, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Medical Department, Division of Psychosomatic Medicine, Berlin, Germany, Berlin, Germany; Madeleine King, School of Psychology, University of Sydney, Camperdown, New South Wales 2006, Australia, Sydney, Australia; Richard Norman, School of Public Health, Curtin University, Perth, Western Australia, Australia, Perth, Australia; Helen McTaggart-Cowan, Faculty of Health Sciences, Simon Fraser University, Burnaby, British Columbia (BC), Canada, Perth, British Columbia, Canada; Bernhard Holzner, Department of Psychiatry, Psychotherapy and Psychosomatics, University Hospital of Psychiatry II, Innsbruck, Austria, Innsbruck, Austria; Matthias Rose, Department of Psychosomatic Medicine, Center for Internal Medicine and Dermatology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany, Berlin, Germany; Georg Kemmler, Division of Psychiatry I, Department of Psychiatry, Psychotherapy and Psychosomatics, Medical University of Innsbruck, Innsbruck, Austria, Innsbruck, Austria; Andrew Bottomley, European Organisation for Research and Treatment of Cancer, Brussels, Belgium; Eva Maria Gamper, European Organisation for Research and Treatment of Cancer, Brussels, Belgium; Eva Maria Gamper, Department of Psychiatry, Psychotherapy and Psychosomatics, University Hospital of Psychiatry II, Innsbruck, Austria, Innsbruck, Austria

**Aims:** To enhance the interpretability of the EORTC QLU-C10D (Quality of Life Utility Core 10 Dimensions), a novel, cancer-specific utility measure based on the widely used EORTC QLQ-C30, we obtained general population norms for six countries, namely Canada, France, Germany, Italy, Poland, and the UK. **Methods:** We used data from a recent international online panel study conducted to develop norm data for the QLQ-C30 and calculated QLU-C10D utilities using the respective national value sets. To investigate the impact of country, sex and age we built a multilinear regression model, including these variables as main effects and accounting for interaction terms. Furthermore, we investigated correlations between utilities and country-specific socioeconomic parameters such as GDP per capita, unemployment rate and health expenditure per capita, retrieved from the WHO database. **Results:** The regression model showed a significant main effect for sex indicating lower QLU-C10D scores in women across all countries. The impact of age differed across countries: whereas in Canada and the UK, an overall increase of scores with age was observed, Poland and Germany decreased with age and no linear impact was found in France and Italy. After controlling for the impact of age and sex there were significant differences ( $p$  values  $\leq 0.045$ ) between countries except between Canada and the UK and France, France and Italy, and Italy and Poland. Overall country specific mean utilities range from 0.724 (SD 0.256) for the UK to 0.843 (SD 0.183) for Italy (highest reachable score would be 1 in all countries representing a current state of full health). Taking country, age and sex differences into account the subgroup with the lowest utility scores were 30 - 39-year-old male Canadians (mean 0.664; SD 0.307) and those with the highest scores were male Italians aged 70 + (mean 0.899; SD 0.128). No relevant correlations were found between QLU-C10D scores and socioeconomic indices. **Conclusion:** Results: showed a varying impact of age and sex on QLU-C10D scores and significant country differences that were not driven by age and sex. The use of national utility scores and reference values is recommended.

### (3116) Validation of the Ferrans and Powers Quality of Life Index for pregnant women in Brazil

Rosa Aurea Fernandes, RN, PhD, University of Guarulhos, Guarulhos, Brazil; Priscilla de Oliveira, RN, Master's Degree student, University of Guarulhos, Guarulhos, Brazil; No elle Freitas, RN, PhD, University of Guarulhos, Guarulhos, Brazil

**Aims:** To assess the psychometric properties of the Ferrans and Powers Quality of Life Index adapted version for pregnant women. **Methods:** Methodological study. The Brazilian version of the Ferrans and Powers Quality of Life Index for pregnant women has 36 items and four domains. Higher scores indicate better quality of life. Content validity was performed by experts committee. The content validity was assessed by means of content validity index (CVI) and the Content Validity Coefficient per item (CVCi). Convergent construct validity was assessed using Pearson's correlation between the scores of the adapted version of Ferrans and Powers Quality of Life Index with the WHOQOL-Bref. The internal consistency was assessed by Cronbach's alpha. Significance level was 0.05. **Results:** Five individuals participated in the expert committee. The value of the CVI was 0.94 and most of the items had a CVCi above 0.80. Participants at the validation step were 280 pregnant women at usual risk. Pearson's correlation between the total score of the Ferrans and Powers Quality of Life Index and the WHOQOL-Bref was positive and strong (0.79;  $p < 0.001$ ). Cronbach's alpha for the Ferrans and Powers Quality of Life Index adapted version was 0.94 for total score and ranged from 0.78 to 0.89 between the domains. **Conclusion:** The Brazilian adapted version of the Ferrans and Powers Quality of Life Index proved to be reliable and valid for use in pregnant women. Future study will carry out the factor analysis of the Ferrans and Powers Quality of Life Index. The Ferrans and Powers Quality of Life Index seems a promising tool for helping Brazilian health professionals and researchers to assess quality of life in pregnant women.

### (3117) Development of a novel, non-disease-specific pediatric faces scale

Shayna Egan, MPH, Endpoint Outcomes, Long Beach, California, United States; Martha Gauthier, MS, Endpoint Outcomes, Boston, Massachusetts, United States; Brooke Witherspoon, Endpoint Outcomes, Boston, Massachusetts, United States; Kayla Scippa, Endpoint Outcomes, Long Beach, California, United States; Christina Graham, Endpoint Outcomes, Boston, Massachusetts, United States; Julie Whyte, MPH, Endpoint Outcomes, Boston, Massachusetts, United States; Maggie Paulich, Endpoint Outcomes, Long Beach, California, United States; Aisling Ryan, Endpoint Outcomes, Boston, Massachusetts, United States; Siobhan McDonold, Endpoint Outcomes, Boston, Massachusetts, United States; Christopher Evans, PhD, Endpoint Outcomes, Boston, Massachusetts, United States

**Aims:** Reviews of pictorial faces scales used for pediatric health assessment confirm that simpler scales are needed for use with young children. Since research shows that visual prompts can aid children in describing their health, it is critical to establish suitable tools for children to use when rating their health status. A novel five-point verbal descriptor scale with corresponding face illustrations was developed for use with children. The objective of this research was to establish the content validity of the newly developed Pediatric Faces Scale within its intended age group in multiple disease areas and to make it available to all researchers. **Methods:** After IRB approval was obtained, children were recruited from various sites across the US for in-person interviews. Participants were presented with five faces and asked to order the faces based on severity, match each face to response options related to pain intensity (i.e., No hurt, A little hurt,

Medium hurt, A lot hurt, Worst hurt) or itching severity (i.e., Not at all itchy, A little itchy, Medium itchy, A lot itchy, Very itchy), and explain their rationale. **Results:** Thirty-three children aged 5–11 (average = 8.6, standard deviation = 1.7) participated in interviews. Comparable subgroups evaluated the pain ( $n = 16$ , 48.5%) and itch response options ( $n = 17$ , 51.5%) for the scale. Nearly all participants were able to correctly sort the faces in the hypothesized order ( $n = 29$ , 87.9%); the majority of participants ( $n = 19$ , 57.6%) correctly matched each response option to the corresponding face. Participants had the most difficulty matching the fourth response option, A lot [of hurt/itchy], to Face 4; however, almost two-thirds of participants ( $n = 21$ , 63.6%) correctly matched this response option to the face. Participants most commonly mixed up the two most severe response options ( $n = 9$ , 27.3%): A lot [of hurt/itchy] ( $n = 3$ ) and [Worst hurt/very itchy] ( $n = 6$ ). Three participants (9.1%) also interchanged the response options A little hurt/itchy and Medium hurt/itchy. **Conclusion:** This research demonstrates the Pediatric Faces Scale was well-understood by children aged 5–11. Future research should explore interpretation of the faces scale for concepts other than pain or itching. The scale is available for use in the public domain by all researchers without any license fees.

Figure 1. Pediatric Faces Scale



### (3118) Developing the Breast Utility Instrument from the EORTC QLQC30 and BR45 to measure patient preferences in breast cancer: confirmatory factor analysis

Teresa Tsui, PhD (cand), Toronto Health Economics and Technology Assessment (THETA) Collaborative, Toronto, Ontario, Canada; Maureen Trudeau, MD, FRCPC, Odette Cancer Centre, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada; Aileen Davis, PhD, Krembil Research Institute, Toronto Western Research Institute, Toronto, Ontario, Canada; Nicholas Mitsakakis, PhD, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada; Andrea Eisen, MD, FRCPC, Odette Cancer Centre, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada; Katarzyna Jerzak, MD, FRCPC, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada; Rossanna Pezo, PhD, MD, FRCPC, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada; Sonal Gandhi, MD, FRCPC, Odette Cancer Centre, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada; Ellen Warner, MD, FRCPC, Odette Cancer Centre, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada; Sofia Torres, MD, Multidisciplinary Oncologic Centre Antwerp, Edegem, Belgium; Danilo Giffoni, MD, Multidisciplinary Oncologic Centre Antwerp, Edegem, Belgium; Danilo Giffoni, MD, Odette Cancer Centre, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada; Lisa Verity, RN, Marion Soloway Rapid Diagnostic Unit, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada; Elizabeth Matheson, PA, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada; Kim Nguyen, PA, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada; Neda Stjepanovic, MD, Odette Cancer Centre, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada; Karen Bremner, BSc, Toronto Health Economics and Technology Assessment (THETA) Collaborative, Toronto, Ontario, Canada; Nim Li, BSc, Physical

Sciences, Sunnybrook Research Institute, Toronto, Ontario, Canada; Kelly Kim, BSc(Hon), Dept of Pharmacology & Toxicology, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada; Murray Krahn, MD, MSc, FRCPC, Toronto Health Economics and Technology Assessment (THETA) Collaborative, Toronto, Ontario, Canada

**Aims:** Breast cancer (BC) and its treatments impair patients' health-related quality of life (HRQoL). Utility is a measure of HRQoL that includes valuation or preferences for health outcomes, not simply their description, and is important in BC decision-making. Generic preference-based instruments lack items that represent BC-specific concerns, and mapping BC-specific psychometric instruments to generic preference-based instruments is unsatisfactory because it estimates generic utilities. Our overall objective is to develop and validate the novel Breast Utility Instrument (BUI), a BC-specific preference-based instrument, derived from the EORTC QLQ-C30 and BR45. **Methods:** A sample of 409 patients who represented the spectrum of BC disease stage were recruited from ambulatory BC clinics at a university-affiliated hospital in Canada. Patients completed the QLQ-C30 and BR45. After assessing the factorability of the data, we performed confirmatory factor analysis (CFA) of the instruments using mean-and variance-adjusted unweighted least squares estimation, based on the QLQ-C30 literature and clinically meaningful factors of the BR45. Residual correlations were iteratively applied between item pairs to improve global model fit. Clinician opinions were consulted to assess the face and content validity of the resulting factors. Overall, CFA evaluates the hypothesized factor (dimensional) structure, and helps to ensure there are low correlations between dimensions of the future BUI. **Results:** Based on global and item fit criteria, a ten-factor model of the combined QLQ-C30 and BR45 items demonstrated 1) good global fit by a non-significant Chi square statistic, 2) good incremental fit: Tucker Lewis Index = 0.958, Comparative Fit Index = 0.962, and 3) adequate parsimony-adjusted fit: root mean square error of approximation = 0.080. The factors in the best-fitted CFA model included: physical and role function, emotional function, social function, body image, pain, fatigue, systemic therapy side effects, endocrine sexual symptoms, arm and breast symptoms, and endocrine therapy symptoms. **Conclusion:** This CFA identified core factors (dimensions) that will inform the construction of the BUI. Our next step is to perform Rasch analyses and incorporate clinician and patient item-importance ratings to select representative items in each dimension for the BUI.

## Methodological Considerations for Quality of Life Research

### (3120) Identifying Frailty Using Self-Reports in a Healthy Aging Population

Mehmet Inceer, PhD, McGill University, Montréal, Quebec, Canada; Nancy Mayo, James McGill Professor, School of Physical and Occupational Therapy, McGill University, Montreal, Quebec, Canada; Jose Morais, Associate Professor, Division of Geriatric Medicine, McGill University, Montreal, Quebec, Canada

**Aims:** Fried's Frailty Phenotype is the predominant method to classify people as frail, requiring 3 of 5 criteria. Two criteria, exhaustion and physical activity, are self-reported; weight loss can be reported or measured; grip strength, and gait speed are performance-based. The latter are often replaced by self-report but the extent to which this provides the same interpretation is not known. To estimate the extent to which gait speed and grip strength can be substituted with self-reported items on limitation in walking and arm use to identify the added value of the inclusion of other personal and functional factors. **Methods:** The data for this study came from a longitudinal study on nutrition and successful aging (NuAge) based in Quebec, Canada. Gait speed and grip strength were directly measured and were classified as frail or not using cut points defined by Fried. Three

explanatory logistic regression models were developed for frail gait speed, one each for self-reported limitation in walking 100 m, 200 m, and 1 km. For frail grip strength, the model used self-reported limitation in lifting and carrying groceries. Variables were added stepwise to these base models for comorbidities, for SF-36 items related to pain, mood, fatigue, and health perception, and for cognition (cut-point score on Mini Mental State of Examination) until the highest prediction (c-statistic) was reached. Models with c-stat  $\geq 0.8$  are considered to have excellent prediction. **Results:** Of the 1754 participants (mean age:74; SD:4.2), 38 (2.2%) people were classified with frail gait and 375 (21.5%) with frail grip. The highest prediction for frail gait was using "limitation in walking one km" (c-stat:0.73 men; 0.79 women). For men prediction improved (c-stat = 0.80) by including the depression item, for women by including pain (c-stat = 0.80). Prediction of frail grip was moderate even with the inclusion of age, pain, and cognition for men (c-stat = 0.69) and age and health perception for women (c-stat = 0.67). **Conclusion:** Self-reported limitation in walking 1 km could substitute for measured gait speed; no SF-36 item could substitute for grip strength but other self-report items do exist which query this directly. Researchers should consider these items when designing survey studies involving older persons.

### (3121) One ruler to measure them all: combining data from multiple forms

Cristian Gugiu, Ph.D., Clinical Outcomes Solutions, Chicago, Illinois, United States; Helen Doll, MSc DPhil, Clinical Outcomes Solutions, Folkestone, Arizona, United Kingdom; Libby Floden, PhD, MPH, Clinical Outcomes Solutions, Tucson, Arizona, United States; Stacie Hudgens, MA, Clinical Outcomes Solutions, Tucson, Arizona, United States; Audrey Dunn-Galvin, University College, Cork, California, United Kingdom; Sarah Donelson, Aimmune Therapeutics, Cork, California, United States

**Aims:** According to the FDA's (2019) draft guidance, "scores generated by different types of clinical outcome assessments (COAs)...cannot be pooled to form a single clinical trial endpoint, even if they are developed to assess the same construct(s)." However, the appropriateness of pooling COAs is a testable hypothesis that should not be unilaterally ruled out without adjudication of its merits. This paper illustrates that multiple versions of the Food Allergy-Related Quality of Life Questionnaire (FAQLQ) measure the same construct and can be pooled successfully. **Methods:** Five versions of the FAQLQ were administered in three clinical trials (PALISADE, ARC003,  $n = 545$ ; RAMSES, ARC007,  $n = 493$ ; ARTEMIS, ARC010,  $n = 168$ ): FAQLQ-CF (children, 8–12 years), FAQLQ-TF (teenagers, 13–17 years), FAQLQ-AF (adults), FAQLQ-PF (parents of children), and FAQLQ-PFT (parents of teenagers). All items were measured using a 7-point response scale, where higher values denote worse outcomes. Rasch hierarchical generalized linear models (HGLM) were used to combine multiple COAs by linking on common items found across the measures and equating unique items to the same scale. Differential item functioning (DIF) analyses were performed to assess whether the common items measured the same construct across multiple COAs. Equipercetile equating was used to equate the FAQLQ-PF and FAQLQ-PFT since they lacked common items. **Results:** Rasch HGLM analyses showed subjects utilized the response scale as expected; namely, the thresholds between response categories increased monotonically and item difficulties were spread out across the continuum (see intercept and item coefficients, respectively, in Table 1). DIF analyses across the FAQLQ-CF, FAQLQ-TF, and FAQLQ-AF constructs indicated all but two items exhibited measurement invariance (see Table 2). Equipercetile equating was used to transform the FAQLQ-PFT into FAQLQ-PF

scores with greater than 99% accuracy (see example Fig. 1 for Emotional Impact domain). Conclusion: The FDA’s objection regarding pooling COAs is rooted in the belief one cannot statistically discern whether different groups of subjects conceptualize a construct in similar ways. This study challenges this notion and shows Rasch HGLM can be used to create a single ruler to measure across multiple COAs developed to measure the same construct, providing real-world benefits for research and ultimately patients; thus, suggesting the FDA’s policy needs revisiting.

Figure 1: Plot of Percentile Ranks by Percentile Scores for EL-PF and EL-PFT

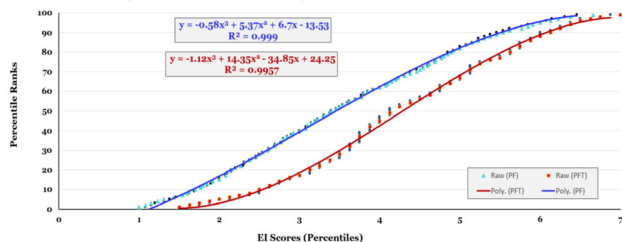


Table 1: Three-Level Rasch HGLM<sup>11</sup> Examining FAQLQ<sup>11</sup> at Baseline by Domain  
Population: Pooled PRP<sup>25</sup>; Domain: Emotional Impact - CTAF

Source	Description	Model 1 <sup>11</sup>	Model 2 <sup>11</sup>	Model 3 <sup>11</sup>
ICC		0.451	0.479	0.462
-2 LL		17340.33	17036.21 *	16814.98 *
AIC		17354.33	17072.21	16864.98
BIC		17386.94	17156.08	16981.18
Intercept <sup>41</sup>	1	-3.452 (0.090)	-3.849 (0.291)	-3.819 (0.377)
	2	-2.396 (0.078)	-2.755 (0.287)	-2.728 (0.374)
	3	-1.430 (0.072)	-1.753 (0.285)	-1.732 (0.373)
	4	-0.390 (0.069)	-0.666 (0.284)	-0.646 (0.372)
	5	0.572 (0.069)	0.343 (0.284)	0.370 (0.372)
	6	1.663 (0.073)	1.492 (0.285)	1.523 (0.372)
Item 15	Less control when eating out		0.427 (0.290)	0.445 (0.292)
Item 16	Carry an epinephrine auto injector		-0.537 (0.306)	-0.550 (0.309)
Item 17	Frightened of an allergic reaction		-0.369 (0.283)	-0.348 (0.287)
Item 18	Frightened of eating the wrong food		-0.175 (0.283)	-0.141 (0.287)
Item 19	Eat something never eaten before		0.841 (0.287) *	0.867 (0.291) *
Item 20	Frightened when eating out		0.097 (0.350)	0.088 (0.350)
Item 21	Never get rid of your food allergy		0.697 (0.295) *	0.730 (0.300) *
Item 22	People don't take it into account		0.227 (0.288)	0.248 (0.292)
Item 23	Disappointed to have food allergy		0.685 (0.295) *	0.719 (0.301) *
Item 24	Discouraged by allergic reaction		0.634 (0.291) *	0.658 (0.293) *
Item 25	You feel you are being a nuisance		0.691 (0.354)	0.687 (0.354)
Item 26 <sup>41</sup>	Apprehensive to eat something new		0.000 (.)	0.000 (.)
Sex <sup>71</sup>	Female		-0.682 (0.136) *	-0.682 (0.136) *
	Male		0.000 (.)	0.000 (.)
Race <sup>71</sup>	Asian		0.455 (0.215) *	0.455 (0.215) *
	Black or African American		0.553 (0.410)	0.553 (0.410)
	Multiple		0.004 (0.369)	0.004 (0.369)
	Native Hawaiian or Other Pacific Islander		-2.706 (2.042)	-2.706 (2.042)
	Other		0.115 (0.267)	0.115 (0.267)
	White		0.000 (.)	0.000 (.)
Age			0.013 (0.015)	0.013 (0.015)

Table 1: Three-Level Rasch HGLM<sup>11</sup> Examining FAQLQ<sup>11</sup> at Baseline by Domain

Abbreviations: HGLM = Hierarchical Generalized Linear Model, FAQLQ = Food Allergy Quality of Life Questionnaire, LL = Log likelihood, PRP = Patient reported outcome population, CTAF = unified Child-Teenager-Adult Form, PF = Parent Form (0-12 years old), PFT = Parent Form - Teenager (13-7 years), AIC = Akaike's Information Criterion, BIC = Bayesian Information Criterion, ICC = intraclass correlation coefficient, PRP = Patient Reported Outcome Population, FAP = Full Analysis Population.

Note: \* Statistically significant, p<0.05

[1] HGLMs are appropriate to use when categorical, non-normally distributed response variables are clustered within groups. Herein, the FAQLQ items for each domain were nested within each subject who, in turn, were nested within clinical trials. The failure to account for nesting has been shown to bias variance estimates, available power to detect treatment or covariate effects, inflate Type I error rates, and lead to substantive errors in interpreting the results of statistical significance tests.

[2] The FAQLQ measures an individual's health-related quality of life as impacted by allergy using one of two perspectives to collect the data—self-reported outcome by the patient him/herself [i.e., either a child (8-12 years old) or teenager (13-17 years old)] or by parent-reported outcome (i.e., the parent's perception) of the child (0-12 years old) or the teenager (13-17 years old). Statements are rated using a 7-point, from 1 = 'not' to 7 = 'extremely', where a higher score indicates a lower quality of life due to the impact of food allergy.

[3] The PRP is a subset of the FAP population and includes subjects who have completed at least one PRO measure at Baseline.

[4] Intercepts are the points on the latent variable where adjacent response categories of the item are equally probable. Since the FAQLQ employed a 7-point response scale, there are six intercepts (also known as, Andrich thresholds) that separate the 7 response categories.

[5] These coefficients represent regression coefficients (standard errors).

[6] This is the reference item against which other item coefficients are compared. Negative values for other items connote they were easier to endorse with high values on the response scale (greater food allergy severity) than this item, and vice versa for positive values.

[7] Males (M) and subjects who are White were used as reference groups for Sex and Race, respectively. The coefficients for the remaining groups are compared against these groups.

Table 2: Differential Item Functioning Examining FAQLQ<sup>11</sup> at Baseline by Domain  
Population: Pooled PRP<sup>25</sup>; Domain: Emotional Impact - CTAF

Item	Description	Form	Coefficient <sup>11</sup>	SE	t-ratio	p-value	alpha <sup>41</sup>
Item 15	Less control when eating out	Adult	-1.2281	0.40136	-3.0600	0.00223	0.00833
		Teenager	0.0000	-	-	-	-
Item 17	Frightened of an allergic reaction	Adult	0.0117	0.38443	0.0304	0.97573	0.00833
		Child	-0.0648	0.19765	-0.3279	0.74302	0.00833
		Teenager	0.0000	-	-	-	-
Item 18	Frightened of eating the wrong food	Adult	0.0000	-	-	-	-
		Child	0.0000	-	-	-	-
		Teenager	0.0000	-	-	-	-
Item 19	Eat something never eaten before	Child	0.7526	0.19532	3.8530	0.00012	0.00833
		Teenager	0.0000	-	-	-	-
Item 22	People don't take it into account	Child	-0.0011	0.20174	-0.0055	0.99559	0.00833
		Teenager	0.0000	-	-	-	-
Item 24	Discouraged by allergic reaction	Adult	-0.0423	0.39814	-0.1063	0.91534	0.00833
		Teenager	0.0000	-	-	-	-

Abbreviations: HGLM = Hierarchical Generalized Linear Model, FAQLQ = Food Allergy Quality of Life Questionnaire, CTAF = unified Child-Teenager-Adult Form, PF = Parent Form (0-12 years), PFT = Parent Form - Teenager (13-7 years), DIF = Differential Item Functioning, PRP = Patient Reported Outcome Population, FAP = Full Analysis Population.

[1] The FAQLQ measures an individual's health-related quality of life as impacted by allergy using one of two perspectives to collect the data—self-reported outcome by the patient him/herself [i.e., either a child (8-12 years old) or teenager (13-17 years old)] or by parent-reported outcome (i.e., the parent's perception) of the child (0-12 years old) or the teenager (13-17 years old). Statements are rated using a 7-point, from 1 = 'not' to 7 = 'extremely', where a higher score indicates a lower quality of life due to the impact of food allergy.

[2] The PRP is a subset of the FAP and includes subjects who have completed at least one PRO measure at Baseline.

[3] This estimate represents the DIF magnitude, also known as a contrast effect. By convention, only estimates greater than 0.5 are deemed non-trivial.

[4] In order to control the Type I error rate, the Bonferroni procedure was used to adjust for the multiple comparisons performed for each domain.

### (3122) Enhancing understanding of Patient-Reported Outcomes Measurement Information System® (PROMIS) pediatric measures of symptoms and functioning in relation to disease activity for youth with juvenile idiopathic arthritis and systemic lupus erythematosus

Elissa R Weitzman, ScD, MSc, Harvard Medical School | Boston Children's Hospital, Boston, Massachusetts, United States; Amy Gaultney, MD, MTS, UCLA Mattel Children's Hospital, Los Angeles, California, United States; Emily von Scheven, MD, University of California, San Francisco, Division of Pediatric Rheumatology, San Francisco, California, United States; Sarah Ringold, MD, Department of Pediatrics, Seattle Children's Hospital, Seattle, Washington, United States; Courtney M Mann, MA, Duke University School of Medicine, Department of Population Health Sciences, Durham, North Carolina, United States; Kara Magane, MPH, Boston Children's Hospital, Durham, Massachusetts, United States; Li Lin, MS, Duke University School of Medicine, Department of Population Health Sciences, Durham, North Carolina, United States; Renee Leverty, BSN, MA, Duke University Health System, Duke Clinical Research Institute, Durham, North Carolina, United States; Anne Denos, MSPH, Duke University Health System, Duke Clinical Research Institute, Durham, North Carolina, United States; Cheryl Rorick, RN, Duke University Health System, Duke Clinical Research Institute, Durham, North Carolina, United States; Alexy Hernandez, RN, Duke University Health System, Duke Clinical Research Institute, Durham, North Carolina, United States; Steven Lippmann, PhD, Duke University School of Medicine, Department of Population Health Sciences, Durham, North Carolina, United States; Alexandra Marin, PhD, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, United States; Rachele Cox, MPH, Boston Children's Hospital, Boston, Massachusetts, United States; Bryce Reeve, PhD, Duke University School of Medicine, Department of Population Health Sciences, Durham, North Carolina, United States; Laura E Schanberg, MD, Duke University Health System, Duke Clinical Research Institute, Durham, North Carolina, United States; submitted on behalf of the CARRA Registry Investigators

Aims: Evaluate construct validity of Patient-Reported Outcomes Measurement Information System® (PROMIS®) measures of symptoms and functioning against measures of disease activity among a multi-site US cohort of youth with juvenile idiopathic arthritis (JIA)

or systemic lupus erythematosus (SLE). Methods: Youth ages 8–17 years enrolled in a clinical disease registry completed PROMIS Pediatric measures of Mobility, Physical Activity, Fatigue, Pain Interference, Family Relationships, Peer Relationships, Depressive Symptoms, Psychological Stress, Anxiety, and Meaning and Purpose using computer-adaptive testing. Cross-sectional analyses of bivariate and multivariate associations among PROMIS and clinical metrics of disease activity were undertaken by condition. Results: Of 451 enrolled participants (average age 13.8 years, 71.4% female), most (87.1%) had JIA and the remainder (12.9%) SLE. Comparing those with inactive and active JIA, on average the latter had worse mobility (53.7 versus 45.2), and peer and family relationships (53.0 versus 50.9, and 52.1 versus 50.3, respectively), and greater fatigue (42.5 versus 47.4), pain interference (41.7 versus 47.7), depressive symptoms (46.4 versus 50.4), anxiety (43.9 versus 46.8), and psychological stress (50.7 versus 54.4) ( $p$  values  $< 0.05$ ). For SLE, comparing those with inactive and active disease, on average the latter had less mobility (50.8 versus 43.8) and greater family relationships (48.2 versus 56.5) ( $p$  values  $< 0.05$ ). Conclusion: This study provides evidence of the validity of a comprehensive set of PROMIS Pediatric measures describing dimensions of symptoms and functioning and distinguishing active from inactive disease among youth with JIA. This study provides initial evidence of the validity of PROMIS Pediatric measures for describing mobility and family involvement among youth with SLE, a group at particularly high risk for poor health outcomes. Results: advance our ability to use PROMIS for evaluating interventions and improve outcomes for youth with JIA and SLE.

### (3123) Comparing the use of patient-reported outcomes in clinical studies in Europe in 2008 and 2018; a literature review

Guro Lindviksmoen Astrup, PhD, Oslo University Hospital, Oslo, Norway; Gudrun Rohde, PhD, Sørlandet Hospital, Kristiansand, Norway; Stein Arne Rimehaug, MSc, Sunnaas Rehabilitation Hospital, Bjørnemyr, Norway; Marit Helen Andersen, PhD, Oslo University Hospital, Oslo, Norway; Tomm Bernklev, Dr. Philos, Vestfold Hospital, Tønsberg, Norway; Kristin Bjordal, Dr. med, Oslo University Hospital, Norway; Ragnhild Sørnum Falk, PhD, Oslo University Hospital, Oslo, Norway; Nina Marie Høyning Jørgensen, University of Oslo, Oslo, Norway; Knut Stavem, PhD, Akershus University Hospital, Lørenskog, Norway; Anita Tollisen, PhD, Lovisenberg Diaconal Hospital, Oslo, Norway; Cecilie Delphin Amdal, PhD, Lovisenberg Diaconal Hospital, Oslo, Norway; Cecilie Delphin Amdal, PhD, Oslo University Hospital, Oslo, Norway

Aims: Purpose: Several guidelines for the use of patient-reported outcomes (PROs) in clinical studies have been published the past decade. This review aimed to investigate whether these guidelines have contributed to an improvement by comparing the number and quality of clinical studies in Europe using PROs published in 2008 and 2018. Secondary, to describe the study designs, PRO instruments used, patient groups studied, and countries where the clinical studies were conducted. Methods: A literature search was conducted in MEDLINE to identify eligible publications. To assess the number of publications, all abstracts were screened for eligibility. To determine the quality of publications and describe secondary aims, a random sample of 150 eligible publications from each year were read in full text and assessed according to the CONSORT-PRO criteria. The studies were stratified on study design and divided in randomized clinical trials (RCTs) and other studies. Results: The search identified 1692 publications in 2008 and 4290 in 2018. After the screening of abstracts, 1240 from 2008 and 2869 from 2018 were eligible. By full-text review, the proportion of both RCTs and other studies discussing PRO-specific limitations and implications was higher in 2018 than in

2008, but there were no differences in the proportions complying with other CONSORT-PRO criteria. In 2018, a higher proportion of studies were longitudinal/cohort studies, included  $\geq 300$  patients, and used electronic administration of PRO than in 2008. A total of 340 PRO instruments and 29 single items or instruments designed for the specific study were used. The most common patient groups studied were those with cancer or diseases of the musculoskeletal system or connective tissue. Patients were most often included from Germany, United Kingdom, and The Netherlands. Conclusion: The number of clinical studies from Europe using PROs was higher in 2018 than in 2008, but there was little difference in the quality of the studies according to compliance with the CONSORT-PRO criteria, even for RCTs. The studies varied in terms of study design and PRO instruments used in both publication years.

### (3124) Patient-reported outcomes: do they provide additional information for individualized care?

Vanina Taliercio, MD, MS, University of Utah, Salt Lake City, Utah, United States; Ashley Snyder, MPH, University of Utah, Salt Lake City, Utah, United States; Kristina Duffin, MD, MS, University of Utah, Salt Lake City, Utah, United States; Zachary Hopkins, MD, Broward Health, Fort Lauderdale, Florida, United States; Gregory Stoddard, MPH, University of Utah, Salt Lake City, Utah, United States; Jacob Kean, PhD, University of Utah, Salt Lake City, Utah, United States; Rachel Hess, MD, MS, University of Utah, Salt Lake City, Utah, United States; Aaron Secrest, MD, PhD, University of Utah, Salt Lake City, Utah, United States

Aims: Patient-reported outcomes (PROs) can capture disease severity from the patient perspective. Despite the benefits of clinical use, including improved patient-clinician communication and shared decision-making, many clinicians remain skeptical of how useful PROs could be for their patients. We explored this utility within dermatology by comparing a validated quality of life (QOL) measure, Skindex-16, to the overall Physician Global Assessment (OPGA), a clinician-reported outcome (CRO) for psoriasis, and describing scenarios where the two aligned poorly. Methods: We retrospectively reviewed charts for 494 patients seen for psoriasis at University of Utah Dermatology since 2016 who completed a Skindex-16 questionnaire. Spearman correlations were calculated between OPGA and Skindex-16. Patients with  $> 1$  dermatology visit had their scores averaged prior to correlation calculations. We calculated median (10th-, 90th-percentile) overall Skindex-16 scores for each OPGA category. We identified two discordant groups: (1) “Clear skin, poor QOL,” signifying OPGA = 0 (clear of psoriasis), but Skindex-16 scores  $> 90$ th percentile, and (2) “Severe skin, good QOL,” meaning OPGA = 4 (severe psoriasis), but Skindex-16 scores  $< 10$ th percentile. Gender, concurrent skin conditions, depression, and rheumatological comorbidities were compared between these groups using Fisher’s exact test. Results: The correlation between overall Skindex-16 score and OPGA was  $r = 0.56$ , and correlations between OPGA and Skindex-16 domains were: symptoms ( $r = 0.53$ ), emotions ( $r = 0.54$ ), and functional ability ( $r = 0.47$ ;  $p < 0.001$  for all). Overall Skindex-16 scores were 6 (0, 43) for OPGA = 0; 16 (2, 49) for OPGA = 1, 32 (7, 72) for OPGA = 2, 55 (19, 86) for OPGA = 3, and 71 (24, 96) for OPGA = 4. Comparing discordant groups, those with “clear skin, poor QOL” were predominantly female (76% vs. 25%,  $p = 0.08$ ) and had at least one other skin condition (84% vs. 0%,  $p = 0.003$ ; Table 1). Conclusion: We found Skindex-16 moderately correlates with OPGA in psoriasis patients. About 6% of patients fell into discordant groups where PRO and CRO scores were mismatched. The existence of such discordance lends support to routine clinical use of PROs in dermatology to aid clinical decisions and improve communication. If noted during a visit, this discordance may prompt

further questioning into other dermatologic processes that may otherwise go undiscussed but be relevant to current treatment decisions.

Table 1. Comparison of discordant cases classified as either "clear skin, poor QOL" or "severe skin good QOL."

	OPGA=0, Skindex>90 <sup>th</sup> %ile "Clear skin, poor QOL"	OPGA=4, Skindex<10 <sup>th</sup> %ile "Severe skin, good QOL"	p-value*
Total, n (%)	25 (100)	4 (100)	---
Female, n (%)	19 (76)	1 (25)	0.08
Other chronic skin conditions, n (%)	21 (84)	0 (0)	0.003
Major depression, n (%)	13 (52)	1 (25)	0.60
Other comorbidities, n (%)	18 (72)	1 (25)	0.10
Median (min-max)	1 (0-8)		
Rheumatological disease	5 (20) 3 psoriatic arthritis 1 RA 1 fibromyalgia	1 (25) 1 psoriatic arthritis	1.0

\*p-values calculated using two-sided Fisher's exact test.

### (3125) Scoping review to inform PRO thresholds for use in research and clinical practice: traumatic brain injury case study

Sally Bradshaw, Centre for Patient Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Grace Turner, PhD, Centre for Patient Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Mel Calvert, PhD, Centre for Patient Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Anita Slade, PhD, Centre for Patient Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Ameeta Rezter, PhD, Centre for Patient Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom; Tony Belli, Institute of Inflammation and Ageing, Birmingham, United Kingdom; Christel McMullan, PhD, Centre for Patient Reported Outcomes Research, Institute of Applied Health Research, University of Birmingham, Birmingham, United Kingdom

**Aims:** Patient-reported outcomes are increasingly used in research and clinical practice to identify early deterioration of symptoms; resulting in a need to identify clinically relevant thresholds to trigger a response. The aim of this review was: i) to identify how thresholds for three commonly used PRO measures, GAD-7, PHQ-9 and PCL-5, were determined and used in traumatic brain injury (TBI) and ii) identify potential key principles for threshold selection for other tools and clinical areas. **Methods:** Using a formal scoping review methodology we systematically searched Medline, Embase, PsycINFO, CINAHL, AHMED, OpenGrey and Google databases for articles where GAD-7, PHQ-9 and PCL-5 were used in a TBI population. In addition, publisher manuals, national level and professional body guidelines including the measures in any population were identified. Data screening and extraction was undertaken by two independent reviewers. The review was registered on Research Registry, including search and screening strategy, explicit inclusion/exclusion criteria, data extraction and appraisal plan. **Results:** A total of 1,011 publications were screened and 64 studies included in the review. A total of three publisher manuals and 12 clinical guidelines were identified and also included. Although the publisher manuals for all three PROMs provided clear information on thresholds, GAD-7 and PHQ-9 are not specific to the TBI population. All the guidelines mentioned the tools but did not always provide further information on how to use the thresholds. In the single studies, there was a little

variation in the thresholds used. In addition, there was limited information regarding the reasons for using these particular thresholds. The single studies also provided limited information on reliability, validity, bias, and limitations of the tools. Implications for threshold review in other clinical disciplines will be presented. **Conclusion:** In order to improve the use of these tools within the field of TBI, authors should be clearer about their motivations behind selecting specific threshold. In addition, more information on applicability to the TBI population as well as more information on the reliability, validity, bias, and limitations of the tools with a TBI population is required. Finally, consideration should be given to applying these thresholds to other tools and clinical areas.

### (3126) Effect of shiftwork on the health and wellbeing of Alberta long-term and assisted-living professional caregivers

Oluwabohunmi Awosoga, PhD, MBA, Faculty of Health Sciences, University of Lethbridge, Lethbridge, Alberta, Canada; Jon Doan, PhD, P.Eng, University of Lethbridge, Lethbridge, Alberta, Canada; Claudia Steinke, PhD, RN, University of Lethbridge, Lethbridge, Alberta, Canada; Christina Nord, MS, PhD(c), University of Lethbridge, Lethbridge, Alberta, Canada; Stephanie Varsanyi, BSc, MS(c), University of Lethbridge, Lethbridge, Alberta, Canada; Jeff Meadows, M.Ed., University of Lethbridge, Lethbridge, Alberta, Canada; Sheli Murphy, PhD, RN, Covenant Health, Edmonton, Alberta, Canada

**Aims:** There are findings regarding the relationship between shift work and health, especially considering the impact of third shift (nighttime) work on workers physical and mental wellbeing. This study aims to identify factors that explain the effect of shiftwork on the health and wellbeing of long-term care (LTC) and assisted-living (AL) professional caregivers in Alberta. In this presentation, we explore whether or not there is a relationship between health status and shifts work among caregivers. **Methods:** The caring for professional caregivers research study was conducted in 39 LTC & AL Facilities across Alberta, using a questionnaire designed by Health Canada. Of the 1,385 questionnaires sent out to LTC & AL Facilities in five Alberta Health Services (AHS) Zones, 933 surveys (67.4%) were returned completed. Questions relating to shiftwork were selected and found to explain variances in the reported health status of caregivers. We next performed a Chi square test for these questions to examine whether there is a relationship between health status and the shifts caregivers worked. **Results:** We found significant differences between the responses from those working different shifts to the three questions, "I feel I have a lot of energy" ( $\chi^2(6) = 29.098, p = .000$ ), "I expect my health to get better" ( $\chi^2(6) = 15.052, p = .020$ ), and "Overall, I am satisfied with my health" ( $\chi^2(6) = 15.430, p = .017$ ). Similarly, there was a significant difference between responses of those working different shifts to the prompts, "I have high self-esteem/feel happy with myself, I am a happy person" ( $\chi^2(6) = 14.238, p = .027$ ), "I have a good level of motivation" ( $\chi^2(6) = 16.637, p = .011$ ), and "I have difficulty falling or staying asleep" ( $\chi^2(6) = 14.622, p = .023$ ). **Conclusion:** Our study provides further evidence that night shift caregivers report lower health status compared to caregivers working other shifts. In addition, caregivers reported negative evaluations of their physical, mental, and emotional health, lower ratings of their quality of life, and negative responses to questions concerning healthy behaviors. These findings will support AHS and other stakeholders to make personnel decisions and outline future policies.

**Qualitative Analysis of transcript**

We analyzed the interview and open-ended questions transcripts using NVivo 12 software to identify the most occurring themes from the participant responses. We read the transcripts several times to successfully obtain the final themes. Four major themes evolved, along with subthemes. The following categories show the resulting themes in participant responses concerning the health status of LTC/AL caregivers and how health status impacts their work, and what they believe can be done to improve it:

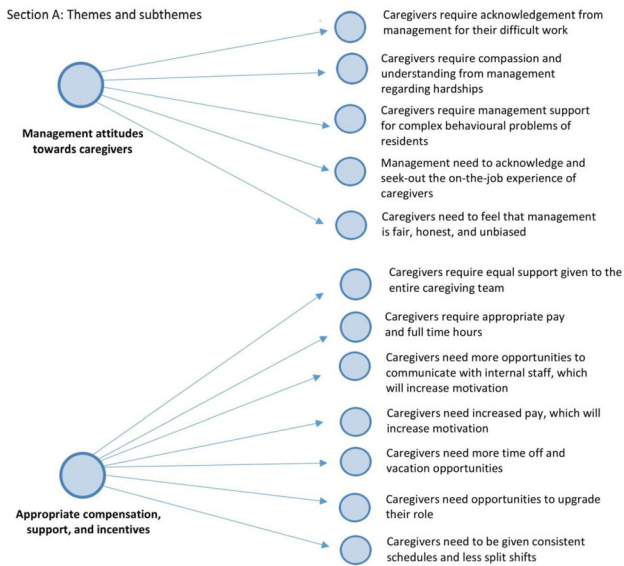


Figure 1: Themes and subthemes from participant interview and open-ended responses

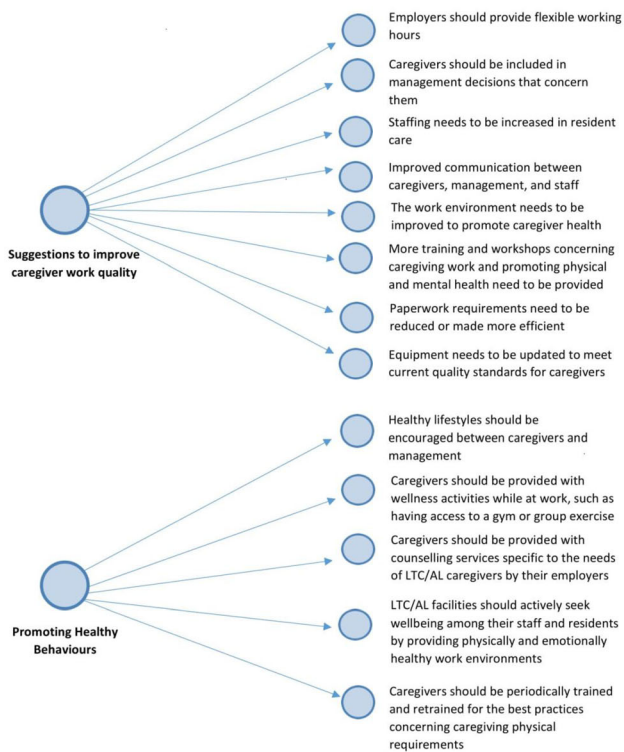


Figure 2: Themes and subthemes from participant interview and open-ended responses

Source: Authors' Construct (2019)

**(3127) Ethical considerations for the use of patient-reported outcomes (PRO) in research and routine practice: a systematic review**

Samantha Cruz Rivera, PhD, Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, United Kingdom; Olalekan Lee Aiyegbusi, PhD, Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, United Kingdom; Rebecca Mercieca-Bebber, University of Sydney, Faculty of Medicine and Health, NHMRC Clinical Trials Centre, Sydney, Australia; Jane Scott, Janssen Global Services, Johnson and Johnson, High Wycombe, United Kingdom; Amanda Hunn, Independent Advisor, London, United Kingdom; Carolyn Ells, Interagency Advisory Panel on Research Ethics, London, Ontario, Canada; Gary Price, Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, United Kingdom; Conrad Fernandez, Dalhousie University and the IWK Health Centre, Halifax, Nova Scotia, Canada; Heather Draper, University of Warwick, Warwick, United Kingdom; Jonathan Ives, University of Bristol, Bristol, United Kingdom; Melanie Calvert, University of Bristol, Bristol, United Kingdom; Melanie Calvert, Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, United Kingdom

**Aims:** Collection of patient-reported outcome (PRO) data in clinical trials and clinical practice can be associated with a number of ethical issues. The aim of this systematic review is to identify ethical considerations associated with PRO assessment in research and clinical practice, which can be used to enhance considerations for the safety of patients participating in PRO assessments and maximize benefits for patients and society. **Methods:** A systematic review of studies to identify ethical considerations and published ethical guidelines associated with PRO assessment was registered on the PROSPERO database (CRD42020176177). The review was conducted in the following databases from inception: MEDLINE (Ovid), EMBASE, Allied and Complimentary Medicine Database (AMED) and CINHALL. Further eligible papers were identified through Traditional Pearl Growing methodology (Schlosser et al., 2006), communication with experts and Google Scholar. Two reviewers independently reviewed titles and abstracts for eligibility. Any papers deemed potentially relevant at that stage were reviewed in full-text to determine eligibility. A thematic analysis approach was used to synthesize the ethical considerations in PRO research and more broadly in routine clinical practice. Patient partners and the broader team were included in the validation of the coding frame. **Results:** 67 papers were screened and 7 papers identified for full text screening, citation and reference searching is ongoing and results of the systematic review will be presented. Initial themes include: participant burden, data security and privacy, management of concerning PRO results and feedback of results to patients. A comparison of ethical issues in trials and routine practice and implications for research ethics guidance will be presented. **Conclusion:** This systematic review will provide a comprehensive assessment of the ethical considerations associated with the use of PROs in clinical practice and research. Findings of the systematic review will be used to inform ethics guidelines for use by Research Ethics Committees (RECs) and institutional review boards (IRBs) to protect patient safety and maximize benefits for patients and society.

**(3128) Review of the patient-reported outcomes instruments in ischemic heart disease**

Yolanda Pardo, CIBER de Epidemiología y Salud Pública (CIBERESP), IMIM (Institut Hospital del Mar d'Investigacions Mèdiques), Universitat Autònoma de Barcelona, Barcelona, Spain; Cristina Oriol, MD, IMIM (Institut Hospital del Mar d'Investigacions Mèdiques), Barcelona, Spain; Gemma Vilagut, PhD, IMIM (Institut



Hospital del Mar d'Investigacions Mèdiques), CIBER de Epidemiologia y Salud Pública (CIBERESP), Barcelona, Spain; Itxaso Alayo, IMIM (Institut Hospital del Mar d'Investigacions Mèdiques), CIBER de Epidemiologia y Salud Pública (CIBERESP), Barcelona, Spain; Víctor Zamora, IMIM (Institut Hospital del Mar d'Investigacions Mèdiques), Barcelona, Spain; Montserrat Ferrer, PhD, IMIM (Institut Hospital del Mar d'Investigacions Mèdiques), CIBER de Epidemiologia y Salud Pública (CIBERESP), Universitat Autònoma de Barcelona, Barcelona, Spain; Olatz Garin, PhD, IMIM (Institut Hospital del Mar d'Investigacions Mèdiques), CIBER de Epidemiologia y Salud Pública (CIBERESP), Universitat Pompeu Fabra, Barcelona, Spain; Group Empro, CIBERESP, CIBERNED, CIBERSAM, IMIM, UAB, UPF, Barcelona, Spain

**Aims:** Ischemic heart disease (IHD) is ranked first in disability-adjusted life years worldwide (Global Burden of Disease Study, 2017). In order to facilitate and improve PRO (Patient-Reported Outcome) measurement in patients with IHD at different settings, the aim of this study was to identify available disease-specific PRO instruments for individuals suffering this pathology, and to describe the construct measured. **Methods:** Disease-specific PRO instruments applicable for IHD were identified by systematic reviews published in Pubmed until 2019 and specialized libraries (BiblioPRO and PROQOLID). Systematic searches were conducted in Pubmed until 2019 for each instrument identified. Detailed information was collected on their general characteristics: author, year and country of publication, target population, construct measured, number of items, dimensions, scores, and versions available for other countries. **Results:** Thirty-three instruments were identified: 15 for IHD patients, 10 developed for heart diseases in general, and 8 designed for patients undergoing a specific treatment (revascularization or surgery, for example). Instruments were published between 1985 and 2014. Regarding the construct measured, 15 of them evaluated health-related quality of life (HRQL), 9 were symptoms scales, 2 measured functional activity, and 7 evaluated other constructs such as depression, attitudes, adjustment, social support or self-efficacy. Regarding the availability of country versions, 24 instruments were only available in the language of development or had one or two cross-cultural adaptations, while the following 3 had more than 30 adaptations: MacNew Heart Disease Health-Related Quality of Life Questionnaire, Seattle Angina Questionnaire, and Heart Quality of Life Questionnaire. Most of the instruments include physical and cognitive/emotional components and some of them, such as the MacNew, also have a social component, or a disease-specific component as the Seattle Angina Questionnaire. **Conclusion:** There are at least 33 PRO instruments to assess the impact that ischemic heart disease has on patients, most of them measure HRQL, and could be used both in the clinical setting and in research. It is important to know the instruments available and their conceptual models in order to select the most suitable for each study, setting and purpose.

### **(3129) Development of knowledge transfer tools to facilitate involvement of patient partners in PRO trial protocol development, according to SPIRIT-PRO**

Samantha Cruz Rivera, Dr, Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, United Kingdom; Richard Stephens, BBMRI-ERIC Stakeholder Group, London, United Kingdom; Rebecca Mercieca-Bebber, Dr, University of Sydney, Sydney, Australia; Claudia Rutherford, Dr, University of Sydney, Sydney, Australia; Gary Price, CPROR, University of Birmingham, Birmingham, United Kingdom; Ameeta Retzer, CPROR, University of Birmingham, Birmingham, United Kingdom; Anita Slade, CPROR, University of Birmingham, Birmingham, United Kingdom; Olalekan Lee Aiyegbusi, Dr, CPROR, University of Birmingham,

Birmingham, United Kingdom; Philip Edge, CPROR, University of Birmingham, Birmingham, United Kingdom; Lesley Gosden, CPROR, University of Birmingham, Birmingham, United Kingdom; Lesley Roberts, CPROR, University of Birmingham, Birmingham, United Kingdom; Lesley Roberts, CPROR, University of Birmingham, Birmingham, United Kingdom; Rav Verdi, CPROR, University of Birmingham, Birmingham, United Kingdom; Roger Wilson, NCRI Consumer Forum National Cancer Research Institute, London, United Kingdom; Melanie Calvert, Prof, CPROR, Birmingham, Birmingham, United Kingdom

**Aims:** Patient-reported outcomes (PROs) are increasingly used in clinical trials. Recent research suggests patient partners would like more engagement in the development of trial protocols with PRO endpoints, and that involving patients in this way may reduce missing PRO data during trial conduct. The SPIRIT-PRO Extension provides recommendations for items that should be addressed in PRO clinical trial protocols. However, there is lack of training materials and tools to support patient partners involved in the co-design of PRO clinical trials. Therefore, the aim of this research was to co-design: a) a user-friendly version of the SPIRIT-PRO Extension guidance; and b) a web-based tool to support the dissemination and uptake of the SPIRIT-PRO Extension for patient partners. **Methods:** A lay summary and glossary for each of the SPIRIT-PRO items was co-developed with patient partners and used to inform discussions at a one-day patient and public involvement session held in November 2019 at the University of Birmingham. Five patient partners co-designed the tools, while two more patient partners were involved in writing the manuscript. The study adhered to INVOLVE guidelines and was reported according to GRIPP 2 checklists. **Results:** Two user-friendly tools were developed to help those patients and members of the public involved in the co-design of PRO clinical trials. The first tool presents a lay version of the SPIRIT-PRO Extension guidance. The second tool depicts the most relevant points, identified by the PPI group, of the guidance through an interactive flow diagram. The involvement of patients and members of the public helped to ensure that the tools focused on issues most relevant to them. They were involved in the design, checked comprehension of the vocabulary and piloted both tools. They also contributed to edits of the paper and are co-authors. **Conclusion:** These tools, if used appropriately, have the potential to facilitate the involvement of patient partners in providing informed input into the development of PRO aspects of clinical trial protocols, in accordance with the SPIRIT-PRO Extension guidelines.

### **(3130) The value of patient-reported outcomes and quality of life measures as endpoints in FDA approvals of new therapies and products**

Patricia Vanderwolf, BA, National Organization for Rare Disorders, Washington, District Of Columbia, United States; Vanessa Boulanger, MSc, National Organization for Rare Disorders, Danbury, Connecticut, United States; Pamela Gavin, MBA, National Organization for Rare Disorders, Boston, Massachusetts, United States; Amelia Robert, MPH, National Organization for Rare Disorders, Washington, District Of Columbia, United States; Danielle Boyce, MPH, National Organization for Rare Disorders, Washington, District Of Columbia, United States

**Aims:** The National Organization for Rare Disorders (NORD), under a grant from FDA Center for Biologics Evaluation and Research (CBER), is leading an innovative project to define best practices and processes for designing a regulatory-grade natural history study. In partnership with rare disease patient advocacy organizations and the FDA, this pilot project translates the FDA Guidance Rare Diseases: Natural History Studies for Drug Development into actionable

impact. **Methods:** The study incorporates sound research practices for building natural history studies, incorporating patient-reported outcomes, and leveraging the cohort data for use as external controls, thereby shifting clinical trial design for rare diseases and providing regulatory-grade evidence to inform FDA decision-making processes. **Results:** This study will make a significant contribution to regulatory science by demonstrating the utility of well-designed natural history studies that incorporate patient-reported outcomes and integrated data streams for use as external controls. Patients and caregivers are provided with a powerful opportunity to contribute directly to research that will enhance understanding of rare disorders, facilitating the development of new diagnostic and treatment options. **Conclusion:** This presentation will 1, provide an overview of patient-focused drug development processes, 2, review the multi-stakeholder programmatic approach used to work in collaboration with the FDA, patient advocacy groups, and clinicians and researchers, 3, highlight the methodology used to develop a natural history study for a rare disease case study, and 4, demonstrate recruitment and retention practices to engage and maintain a diverse patient population using a longitudinal study design. The presentation will also cover the methods used to reduce patient burden and enhance data collection through electronic recruitment, web-based surveys, supplemental health information, event-driven mobile data collection, and linkages to electronic health records. In addition, the evaluation of patient-reported outcomes for use as endpoints will be compared to those reported by clinicians.

#### Methodological Considerations in Cancer Research

##### (3131) ‘Looking under the hood’ of anchor-based assessment of clinically important change: Role of domain and attribution in explaining Global-Assessment-of-Change scores

Roland Stark, M.Ed., DeltaQuest Foundation, Concord, Massachusetts, United States; Joel Finkelstein, M.D., M.Sc., FRCS(C), Sunnybrook Health Sciences Center, Toronto, Ontario, Canada; Bruce Rapkin, Ph.D., Albert Einstein College of Medicine, Bronx, New York, United States; Wesley Michael, M.B.A., Rare Patient Voice, LLC, Towson, Maryland, United States; Carolyn Schwartz, Sc.D., DeltaQuest Foundation, Concord, Massachusetts, United States

**Aims:** The Global Assessment of Change (GAC) item has facilitated the interpretation of change in patient-reported outcomes (PRO). Querying patients’ evaluation of recent health change, GAC has provided a useful anchor for computing minimally important PRO differences. GAC’s construct validity has been documented via disease-specific PRO change. We sought to explain GAC ratings using a variety of sociodemographic factors; health-related quality of life (HRQOL) domains; attributions of health-related change; and QOL appraisal processes. **Methods:** This secondary analysis examined data from 1481 chronically ill patients and caregivers (mean age 50, SD 13; 86% female; mean no. comorbidities = 4.1 for patients and 3.3 for caregivers) who completed a web-based survey at baseline and 17 months. Items queried change since baseline in overall disease-related symptoms (GAC), as well as in Physical-, Emotional-, and Social-functioning domains. Candidate predictors included sociodemographic factors; binary attributions of change including health, life circumstances, supports, etc.; items and second-order principal component scores representing QOL appraisal. LASSO (least absolute shrinkage and selection operator) and bootstrapping tested 77 predictors’ effectiveness and stability in accounting for variance in GAC. **Results:** GAC worsening was notably associated with being disabled ( $\beta = -0.24$ ) and having difficulty paying bills ( $\beta = -0.13$ ). Subsequent models controlled for these and six other sociodemographics. GAC was better explained by the Physical domain than the Emotional

or Social domains ( $\beta = 0.67, 0.10, \text{ and } 0.03$  and  $p < .0005, < .0005, \text{ and } = .20$ , respectively;  $R^2_{\text{adj}} = 0.63$ ). In a separate model ( $R^2_{\text{adj}} = 0.18$ ), GAC variance was explained by goals related to solving problems with healthcare and keeping up activities; attributions about changing health and changing response of one’s health team; and appraisal about things getting better ( $\beta = -0.07, 0.05, -0.14, 0.08, 0.21$ , respectively,  $p$  range  $\sim 0.0005\text{--}0.05$ ). Caregivers were more likely than patients to attribute GAC to changing responsibilities and support from others. **Conclusion:** The GAC primarily reflects the Physical domain of HRQOL, and it reflects goals, attributions, and patterns of emphasis related to change in health and healthcare. Our findings have bearing on the construct validity of GAC in facilitating interpretation of PRO change. They suggest that many other unmeasured factors may be relevant to explaining GAC scores.

##### (3132) Usability of an electronic data assessment software for collection of patient-reported outcomes (PRO) and common terminology criteria for adverse events (CTCAE) in a randomized controlled trial

Lisa M. Wintner, Medical University of Innsbruck, Innsbruck, Austria; Johannes M. Giesinger, Medical University of Innsbruck, Innsbruck, Austria; Andrew Bottomley, European Organisation for Research and Treatment of Cancer (EORTC Headquarters), Brussels, Belgium; Yuichiro Kikawa, Kansai Medical University Hospital, Osaka, Japan; Guillaume Mouillet, Centre hospitalier régional universitaire de Besançon, Besançon, France; August Zabernigg, Kufstein County Hospital, Besançon, Austria; Heike Schmidt, Institute of Health and Nursing Science and University Clinic and Outpatient Clinic for Radiotherapy; Medical Faculty; Martin Luther University Halle-Wittenberg, Halle, Germany; Gerhard Rumpold, Medical University of Innsbruck, Austria, Austria; Bernhard Holzner, Medical University of Innsbruck, Innsbruck, Austria; submitted on behalf of the EORTC Quality of Life Group

**Aims:** A randomized controlled trial currently underway is investigating whether the provision of patient-reported outcome (PRO) data improves the inter-rater reliability between two independent ratings of adverse events (according to common terminology criteria for adverse events, CTCAE) (trial registration number). To enable real time data assessment, processing and visualization, the Computer-based Health Evaluation System (CHES) is used to collect PRO data, calculate scores and present them alongside the electronic CTCAE rating. Before starting data collection, the usability of the software was evaluated to identify shortcomings of software functionalities and further need for user instruction. **Methods:** The usability test included a five page explanation of the trial procedure and software functionalities, login data for a CHES preview version, and the request to log on the system and complete a short list of tasks. Finally, raters completed the System Usability Scale (SUS) and four comprehension questions evaluating the understanding of trial information material. The SUS is a 10-item scale using a 5 point Likert scale for evaluating the usability of computer systems, which is frequently used and has received the maximum score in a recent quality appraisal. **Results:** So far, 16 raters from Austria (37.5%), France (18.8%) and Japan (43.8%) completed the usability test. Most participants were physicians and had extensive experience in participating in clinical studies (each 81%). All except one rater had at least “a little” experience with CTCAE ratings (93%). The half were female (50%) and on average 41 years old (SD 8.4), with an average of 12 years of professional experience (SD 6.7). The overall SUS score was 80.8 points (SD 14.0). All comprehension questions have been answered correctly by 37.5% of raters and 56.3% answered 3 out of 4 correctly. **Conclusion:** According to published thresholds, CHES proved to have

a good usability and basic trial information showed to be well comprehended. Thus, there were no changes of software functions or instruction materials. Data collection for usability testing is ongoing in three more centers (Germany, Italy, Jordan) and final results will be presented at the conference.

### **(3133) Burden and quality of life (QOL) of informal caregivers of cancer patients: An overview of self-reporting instruments**

Eva Brown Hajdukova, PhD, Parexel, London, United Kingdom; Ewelina Rzepa, PhD, Parexel, London, United Kingdom; Katja Rudell, PhD, Parexel, London, United Kingdom

**Aims:** Caregivers of patients with cancer have multiple, varied, and serious unmet financial, physical, emotional, and social needs. Without adequate psycho-social support, distressed caregivers are at risk for psychiatric and medical morbidity, which can adversely affect patient outcomes and health care costs. While the burden of cancer on informal caregivers' quality of life (QoL) has been recognized, clinical trials rarely assess these effects. Limited information on caregiver burden is available when evaluating the cost-effectiveness of treatments. To guide researchers and clinicians in their choice of caregiver burden and QoL screening tools, this poster provides an overview of instruments identified in cancer clinical trials and research studies and describes their content and relevance for this caregiver population. **Methods:** A targeted literature review was performed using PsycINFO and MEDLINE to identify cancer studies published in English between 1990 and 2019 that included self-reported assessment measures of caregiver burden and QoL. Additional searches were conducted on clinicaltrials.gov. **Results:** Utilizing inclusion criteria, 28 QoL and 6 caregiver burden instruments were identified and evaluated regarding their content and relevance for this carer population. Nearly all reviewed studies and clinical trials implemented one of the following: EQ-5D-5L, PROMIS Depression Short Form, PROMIS Anxiety Short Form, Hospital Anxiety and Depression Scale (HADS), the Zarit Burden Short Form, Caregiver Well-Being Scale (CWBS), the Short Form 12 (SF12), the Caregiver Quality of Life Index-Cancer (CQOLC), Work Productivity and Activity Impairment—modified (WPAI:CG) and the Montgomery Borgatta Caregiver Burden scale. **Conclusion:** There are a wide range of instruments considering different aspects of the impact of cancer on informal caregivers, however further research into the appropriateness and validity of using non-specific cancer measures is needed to ascertain which instruments may provide the most comprehensive and parsimonious solution for caregiver burden assessment. A solution will be proposed in the poster.

### **(3134) Managing symptoms and functions in colorectal cancer survivors: mixed-methods study of stakeholder opinions and current practices**

Claudia Rutherford, PhD, The University of Sydney, School of Psychology, Sydney, Australia; Angela Ju, PhD, The University of Sydney, School of Psychology, Sydney, Australia; Lisette Wiltink, PhD, The University of Sydney, School of Psychology, Sydney, Australia; Nasiba Faiz, BN, The University of Sydney, Sydney Nursing School, Sydney, Australia; Madeleine King, PhD, The University of Sydney, School of Psychology, Sydney, Australia; Kate White, PhD, The University of Sydney, Sydney Nursing School, Sydney, Australia

**Aims:** Colorectal cancer (CRC) is prevalent in the developed world, with unhealthy lifestyles and diet contributing to rising incidence. Advances in effective diagnosis and treatments have improved survival rates but CRC survivors experience poorer physical and psychological function than the general population and suffer from long-term gastrointestinal (GI) symptoms. Managing these important patient-reported outcomes (PROs) is key to improving survivors' quality of life. However, the current state of care provision and management options remains unclear. We aimed to identify how GI symptoms and functioning impairments are currently managed in CRC survivors and what interventions are available for relevant PROs. **Methods:** We conducted: 1) searches in five electronic databases for studies describing/evaluating interventions to address GI symptoms and functioning in CRC; and 2) conducted an international survey plus structured interviews with CRC survivors and managing health professionals on current practices and opinions about available interventions and gaps. **Results:** We identified 21 interventions for managing GI symptoms and functional outcomes in individuals treated for CRC (10 behavioral, 5 complementary, 4 pharmacological, 2 psychological). Few interventions are available for managing fatigue, bowel and sexual function. Responses from CRC survivors revealed that while survivors have ongoing debilitating problems such as neuropathy, impaired bowel and sexual function, and anxiety, few seek professional help and often self-manage their symptoms. Health professionals stressed the need for methods to monitor and detect these PROs and affordable and established clinical pathways to specialist care such as nurse-led clinics or comprehensive allied health services to provide appropriate holistic care. **Conclusion:** Patients seldom seek professional help and few supportive care interventions or services are routinely offered. Follow-up care for CRC survivors should integrate routine monitoring of symptoms and functions, with the goal of earlier detection and amelioration of these problems to improve quality of life.

### **(3135) Quality of Life in cancer patients: the role of the patients' evaluation of the information received from health professionals. An international EORTC study**

Johannes Giesinger, PhD, Medical University of Innsbruck, Innsbruck, Austria; Iqbal Bahar, Cachar Cancer Hospital and Research Centre Silchar, Silchar, India; Eva Greimel, Medical University Graz, Graz, Austria; Omar Shamieh, King Hussein Cancer Center, Amman, Jordan; Michael Koller, University Hospital Regensburg, Regensburg, Germany; Anne Bredart, Institute Curie and University Paris Descartes, Regensburg, France; Karin Kuljanic, University Hospital Center Rijeka, Rijeka, Croatia; Anna Costantini, Andrea Hospital, Sapienza University, Rome, Italy; Monika Sztankay, Medical University of Innsbruck, Innsbruck, Austria; Lisa Wintner, Medical University of Innsbruck, Innsbruck, Austria; Uxue Zarandona, Medical University of Innsbruck, Innsbruck, Austria; Uxue Zarandona, Complejo Hospitalario de Navarra, Pamplona, Spain; Marina Carreiro de Sousa, Universidade dos Açores, Ponta Delgada, Portugal; Hiroto Ishiki, National Cancer Center, Tokyo, Japan; Meropi Kontogianni, Harokopio University, Athens, Greece; Maja Wolan, University of Rzeszow, Rzeszów, Poland; Yuichiro Kikawa, Kobe City Medical Center General Hospital, Kobe, Japan; Dagmara Kulis, EORTC Quality of Life Department, Brussels, Belgium; Juan Ignacio Arraras, Complejo Hospitalario de Navarra, Pamplona, Spain

**Aims:** Quality of Life (QOL) is a key outcome in cancer patients and widely assessed in clinical trials and daily practice. Patients'

perception of their QOL may be affected by how the patient evaluates the health professionals' view on his/her disease and treatment. The aim of the present work is to study in an international sample the impact of the patient's evaluation of the information on their disease and treatment received by professional on the patient's QOL ratings. Methods: Cancer patients from 12 countries in Europe and Asia completed the EORTC QLQ-C30 and an item on the patients' evaluation of the health professional's (doctors and nurses) views about their disease and treatment (response options ranging from 1 very negative to 10 very positive) during the consultation. Patients were evaluated after a consultation with the doctor: at first follow-up after radiotherapy or three months after chemotherapy start. Using previously published thresholds for clinical importance, we created binary variables (clinically important problem no/yes) from each QLQ-C30 scale to be used as dependent variables in a multivariate logistic regression analysis. The key independent variable in this model was the item on the health professional's view on disease/treatment (adjusted for age, gender, disease stage, comorbidity, curative/palliate treatment and previous treatment). Results: We analyzed data from 717 cancer patients (mean age 59.3 years; 53.8% females; 52.6% curative treatment, and 57.9% had received previous treatment). The mean score for the item on the patient's evaluation was 7.3 (SD 1.9). In the multivariate model, negative ratings of the professionals' views predicted clinically important symptoms/problem for 11 QLQ-C30 domains: physical (odds ratio (OR) 0.72), role (OR 0.89), social (OR 0.83) emotional (OR 0.88) and cognitive (OR 0.78) functioning; summary score (OR 0.78); fatigue (OR 1.18), nausea/vomiting (OR 1.22), pain (OR 1.10), dyspnoea (OR 1.14) and appetite loss (OR 1.19). All  $p$  values < 0.05. Conclusion: After adjusting for key clinical variables, patients' evaluation on the professional's views was found to be a QOL determinant in an international sample. Patients' interpretation might be addressed in the consultations through adequate information and communication strategies.

### (3136) Understanding patient adherence to routine online PRO completions during cancer treatment- insights from qualitative feedback

Kate Absolom, University of Leeds, Leeds, United Kingdom; Lorraine Warrington, PhD, University of Leeds, Leeds, United Kingdom; Marie Holmes, University of Leeds, Leeds, United Kingdom; Andrea Gibson, University of Leeds, Leeds, United Kingdom; Rosemary Peacock, University of Leeds, Leeds, United Kingdom; Beverly Clayton, University of Leeds, Leeds, United Kingdom; Zoe Rogers, University of Leeds, Leeds, United Kingdom; Sarah Dickinson, University of Leeds, Leeds, United Kingdom; Patricia Holch, Leeds Beckett University, Leeds, United Kingdom; Galina Velikova, University of Leeds, Leeds, United Kingdom; University of Leeds, Leeds, United Kingdom

Aims: Patient-Reported Outcome (PRO) assessments are increasingly used in clinical care. It remains essential we continue to capture patients' experiences of their application and learn from feedback. The eRAPID intervention was designed to allow cancer patients to complete symptom-reports online from home. Immediate tailored advice for prompting patient action (self-management or seeking medical advice) is provided. Symptom data are made available to clinical teams via electronic patient records. The eRAPID systemic RCT evaluated the intervention during chemotherapy. Findings from the embedded qualitative substudy are presented here. Views of eRAPID were explored through patient interviews and written feedback to understand acceptability, adherence and future recommendations. Methods: Patients starting treatment for breast, gynecological or colorectal cancers ( $n = 508$ ) were recruited to an RCT comparing Usual Care (UC) with UC plus the eRAPID

intervention. Over 18-weeks, intervention patients ( $n = 256$ ) were asked to use eRAPID by completing weekly online symptom-reports. In addition to main trial outcomes (quality of life/clinical processes/use of resources) participant feedback on the intervention was gathered through end of study interviews and written feedback forms. These data were collated and analyzed thematically. Results: Interviews were conducted with  $n = 44$  patients and written comments obtained from  $n = 175$ . Feedback could be summarized under three main interconnecting themes to describe patient views on the value of eRAPID and adherence to online symptom-reporting: (1) IT functioning, (2) Personal Benefit and (3) Medical/clinical use. Some patients were highly positive and motivated by the support provided by the personal symptom monitoring processes and tailored advice. Others identified as research participants who were helping to benefit future patients. Criticisms around the clinical use of the symptom reports and limitations in the capability of the IT system emerged, particularly from written comments. Conclusion: The eRAPID qualitative data provides important patient-centered insight which aids interpretation of the main trial findings. Although many patients felt eRAPID provided added value to care experiences others were deterred by limited clinician use and restrictions in IT functions. Collecting written feedback in addition to interviews gave more participants an opportunity to share honest views of the intervention. The findings provide valuable guidance for future intervention development and implementation.

### (3137) The lived experience of clinical research associates with missing patient-reported outcome data in cancer randomized clinical trials

Michael Palmer, PhD, Queen's University, Kingston, Ontario, Canada; Terry Krupa, PhD, Queen's University, Kingston, Ontario, Canada; Harriet Richardson, PhD, Queen's University, Kingston, Ontario, Canada; Michael Brundage, MD, Queen's University, Kingston, Ontario, Canada

Aims: Missing patient-reported outcome (PRO) data threaten the validity of PRO-specific findings and conclusions from randomized clinical trials (RCTs) by introducing bias due to data missing not at random. Clinical Research Associates (CRAs) are a largely unexplored source for informing understanding of potential causes of missing PRO data. The purpose of this qualitative study was to describe factors that influence missing PRO data as revealed through the lived experience of CRAs. Methods: Maximum variation sampling was used to select CRAs having a range of experience with missing PRO data from university-based or community centers in different geographic locations of Canada. Semi-structured interviews were audio-recorded, transcribed verbatim, and analyzed according to descriptive phenomenology. Steps included: (1) identifying relevant statements by considering how each relates to the research question, (2) evaluating the uniqueness of each relevant statement, (3) identifying 'invariant meaning units' by removing the redundant statements, (4) organizing invariant meaning units into clusters, and grouping related clusters into themes. Data collection continued until saturation. Authors performed a face-validity exercise reaching consensus that each cluster was a factor having relevance to missing PRO data, and every theme reflected the associated factors. Results: Eleven CRAs were interviewed. Most had more than ten years of experience. CRAs did not describe a predominant influence but instead revealed several factors that influence missing PRO data, organized within the following themes. PROs for routine clinical care compete with PROs for RCTs. Both the paper format and the electronic format have benefits and drawbacks. Missing PRO data are influenced by characteristics of the instruments, and by personal characteristics of the patients. Assessment of PROs at progression of disease is particularly

difficult. Deficiencies in center research infrastructure can contribute. CRAs develop relationships with patients that may help reduce missing PRO data. It is not always possible to provide sufficient time to complete the instrument. There is a need for field guidance, and a motivation among CRAs to contribute their knowledge to address issues. Conclusion: These results enhance understanding of which factors influence missing PRO data and have important implications for designing solutions to minimize missing PRO data on cancer RCTs.

**(3138) Is there anything else you would like to share?—Experiences of being a family caregiver to a patient treated for oesophageal cancer**

Cecilia Haddad Ringborg, Karolinska Institutet, Department of Molecular Medicine and Surgery, Surgical Care Science, Stockholm, Sweden; Anna Schandl, Karolinska Institutet, Department of Molecular Medicine and Surgery, Surgical Care Science and Department of Anaesthesiology and Intensive Care, Södersjukhuset Stockholm Sweden, Stockholm, Sweden; Pernilla Lagergren, Karolinska Institutet, Department of Molecular Medicine and Surgery, Surgical Care Science and Department of Surgery and Cancer, Imperial College London, London, United Kingdom, Stockholm, Sweden

**Aims:** There is a need to put family caregivers on the cancer survivorship research agenda as they are a part of the survivorship concept. Because the literature about family caregivers to oesophageal cancer patients is scarce, the aim of this study is to explore the experiences of being a family caregiver to a patient with oesophageal cancer. **Methods:** This qualitative study was based on the ongoing nationwide and prospective cohort study entitled Oesophageal Surgery on Cancer patients—Adaptation and Recovery study (OSCAR) including patients surgically treated for oesophageal cancer in Sweden and their closest family caregiver. One year after the patient's surgery each family caregiver received a self-report questionnaire-kit to fill in and return by postal mail. For the purpose of this study the responses of the last open-ended question "Is there anything else you would like to share?" was used and analyzed by conducting thematic analysis. **Results:** In total 114 responses of the open-ended question were transcribed and analyzed. The text rendered four themes. The health care contact—mostly a positive experience before surgery. However, after discharge from hospital they felt being left alone, fully responsible for the patient's care. A changed life—an unprepared changed life situation after the patient's surgery. A feeling that nothing will ever be the same and like your personal self is eradicated. You live to take care of your family member. Psychological distress—was described as a feeling of being alone. Family caregivers felt invisible and no longer important to family and friends. The patient was the one that mattered. Ways to cope—stay positive and talk about the disease and death. However, sometimes allow to feel down and having a bad day. **Conclusion:** This study indicates that the healthcare is missing the family caregivers' perspectives and more studies are needed to further investigate the situation of family caregivers to patients with oesophageal cancer.

**(3139) Impact of open-label design on patient-reported outcomes (PROs) data in randomized clinical trials of immuno-oncology (IO) agents in patients with advanced or metastatic cancer: a 10-year systematic literature review (SLR)**

Amélie Anota, PhD, Methodology and Quality of Life in Oncology Unit (U1098) & French National Platform Quality of Life and Cancer,

University Hospital of Besançon, Besançon, France; Astrid Pozet, Methodology and Quality of Life in Oncology Unit (U1098), University Hospital of Besançon, Besançon, France; Hervé Lemasson, PhD, Bristol-Myers Squibb, Market Access, Rueil-Malmaison, France; François-Emerly Cotte, PhD, Bristol-Myers Squibb, Market Access, Rueil-Malmaison, France; Stéphane Guerzider, MD, Chest Disease Department, University Hospital of Besançon, Besançon, France; Guillaume Mouillet, MD, PhD, Methodology and Quality of Life Unit in Oncology (U1098) & Oncology Department, University Hospital of Besançon, Besançon, France; Guillaume Eberst, MD, Chest Disease Department, University Hospital of Besançon, Besançon, France; Emilie Charton, Methodology and Quality of Life in Oncology Unit (U1098), University Hospital of Besançon, Besançon, France; Virginie Westeel, MD, PhD, Methodology and Quality of Life in Oncology Unit (U1098) & Chest Disease Department, University Hospital of Besançon, Besançon, France

**Aims:** The unblinded nature of open-label clinical trials and the knowledge of the treatment received could bias PRO measures by influencing patients' propensity to complete the questionnaire and their quality of life (perception). In this context, we conducted a SLR of IO trials across indications, to assess and quantify the potential impact of open-label vs double-blind designs on PRO data. **Methods:** A systematic search of indexed literature published from January 2009 to May 2019 was conducted using PubMed/MEDLINE, Cochrane Library, and EMBASE database. A detailed search algorithm identified all randomized clinical trials (RCTs) of IO therapies focused on advanced or metastatic cancer patients reporting PRO data. We conducted descriptive analyses quantifying differences at baseline and over time according to the type of study (open-label vs blinded) in terms of questionnaire completion rate and on obtained PRO scores. **Results:** Overall, we identified 8,284 references. After removal of duplicates and screening papers according to predefined criteria, 23 studies were identified. Among them, 15 (65%) were open-label studies. The main cancer sites were melanoma ( $n = 10$ , 43%) and non-small cell lung cancer ( $n = 5$ , 22%). The principal IO-drugs investigated, alone or as part of a combination therapy regimen, were nivolumab ( $n = 11$ , 48%) and ipilimumab ( $n = 7$ , 30%). The first PRO questionnaire was administered at randomization or before treatment start in majority of studies (57%), or at the beginning of the treatment (22%). The PRO compliance rate was reported in most studies (74%), and particularly in open-label studies (80%). At baseline, no clear difference in compliance rate was observed between experimental and control arms whatever the type of studies (absolute mean difference of 2.2% for blinded studies vs 2.8% for open label studies). No clinically significant difference of baseline PRO level was observed between treatment arms. Over time, impact on PRO measures cannot be observed due to the limited number of studies and heterogeneity of questionnaires and tumor types. **Conclusion:** No impact of the design was highlighted at baseline regarding compliance rate and PRO measures. Future research should involve analyses by specific cancer types and ideally compare individual data from two closed RCTs (blinded vs open-label).

**(3140) Reliability Generalization Study on the WHOQOL-BREF Taiwan Version**

Li chung Lin, master, National Taiwan University, Taipei, Taiwan; Grace Yao, professor, National Taiwan University, Taipei, Taiwan; Yi-Ching Su, master, National Taiwan University, Taipei, Taiwan

**Aims:** From the psychometric viewpoint, reliability is the property of scale scores, but not the scale itself (Vacha-Haase, 1998). In other words, score reliability varies across samples and test conditions. It is

important to examine the score reliability for individual studies. The WHOQOL-BREF is a popular scale in many fields. So far, there is no study on the influential factors of the score reliability for the WHOQOL-BREF. Reliability generalization (RG) is a powerful approach to examining the psychometric properties of the score reliability and assessing the characteristics associated with score reliability. The present study utilized RG to examine the score reliability of the four domains of the WHOQOL-BREF Taiwan version. Methods: We used the keywords, WHOQOL and World Health Organization Quality of Life, to select 1,332 doctoral and master theses (from 1998 to May 2019) from two databases. The reason for selecting doctoral and master theses instead of published articles was that more necessary information appeared in the theses. After excluding the theses with missing information, 248 theses were left. In addition to calculating the mean and standard deviation (std) of score reliabilities on each domain across studies, we conducted several statistical analyses (such as regression analysis, correlation analysis, t-test, and analysis of variance) to examine the relationship between score reliability (Cronbach’s alpha) and eleven characteristics (such as mean and std of domain scores, sample size, mean and std of age, gender ratio, education, religion, marriage status, sample type, living area.). Results: The mean (> 0.7) and std (< 0.09) for each domain indicate that the WHOQOL-BREF Taiwan version has acceptable and consistent reliability across studies. Moreover, the std of the domain scores significantly predict their score reliability for the four domains. Education level, religious beliefs, age, and the std of domain scores are correlated with the score reliability, especially for physical and environmental domains. Psychological and social domains are more robust to the eleven characteristics. Conclusion: The WHOQOL-BREF Taiwan version is a reliable scale. The study results allow the researchers to evaluate the influential factors on score reliability and modify their study design to maximize its score reliability.

**Musculoskeletal research**

**(3141) Association between the Surgical Invasiveness Index and patient-reported outcomes after lumbar spine surgery**

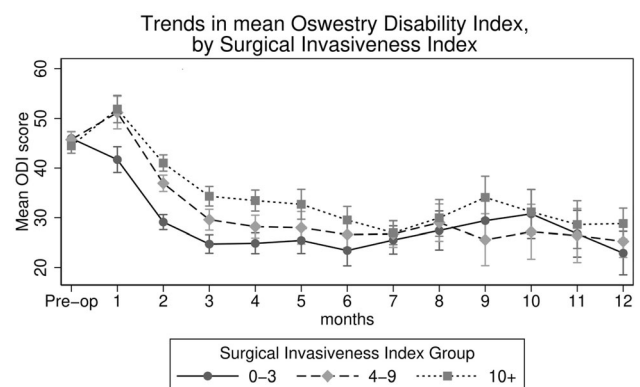
Brook Martin, PhD MPH, University of Utah, Salt Lake City, Utah, United States; Darrel Brodke, MD, University of Utah, Salt Lake City, Utah, United States; Amy Cizik, PhD, University of Utah, Salt Lake City, Utah, United States; Julie Fritz, PT, PhD, FAPTA, University of Utah, Salt Lake City, Utah, United States

**Aims:** The Surgical Invasiveness Index is a measure of the complexity of spine surgery that correlates with operating time, blood loss, and surgical site infection. No previous study has examined its association with patient-reported outcomes after lumbar spine surgery. We sought to examine the association between surgical invasiveness and patient-reported measures of Pain Interference and Physical Function using the Patient-Reported Outcomes Measurement Information System (PROMIS), and back-related disability using the Oswestry Disability Index (ODI). We further describe differences in response rates over a range of clinically relevant treatment thresholds. **Methods:** We prospectively collected outcomes during routine appointments from 1,774 patients undergoing lumbar spine decompression with or without fusion. The Surgical Invasiveness Index was calculated using a previously validated weighting of Current Procedural Terminology (CPT) codes, and then categorized into three groups based on their ranking. Mixed-effects regressions were used to

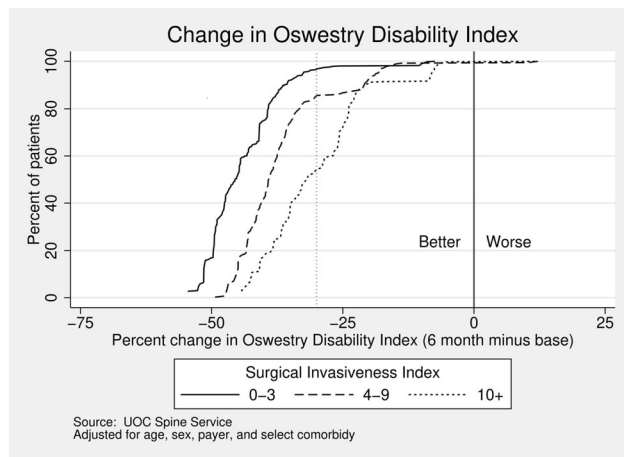
test the association between invasiveness and PROMIS or ODI outcomes, adjusting for age, sex, insurance and select comorbidities. Cumulative distribution functions of the adjusted outcomes describe differences in group-level treatment response over the range of treatment effect thresholds. Results: The mean age of the cohort was 57.5 (sd 16.5), with 45.7% female, and 45% from a commercial payer. The low invasiveness group had significantly lower levels of baseline diabetes and depression. Compared to the high invasiveness group, those with low invasiveness reported a quicker improvement in disability and physical function during the first six postoperative months. For each measure, responder rates decrease as the threshold for defining improvement increases, with response among those in the high invasiveness group declining more precipitously. For example, in the high invasiveness group only 55% achieved a 30% improvement in disability, compared to over 95% of patients in the low invasiveness group. Conclusion: Greater surgical invasiveness, a reflection of the complexity of surgical treatment and patient pathology, was associated with slower short-term improvement in physical function and disability, as well as lower treatment response rates. Understanding the influence of surgical invasiveness might help advise patients about their expected outcomes and response to surgery.

**Table 1:** Cohort characteristics by surgical invasiveness

	0-3	4-9	10+	All	p-value
<i>Age</i>					
mean	51.0	60.3	62.8	57.5	p<0.001
<i>Sex</i>					
Female (%)	35.2	51.5	53.5	45.8	p<0.001
<i>Primary payer</i>					
Medicare (%)	25.5	46.1	57.4	41.7	p<0.001
Private (%)	58.4	40.2	32.5	44.9	
Medicaid (%)	11.0	9.1	5.9	8.8	
Other (%)	5.1	4.6	4.3	4.7	
<i>Comorbidities</i>					
Diabetes (%)	14.9	24.1	22.7	20.1	p<0.001
Cancer (%)	4.2	5.2	6.7	5.3	p=0.130
Depression (%)	20.5	26.6	24.7	23.6	p=0.034
Anxiety (%)	10.5	13.9	13.0	12.3	p=0.177
Tobacco (%)	9.1	10.6	8.5	9.4	p=0.475
Alcohol (%)	5.2	4.8	4.8	5.0	p=0.936
Opioid use disorder (%)	1.9	1.9	2.0	1.9	p=0.995



Source: UOC Spine Service  
 Error bars represent 95% Confidence Interval  
 Adjusted for age, sex, payer, and select comorbidities



### (3142) The relationship between a measure of therapeutic alliance assessed during an episode of physical therapy for low back pain with patient-reported outcomes

Julie Fritz, PT, PhD, University of Utah, Salt Lake City, Utah, United States; Faris Alodaibi, PT, PhD, Rehabilitation Science Department, King Saud University, Riyadh, Saudi Arabia; Rett Holmes, PT, Physical Therapy at St. Luke's University Health Network, Bethlehem, PA, Pennsylvania, United States; Stephen Karehla, PT, PhD, Physical Therapy at St. Luke's University Health Network, Bethlehem, Pennsylvania, United States; Jason Beneciuk, PT, PhD, MPH, University of Florida Department of Physical Therapy,, Gainesville, Florida, United States

**Aims:** Therapeutic alliance (TA) describes the working relationship between patient and provider established through collaboration, communication, agreement around goals, mutual trust and respect. The TA has been studied extensively in psychotherapy, demonstrating that TA is related to patient-centered outcomes and treatment adherence. Little research on TA has been done in physical therapy. Existing studies have been done in experimental settings. The aim of this study was to examine the relationship between TA assessed during an episode of physical therapy care under routine clinical circumstances and patient-reported outcomes at the conclusion of the episode for patients with low back pain. **Methods:** Outpatient physical therapy clinics in one integrated health system routinely collect patient-reported data at initial, interim and final visits. The lumbar computer-adapted test (LCAT) is the patient-reported measure of function collected for patients with low back pain. The TA was assessed at interim assessments using the Working Alliance—Short Revised (WAI-SR), reported as three subscales (Goals, Task, Bond) and Total score. Descriptive statistics and bivariate correlations were examined. Regression models examined if interim WAI-SR scores explained outcome variance beyond a previously validated multivariate prediction model. **Results:** Analysis included 676 patients (mean age = 55.6, SD = 16.1, 55.9% female). Substantial ceiling effects were observed. Interim WAI-SR scores were not correlated with baseline patient characteristics or LCAT score ( $p = 0.03$ – $0.05$ ), but were correlated with the final LCAT and LCAT change score from initial to final assessment ( $p = 0.17$ – $0.25$ ). The WAI-SR Total score (Adj.  $R^2 = 0.36$ ,  $p < 0.001$ ), and Task (Adj.  $R^2 = 0.38$ ,  $p < 0.001$ ), and Goal subscales (Adj.  $R^2 = 0.35$ ,  $p < 0.001$ ) explained additional variance in outcome beyond the multivariate base model (Adj.  $R^2 = 0.33$ ,  $p < 0.001$ ). **Conclusion:** Our findings are consistent with studies in psychotherapy in finding that the TA is related to clinical outcomes. Our study replicated these findings under routine PT clinical circumstances. While the TA and its subconstructs, particularly the Task construct, were related to patient-centered outcomes, we found challenges in the measurement of TA using the WAI-SR, most notably substantial ceiling

effects. Despite measurement difficulties, the potential utility of understanding TA and its role in optimizing patient-centered outcomes warrants further research.

### (3143) Patient-reported outcome measures for rheumatoid arthritis disease activity: a systematic review

Tim Pickles, MSc, Centre for Trials Research, Cardiff University, Cardiff, United Kingdom; Tanzeel Ijaz, MBBS, Hywel Dda University Health Board, Llanelli, United Kingdom; Mike Horton, PhD, Psychometric Laboratory for Health Sciences, University of Leeds, Leeds, United Kingdom; Karl Bang Christensen, PhD, Section of Biostatistics, University of Copenhagen, Copenhagen, Denmark; Rhiannon Phillips, PhD, Cardiff School of Sport and Health Sciences, Cardiff Metropolitan University, Cardiff, United Kingdom; David Gillespie, PhD, Centre for Trials Research, Cardiff University, Cardiff, United Kingdom; Ernest Choy, FRCP, Division of Infection and Immunity, Cardiff University, Cardiff, United Kingdom

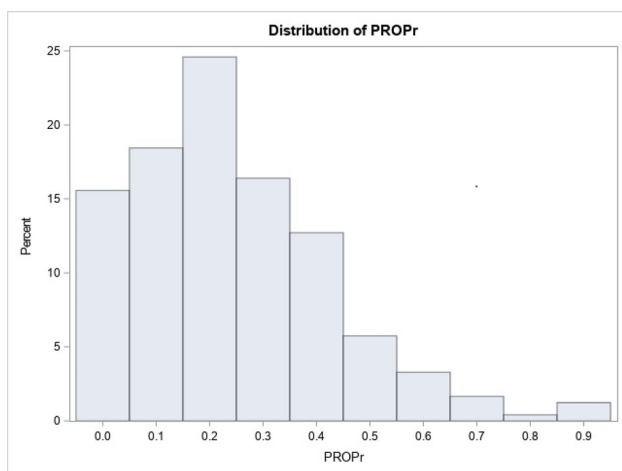
**Aims:** Previous research has suggested that Patient-Reported Outcome Measures (PROMs) are the most informative way to assess Rheumatoid Arthritis (RA) Disease Activity (DA) and allow for a more efficient use of NHS resource. PROMs are critical to research and clinical care, as recognized by the US Food and Drug Administration (FDA), who have published guidelines on how to develop and validate PROMs. The objective was to identify all PROMs for RA DA and evaluate their methodological quality and measurement properties. **Methods:** The PubMed and EMBASE searches undertaken by a previous systematic review were expanded to include all papers up to January 2019 (rather than the previous June 2014 date) and those before January 1994. Some papers from the review were excluded as they involved some clinician or laboratory input. All papers selected for data extraction were assessed for methodological quality and measurement properties through the COSMIN guidelines by two independent reviewers. **Results:** 702 abstracts were retrieved: 310 from both PubMed and EMBASE, 230 from PubMed alone and 162 from EMBASE alone. 107 were selected for abstract review; 62 for full article review; and, 39 for data extraction. These provided information on 12 different PROMs for RA DA. 30 matched the articles in the review (with 4 papers excluded for involving some clinician or laboratory input), and 9 were published after June 2014. The methodological quality in these reports varied, with COSMIN domains largely unknown in the majority and no PROM providing evidence across the necessary domains. **Conclusion:** While there are many RA DA PROMs currently available, the evidence for their quality is not adequate. These were developed using standard validation techniques and do not meet the stringent measurement criteria required under item response theory or Rasch measurement theory required by the FDA. RAPID3, RADAI and RADAI5 had the strongest measurement properties. This search also identified articles on the related area of RA Flare and an older assessment whose items did not form a PROM. There were also inconsistencies or lack of reporting on the items forming the PROM and many reporting categorization agreement of measures, which is not relevant under the COSMIN guidelines.

### (3144) Health valuation in spine surgery: PROMIS-preference score association with socio-demographic and clinical features

Richard Skolasky, Sc.D., Johns Hopkins University, Baltimore, Maryland, United States; Brian Neuman, M.D., Johns Hopkins University, Baltimore, Maryland, United States; Barry Dewitt, Ph.D.,

Carnegie Mellon University, Pittsburgh, Pennsylvania, United States; Janel Hamner, M.D., Ph.D., University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania, United States

**Aims:** Increasing trends of spine surgery use and costs place greater pressure on healthcare payers. Coverage models are proposed to control costs by requiring evidence of benefit for reimbursement. This value-based purchasing requires estimates of cost and “value.” Quality-adjusted life years, a patient-centered value measure, integrate time and quality of life (QOL) change using health utility. Hanmer developed PROMIS-Preference (PROPr) scoring, a societal preference-based measure based on 7 PROMIS domains, to produce health utilities with range of  $-0.022$  to  $1.0$  (0 equivalent to the utility of “dead” and  $1.0$  equivalent to the utility of “full health”). Collecting PROMIS/PROPr would provide domain-level scores to assist clinical decision-making and health utility scores to support economic analyses for policy decisions (e.g., comparative-effectiveness analyses). **Methods:** Evaluate face validity of PROPr in spine surgery patients presenting at an academic center by examining relationship of PROPr with demographics (age, gender, race/ethnicity, patient activation), behavioral (exercise) and clinical features (comorbidity, pain severity and treatment expectations), and patient-reported outcomes (Oswestry/Neck Disability Index (O/NDI)). PROPr score was calculated from PROMIS health domains. Relationship to demographics, clinical features, and other patient-reported outcomes assessed using correlation coefficient ( $r$ ) for continuous and analysis of variance (ANOVA) for categorical measures. Significance at  $0.001$  ( $r < 0.30$  not considered). **Results:** Mean PROPr score was .22 (SD .17, range  $-.02, .90$ ) (Fig. 1) and not associated with age, gender, race/ethnicity, or comorbidity. There was weak to strong correlation in lumbar (back pain,  $r = -.39$  and ODI,  $r = -.76$ ) and cervical (neck pain,  $r = -.38$  and NDI,  $r = -.72$ ) patients but weak correlation with expectations ( $r = .35$ ,  $p < .001$ ). Regular exercise ( $> 2x/wk$  vs. none) had higher score (.27 vs. .14). Those most engaged (PAM stage 4 vs. others) had higher scores (.23 vs. .17, .18, and .21). **Conclusion:** PROMIS health domains can be used to estimate health utility among surgery patients. Estimates show face validity; PROPr is related to severity of spine disease, expectations, regular exercise, and patient activation. Routinely collected PROMIS health domains are a powerful tool to estimate health utility in surgery patients and allow clinicians, researchers, and health economists to assess the value of care in an ever increasingly value-conscious healthcare environment.



### (3145) Psychometric properties of the Patient Assessment of Chronic Illness Care Questionnaire (PACIC) in people with chronic low back pain: a Rasch based analysis

Maitri Modi, Msc in Rehabilitation Science, McGill University, Montreal, Quebec, Canada; Kedar Mate, Post Doc, McGill University, Montreal, Quebec, Canada; Andre Bussière, McGill University, Montreal, Quebec, Canada; Sara Ahmed, McGill University, Montreal, Quebec, Canada

**Aims:** Measuring patients’ experience with health services and care received is an essential element of performance assessment and service improvement when reporting on the quality of care and defining health policy. The Patient Assessment of Chronic Illness Care (PACIC) questionnaire is widely used to measure patients’ experience of health services in patients with chronic pain. However, there have been conflicting findings regarding its psychometric properties. Therefore, re-validating the scale to assess its psychometric properties is necessary. The aim of this study is to assess the extent to which the PACIC items measure a single construct and fit an underlying theoretical hierarchy that forms a linear continuum with interval-like properties. **Methods:** This longitudinal study design data on PACIC measure from 135 people with low back pain who participated in a 6-month interdisciplinary program implemented in four primary care settings in Montreal, Canada. The PACIC questionnaire is a 20 items tool rated on a five-point Likert scale. The items form 5 subscales. Participants rated their experiences with chronic care at baseline and 6 months. The measurement properties of the PACIC were evaluated by Rasch analysis using RUMM2030 software. Fit statistics were assessed according to fit residuals, Chi square, and item characteristic curve (ICC). **Results:** Preliminary results show that there were several indicators of misfit for a full scale with 20 items: 13 items showed disordered thresholds, the fit residuals were outside of range  $-2.50$  to  $+2.50$  for 7 items, and 12 items failed to have observed scores for the inter classes close to the ICC. The items (“asked for my ideas,” “about my health habits,” “carry out my daily life,” “affects my life,” “refer to services” and “other specialist”) showed misfit on all three tests. **Conclusion:** This study will provide evidence of whether the tool fits reasonably well to the Rasch Measurement Theory, and which items need to be revised or replaced. The results will contribute to revising the PACIC to enhance its performance to provide researchers, clinicians, and the healthcare system decision-makers with a tool that can be trusted to assess patient-centered care in individuals with chronic pain.

### (3146) Patient-reported outcome measures for rheumatoid arthritis symptom severity: development of a computer-adaptive test from an item bank using Rasch measurement theory

Tim Pickles, MSc, Centre for Trials Research, Cardiff University, Cardiff, United Kingdom; Ernest Choy, FRCP, Division of Infection and Immunity, Cardiff University, Cardiff, United Kingdom; Mike Horton, PhD, Psychometric Laboratory for Health Sciences, University of Leeds, Leeds, United Kingdom; Karl Bang Christensen, PhD, Section of Biostatistics, University of Copenhagen, Copenhagen, Denmark; Rhiannon Phillips, PhD, Cardiff School of Sport and Health Sciences, Cardiff Metropolitan University, Cardiff, United Kingdom; David Gillespie, PhD, Centre for Trials Research, Cardiff University, Cardiff, United Kingdom

**Aims:** Patient-reported Outcome Measures (PROMs) are critical to clinical care. Although there are many PROMs currently used in clinical trials in rheumatic diseases, they were developed using



standard validation techniques and do not meet the stringent measurement criteria required under Rasch model methodology set by the US Food and Drug Administration. Rheumatoid Arthritis (RA) is a chronic, disabling, autoimmune disease that can attack the entire body. RA affects 1% of the population and no cure is available, so disease modification and symptom management are key for patients. NICE and EULAR RA monitoring guidelines suggest a frequency far more often than patients are currently seen in UK hospitals. RA patients locally and nationally have expressed a desire to have a simple PROM for monitoring their own disease. A computer-adaptive test (CAT), built from items collated into a single-item bank has the potential to transform clinical care in the future. Methods: From a systematic review, the existing PROMs, and items within them, measuring the construct of RA symptom severity have been identified. These were discussed with a National Rheumatoid Arthritis Society organized focus group and study PPI Stakeholders, who suggested the need for items on discomfort of walking, standing and exercising, plus fear of falling, and provided a new example of a pain to scale to rate across joint areas. These, along with the Rheumatoid Arthritis Flare Questionnaire and additional fatigue items, form the items of a questionnaire that will be sent out to adult RA patients in the Cardiff and Vale University Health Board. These data from will be analyzed under Rasch measurement theory to determine which items from existing PROMs can form an item pool, and their content validity will be assessed in discussion with RA patients. Co-calibration of the item pool, again under Rasch measurement theory, will be used to develop an item bank. Lastly, a prototype CAT will be developed, including initial user testing with patients. Results: Rasch analyses of data collected with the RADA15 PROM through the Austrian BioREG registry suggest that this is not a suitable tool to measure the construct of RA symptom severity. Conclusion:

#### **(3147) Responsiveness of the Portuguese version of Pain and Disability Index (SPADI) in patients with shoulder dysfunction**

Luís Cavalheiro, Centre for Health Studies and Research of the University of Coimbra, Coimbra, Portugal; Rui Gonçalves, Polytechnic Institute of Coimbra, Coimbra Health School, Coimbra, Portugal, Coimbra, Portugal; Inês Primo, Polytechnic Institute of Coimbra, Coimbra Health School, Coimbra, Portugal, Coimbra, Portugal; Ana Fernandes, Polytechnic Institute of Coimbra, Coimbra Health School, Coimbra, Portugal, Coimbra, Portugal; João Matos, College of Health Dr. Lopes Dias, Castelo Branco, Portugal, Castelo Branco, Portugal; Pedro Ferreira, Centre for Health Studies and Research of the University of Coimbra, Castelo Branco, Portugal

**Aims:** Shoulder's complexity and the nature of the demands on it make it susceptible to the appearance of several lesions, characterized by pain and movement restriction. The primary complaints associated with shoulder problems are largely subjective in nature, so it is becoming apparent that the patient's subjective perception of their disease status is decisive for both the diagnostic work-up and the subsequent therapeutic management. The Shoulder Pain and Disability Index (SPADI), is deemed a quality of life measure and comprises two subscales: pain and function. Is one of the higher quality questionnaires for evaluation of shoulder and is frequently used. The aim of this study was to analyze the responsiveness of the portuguese version of SPADI, and to test its reliability and validity in patients with shoulder dysfunction. **Methods:** The portuguese version was obtained by forward/backward translations, consensus panel and cognitive debriefing. The obtained version together with a OSS and SF-12 were applied to 93 adults with shoulder problems. The 48-h reproducibility was tested by ICC and internal consistency by Cronbach's alpha coefficient. The validity was tested through the relations obtained with the other PRO measures. The

responsiveness was evaluated after 4 weeks of physical therapy in 57 subjects through the indicators Effect Size Standardized (ESS), Response Mean Standardized (RMS) and Minimal Important Change (MIC). Results: The results for validity showed correlation coefficients between the scores of SPADI Total and OSS ( $r = 0.78$ ). We have also obtained significant correlation between SPDAI and SF-12 summary ( $r = 0.53$  and  $r = 0.28$ ). In assessing the reliability we found high internal consistency and reproducibility with 0.93 for SPADI Total, 0.79 for Pain subscale, and 0.92 for Function subscale. ICC values range to 0.84 (SPADI total – 95% CI 0.71–0.91) and 0.77 (SPDI Function–95% CI 0.59–0.87). Effect size was moderate with SEM between 0.51 (95% CI 0.34–0.72) for SPADI Pain, and 0.63 (95% CI 0.41–0.86) for SPADI Function; and RMS between 0.69 (95% CI 0.42–0.94) for SPADI Pain, and 0.76 (95% CI 0.49–1.00) for SPADI Total. The MIC was approximately 17 for SPADI Total. Conclusion: The portuguese version of SPADI evidenced acceptable levels of reliability and validity. It also showed responsiveness for patients with shoulder dysfunction.

#### **(3148) Cross-cultural adaptation, validity and reliability of the Yoruba version of the Physical Activity Scale for the Elderly**

Adesola Odole, Department of Physiotherapy, College of Medicine, University of Ibadan, Ibadan, Nigeria; Odunaiya Nse, Department of Physiotherapy, College of Medicine, University of Ibadan, Ibadan, Nigeria; Ifeoma Blessing Nwosu, MSc in orthopedic and sports Physiotherapy, Department of Medical Rehabilitation, Faculty of Health Sciences and Technology, Nnamdi Azikiwe University Awka, Nigeria, Awka, Nigeria; Adefemi Akinseloyin, Global consult and Raypower radio station, Ibadan Nigeria, Ibadan, Nigeria; Clara Fatoye, Department of Health Professions, Manchester Metropolitan University, Manchester, UK, Manchester, United Kingdom; Folake Olubiye, Department of Physiotherapy, College of Medicine, University of Ibadan, Ibadan, Nigeria, Manchester, Nigeria; Ezinne Ekediegwu, Department of Medical Rehabilitation, Faculty of Health Sciences and Technology, Nnamdi Azikiwe University Awka, Nigeria, Awka, Nigeria; Chidozie Mbada, Department of Medical Rehabilitation, College of Health Sciences, Obafemi Awolowo University, Ile-Ife, Nigeria, Ile-Ife, Nigeria; Odunola Akinola, Department of Physiotherapy, Bowen University, Iwo, Niger, Iwo, Niger

**Aims:** The aim of this study was to translate, cross culturally adapt and validate the Yoruba version of the Physical Activity Scale for the Elderly (PASE). **Methods:** Following the Beaton guidelines, the English-version of the PASE was adapted into Yoruba through forward-back translations, experts' committee meetings, pretesting and cognitive debriefing interview. Fifty-four purposively selected older adults who had scored  $\geq 24$  on the Mini Mental State Test participated in a cross-sectional survey. The Test retest reliability for Yoruba PASE was investigated using the Intra-Class Correlation (ICC) coefficient while the known group (construct) validity was investigated with the Spearman rank correlation coefficient. Results: The Yoruba version of the PASE has semantic, idiomatic, linguistic and conceptual equivalence with the English PASE. Test retest reliability for Yoruba PASE yielded Intra-Class Correlation (ICC) value of 0.993. The analysis of the sum scores of both the English and Yoruba versions of the PASE for the known group validity yielded a spearman correlation value of  $\rho = 0.989$ ,  $P = 0.001$ . Conclusion: The Yoruba version of the PASE is a valid and reliable outcome instrument and may be used in research and practice fields for the measurement of physical activity among the Yoruba speaking older adults.

**Table 1: Correlations<sup>[1]</sup> between the Standard<sup>[2]</sup> and Revised<sup>[3]</sup> FAQIG<sup>[4]</sup> Domain Scores at Baseline**  
Population: Pooled PPT<sup>[5]</sup>

Domain	Correlation
<b>Combined Child/ Teenager-Adult Forms (CAF)</b>	
Allergen Avoidance/ Dietary Restrictions (AA-CAF)	1.000
Allergen Avoidance (AA-CAF)	0.957 with AA-IR-CAF
Dietary Restrictions (DR-CAF)	0.930 with AA-IR-CAF
Emotional Impact (E-CAF)	1.000
Food Allergy-related Health (FH-CAF)	1.000
Risk of Accidental Exposure (RAE-CAF)	1.000
<b>Total Score</b>	1.000
<b>Parent Form (PF)</b>	
Emotional Impact (E-PF)	0.966
Food Anxiety (FA-PF)	0.988
Social and Dietary Limitations (SDL-PF)	0.999
<b>Total Score</b>	1.000
<b>Parent Form - Teenager (PFT)</b>	
Emotional Impact (E-PFT)	1.000
Food Anxiety (FA-PFT)	1.000
Social Restrictions (SR-PFT)	0.933 with SDL-PF
Dietary Restrictions (DR-PFT)	0.879 with SDL-PF
<b>Total Score</b>	1.000

**Abbreviations:** PPT = Patient-reported outcome population, ITT = Intend-to-treat population, PRO = Patient-reported outcome, CAF = United Children/ Teenager-Adult Form, PF = Parent Form, PFT = Parent Form - Teenagers  
 [1] A Spearman rank correlation was calculated for each domain. Since the revised scoring algorithm combined two sets of domains (AA-CAF and DR-CAF into AA-IR-CAF; SR-PFT and DR-PFT into SDL-PFT), correlations for each set are presented for the separate domains (standard) with the combined (revised) domains. Hence, these correlations only capture the impact of how revised values were scored. Based on Cohen's recommendations, the following values represent the magnitude of association: small = 0.1, moderate = 0.3 and large = 0.5.  
 [2] For the standard scoring, domain score = sum of item scores / number of completed items, except for the FAQIG-PF in which domain score = sum of item scores / number of items. The Allergen Avoidance and Dietary Restrictions domains are scored separately in the PF but together in the CAF and TF Forms; the Dietary Restrictions and Social Restrictions domains are scored separately in the PFT but together in the PF. The Total score = the sum of all items / number of completed items, except for the FAQIG-PF in which Total score = sum of domain scores / number of domains.  
 [3] For the revised scoring, domain score = sum of item scores / number of completed items. The Allergen Avoidance and Dietary Restrictions domains are scored together in the CAF, FA, and TF Forms; the Dietary Restrictions and Social Restrictions domains are scored together in the PFT and PF. The Total score = sum of the domain scores / number of domains.  
 [4] The FAQIG measure an individual's HRQL using data from the patient or by the child's parent. Statements are rated using a 7-point scale, from 1 = 'not at all' to 7 = 'extremely', where a higher score indicates a lower quality of life due to the impact of food allergy.  
 [5] The PPT (N=1206) is a subset of the PAF (ITT) and included subjects who completed at least one PRO measure at baseline.

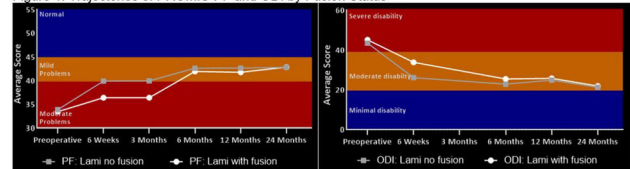
**(3149) Trajectories of PROMIS physical function, pain interference, and participation in social roles in 24 months following lumbar spine surgery**

Jacquelyn S. Pennings, Ph.D., Vanderbilt University Medical Center, Nashville, Tennessee, United States; Rogelio A. Coronado, Ph.D., PT, Vanderbilt University Medical Center, Nashville, Tennessee, United States; Master Hiral, Ph.D., Vanderbilt University Medical Center, Nashville, Tennessee, United States; Wegener Stephen T., Ph.D., Johns Hopkins Medicine, Baltimore, Maryland, United States; Kristin R. Archer, Ph.D., DPT, Vanderbilt University Medical Center, Nashville, Tennessee, United States

**Aims:** PROMIS tools offer alternatives to legacy outcome measures. Recent studies suggest PROMIS tools may be appropriate for long-term outcome assessment in spine surgery populations. The aims are to extend previous validation work on the use of PROMIS in spine surgery cohorts by (1) evaluating the responsiveness of three PROMIS domains from preoperative to 24 months after lumbar spine surgery; (2) comparing the trajectory of PROMIS scores over time to legacy measures of disability and pain for patients with and without fusion. **Methods:** This secondary analysis from a randomized controlled trial comparing postoperative cognitive-behavioral based physical therapy (n = 124) and an education program (n = 124). Participants completed PROMIS CAT questionnaires preoperatively, 6 weeks, 6-, 12-, and 24-months after surgery for Physical Function (PF), Pain Interference (PI), and Ability to Participate in Social Roles and Activities (SR). Measures of disability (Oswestry Disability Index: ODI) and pain (NRS and BPI) were collected at the same timepoints. Mean scores were plotted at each timepoint for all outcome. Patients were stratified by fusion status. **Results:** Of 248 patients, 164 (66%) patients had laminectomy with fusion and 84 (34%) had laminectomy without fusion. PROMIS PF T-scores at preop for fusion and non-fusion were 33 (SD = 6), indicating moderate problems. At 24 months, mean scores for both groups for PROMIS PF were 43 (SD = 7), indicating mild problems. At preop, patients with and without fusion had PROMIS PI T-scores indicating moderate problems (both groups M = 66, SD = 6) and PROMIS SR T-scores indicating mild problems (no fusion: M = 41, SD = 7; fusion: M = 41, SD = 9). At 24 months PROMIS scores were in normal range (means = 52-53, SD = 7-8). The fusion group showed less improvement at 6 weeks, but both groups had similar scores by 6-12 months after surgery. The ODI and pain score trajectories showed a similar pattern of responses over time. **Conclusion:** PROMIS PF, PI, and SR show good responsiveness following lumbar spine surgery. Patients improve towards less symptomatic categories and exceed currently published meaningful change standards. These trajectories align with legacy outcome trajectories over the same time

period both overall and when stratifying by fusion status and support the use of PROMIS measures to assess outcomes in this population.

**Figure 1. Trajectories of PROMIS-PF and ODI by Fusion Status**



**(3150) Towards a core outcome set for distal tibia and ankle fractures (COSTA): a qualitative review exploring patient experience of injury and recovery**

Nathan Pearson, PhD, University of Warwick, Coventry, United Kingdom; Elizabeth Tutton, Oxford University, Oxford, United Kingdom; Alexander Joeris, AO ITC, Clinical Science, AO Foundation, Davos, Switzerland; Stephen Gwilym, Oxford University, Oxford, United Kingdom; Kirstie Haywood, University of Warwick, Coventry, United Kingdom

**Aims:** There is currently no guidance for distal tibia and ankle fractures relating to a core outcome set, which are evident in other fractures such as hip, distal radius, open tibia fractures. Little is known about what matters to patients, current qualitative evidence highlights the risk of persisting psychological distress and reduced physical and social functioning. This review is a synthesis of qualitative studies that explore patient experience of injury and recovery from distal tibia and ankle fractures. The findings will be used to develop a list of patient important outcome priorities for consideration in the development of a core outcome set for distal tibia and ankle fractures (COSTA). **Methods:** 5 databases were searched from inception to April 1, 2020. Methodological quality was assessed using the Critical Appraisal Skills Programme (CASP) checklist. A thematic synthesis of findings was performed, and a separate list of outcome priorities were extracted and mapped against the Outcome Measures in Rheumatology (OMERACT) taxonomy. Confidence in the findings of the thematic synthesis was assessed using the GRADE-CERQual checklist. **Results:** From 5842 hits, 12 studies met the inclusion criteria and were reviewed as full texts; 3 were included in the synthesis. Methodological quality was good. Findings identified the overarching concept of being changed conveyed through three themes: 1) challenges of adaption and resistance to changes; 2) being proactive, a determination to function within their capability, and 3) being the same but different, in all aspects of life. A range of patient important outcomes were identified for inclusion in future work. **Conclusion:** This review identifies how patients make the transition to a new way of being as a result of injury. Multidisciplinary interventions are required to support this transition, particularly for those who may not be able to return to their usual activities. A list of outcomes important to patients will be used to support future work to develop COSTA.

**Neurological & Mental Health Research**

**(3152) Individualized Goal Attainment Scaling as a patient-reported outcome measure for Alzheimer disease and dementia**

Justin Stanley, DGI Clinical Inc, Halifax, Nova Scotia, Canada; Sanja Stanojevic, PhD, DGI Clinical Inc, Halifax, Nova Scotia, Canada; Chere Chapman, MPH, MBA, DGI Clinical Inc, Halifax, Nova Scotia, Canada; Kari Knox, BScN, DGI Clinical Inc, Halifax, Nova Scotia, Canada

Scotia, Canada; Taylor Dunn, MSc, DGI Clinical Inc, Halifax, Nova Scotia, Canada; Susan Howlett, PhD, Dalhousie University, Halifax, Nova Scotia, Canada; Kenneth Rockwood, MD, Dalhousie University, Halifax, Nova Scotia, Canada

**Aims:** Interpreting small changes measured by traditional outcomes is a challenge for patients, caregivers, regulators and payers. In contrast, patient-reported outcomes measure treatment effects that are relevant to them. Goal Attainment Scaling (GAS) enhances relevance by being both patient-reported and individualized. With GAS, patients identify and track only the outcomes that are meaningful to them, while still offering a standardized scoring method. Here, we evaluate the utility of GAS in Alzheimer disease and dementia. **Methods:** Data from three studies were evaluated: Atlantic Canada Alzheimer Disease Investigation of Expectations (ACADIE), Video Imaging Synthesis of Treating Alzheimer disease (VISTA) and Vascular Aspects of dementia (VASPECT). GAS was used traditionally in ACADIE and VISTA whereas VASPECT used SymptomGuide®-dementia, a symptom tracking tool. A GAS score of 50 indicated no change from baseline. Responsiveness was evaluated with Standardized Response Means (SRM). Construct validity was evaluated with correlations between GAS scores and changes in standardized outcome measures and global measures of change. **Results:** Subjects were comparable between VISTA, ACADIE, and VASPECT ( $75.9 \pm 7.8$ ,  $77.2 \pm 7.8$  and  $75.4 \pm 9.2$  years of age; 73%, 63% and 52% women; baseline Mini-Mental State Examination [MMSE] scores  $19.7 \pm 5.2$ ,  $20.3 \pm 3.8$  and  $23.5 \pm 4.4$ , respectively). Mean GAS scores showed net goal attainment in all three trials. Mean  $\pm$  standard deviation GAS scores were  $54.3 \pm 9.5$  (ACADIE),  $53.7 \pm 12.5$  (VISTA) and  $54.9 \pm 16.2$  (VASPECT) which reflected good responsiveness (SRM = 0.45, 0.29 and 0.30, respectively). GAS scores were correlated with measures of clinical global impressions of change ( $\rho = -0.57$ ,  $-0.70$  and  $-0.66$ , respectively). GAS scores were also significantly correlated to standardized measures of function ( $r = 0.27$ ,  $0.38$ ,  $0.30$ ) and cognition ( $r = 0.26$ ,  $0.43$ ,  $0.37$ ) in each trial, respectively. **Conclusion:** GAS scores were valid and responsive to change in people with Alzheimer disease and dementia. GAS is an individualized patient-reported outcome that can be used to effectively measure outcomes that are meaningful to patients.

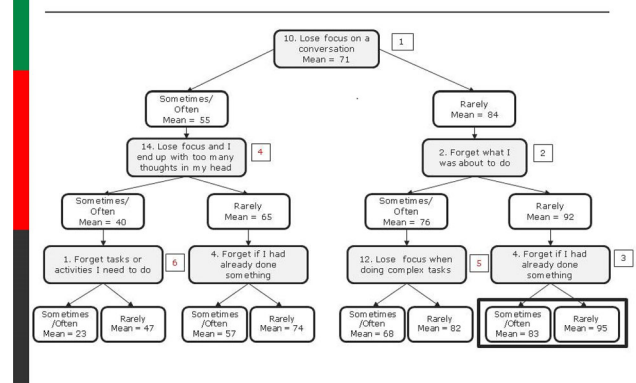
### (3153) Developing a computerized adaptive test (CAT) for the Communication Cognitive Concerns (C3Q) measure using decision regression trees

Nancy Mayo, McGill University, Montreal, Quebec, Canada; Lyne Nadeau, MSc, Center for Outcomes Research and Evaluation, Research Institute of McGill University Health Center, Montreal, Quebec, Canada; Marie-Josée Brouillette, MD, Department of Psychiatry, McGill University, Montreal, Quebec, Canada; Lesley Fellows, MD PhD, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada; Stanley Hum, PhD, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada; Sorayya Askari, OT PhD, School of Occupational Therapy, Dalhousie University, Montreal, Nova Scotia, Canada

**Aims:** As an alternative to Rasch analysis or Item Response Theory (IRT) for the development of computerized adaptive tests (CAT), we propose to use decision regression trees. This method is readily available in standard statistical packages and does not require a substantial item bank and is hence suitable for extant measures that would be made more useful if shortened. **Methods:** The Communication Cognitive Concerns (C3Q), a self-reported “voice-of-the-patient” measure of cognitive ability, was developed for use with people aging with HIV as brain health in this population is threatened from the infection and its treatment. Best-practices for measurement

development were used with almost 2000 people contributing to the content, item refinement, estimation of item hierarchy, and inter-pretability of scores with respect to convergent constructs. A total of 18 items with 3 response options for frequency of occurrence in the past week fit the Rasch model with some skew towards a greater proportion of people at the high end of cognitive ability (scored to range from 0 to 100). A shorter version would be clinically welcome especially one that could be part on a self-assessment app. Decision regression trees, a supervised machine learning model, is a method used to identify item responses that maximize the split of the total score into meaningful leaves. **Results:** Using all the data available from people with HIV from the Positive Brain Health Cohort, some 1288 data points. Figure 1 shows the results of analysis done using PROC HPSPPLIT (SAS 9.4), which identified 6 items that yielded 8 terminal leaves with mean scores of 95, 83, 82, 74, 68, 57, 47, and 23. Values on the C3Q above 82 have been shown to be related to successful aging in people with HIV. Three items (labeled 1, 2, 3) need to be administered to identify successful agers; 3 items (labeled 1, 4, and 6) need to be administered to identify people with serious cognitive concerns. **Conclusion:** This branching format would be very easy to put into a self-assessment app that patients and clinicians could use regularly to identify concerns patients have about their cognition and flag need for further evaluation.

Figure 1. Regression Tree of 6 C3Q Items Yielding 8 Distinct Groups



### (3154) Investigating different methods of calculating the impact of long-term conditions on quality of life and comparing the results across patient groups with eye disease, Parkinson's disease and dementia

Andrea Gibbons, PhD, University of Winchester, Winchester, United Kingdom; Dominic Ffytche, MD, Kings College London, London, United Kingdom; Clare Bradley, PhD, HPR Ltd and Royal Holloway University of London, Egham, United Kingdom

**Aims:** It is often argued that generic measures are needed to compare the impact of different conditions on quality of life even though condition-specific measures provide greater relevance. However, the individualized condition-specific ADDQoL (Audit of Diabetes-Dependent Quality of Life © Bradley) shares its template and scoring metrics with comparable measures for other conditions. We evaluate three different methods of scoring QoL across conditions using disease-specific -DQoL measures. **Methods:** In SHAPED (Study of Hallucinations in Parkinson's, Eye Disease and Dementia) 89 people with eye disease, 59 with Parkinson's disease and 49 people with dementia

completed the EyeDQoL, Parkinson'sDQoL, or MemoryDQoL respectively. These questionnaires are individualized by providing 'not applicable' options for items not universally relevant and by weighting impact ratings (-3 to +1) by importance to the individual's QoL (3 to 0). The average weighted impact (AWI) of the condition on QoL is usually calculated by averaging the weighted-impact scores for all applicable domains. However, -DQoLs for different conditions vary in the number of items which may affect comparability across conditions. Here AWI scores were calculated in three ways: 1) the standard method (above); 2) the ten items eliciting the most negative weighted-impact scores per patient group were identified, and only these items used to calculate AWI; and 3) the ten items eliciting the most negative weighted-impact for each individual were used to calculate AWI. One-way ANOVAs assessed differences between the three methods of calculating AWI across conditions. Results: As expected, across all three conditions, method 3 shows the most negative impact of the condition on QoL and method 1 the least. When comparing the three conditions, differences were only found with AWI scores calculated using the ten items with the most negative weighted impact per individual. The dementia group reported significantly less negative weighted impact than the eye group ( $p = 0.048$ ) and Parkinson's group ( $p = 0.024$ ). The eye and Parkinson's groups did not differ significantly. Conclusion: These template-sharing condition-specific QoL measures can be used to compare QoL impact across conditions. Using the most individualized method of scoring increased both the negative impact of the condition on QoL and differentiation between conditions.

### (3155) Impacts of acute stress reaction and self-leadership on quality of life in college students during the COVID-19 outbreak in China

Yuan He, MS, Université Paris-Saclay, Paris, France; Yanbo Zhu, PhD, MD, Beijing University of Chinese Medicine, Beijing, China; Qian Zhang, PhD, Beijing University of Chinese Medicine, Beijing, China; Xiaoying Lv, PhD, Beijing University of Chinese Medicine, Beijing, China

Aims: Self-leadership is referred to as self-influencing actions and thoughts towards individual objectives. This study examines the impact of acute stress reaction (ASR) and self-leadership on college students' quality of life (QOL) during the COVID-19 outbreak in China. Methods: 274 valid samples of college students were collected from 10th February to 5th March 2020 through a set of internet-based questionnaires, which includes general information, epidemic risk perceptions, the Stanford Acute Stress Reaction Scale (SASRQ), the Revised Self Leadership Questionnaire (RSLQ), and the WHOQOL-BREF. Results: (1) 274 college students' score of ASR towards the COVID-19 was  $(23.09 \pm 20.46)$ , and there were 52 students with acute stress disorder (ASD) (SASRQ total score  $\geq 40$  points, accounting for 18.91%); the average score of the items from 9 dimensions of RSLQ (self-goal setting, self-reward, self-punishment, self-observation, self-cueing, focusing thoughts on natural rewards, visualizing successful performance, self-talk, evaluating beliefs and assumptions) was  $(3.35 \pm 0.79)$  to  $(3.86 \pm 0.70)$  points, which was in the upper-middle level. The quality of life score assessed by WHOQOL-BREF was: physical points  $(67.77 \pm 15.66)$ , psychological points  $(63.85 \pm 16.62)$ , social points  $(65.12 \pm 17.53)$ , environmental points  $(62.57 \pm 15.39)$ , and the total score was  $(64.83 \pm 14.07)$ . (2) Correlation analysis results showed that ASR was negatively correlated with QOL ( $r = -0.558$ ,  $p < 0.05$ ); it was negatively correlated with self-goal setting, self-observation, focusing thoughts on natural rewards, evaluating beliefs and assumptions, and visualizing successful performance ( $r = -0.138$  to  $-0.254$ ,

$p < 0.05$ ); it was positively correlated with self-punishment ( $r = 0.130$ ,  $p < 0.05$ ). The self-punishment under the theory of self-leadership ( $r = -0.160$ ,  $p < 0.05$ ) was negatively correlated with QOL, and the other 8 dimensions were positively correlated with QOL ( $r = 0.142$  to  $r = 0.415$ ,  $p < 0.05$ ). (3) Stratified stepwise regression analysis showed that the total explanatory variation of the independent variables included in this study was 48.8%; of which the ASR negatively affected QOL ( $b' = -0.550$ ,  $t = -10.99$ ,  $p = 0.000$ ), the explainable variation was 25.8%; the explainable variation of self-leadership was 16.4%, of which self-punishment negatively affected with QOL ( $b' = -0.264$ ,  $t = -5.51$ ,  $p = 0.000$ ), and focusing thoughts on natural rewards positively affected with QOL ( $b' = 0.431$ ,  $t = 8.91$ ,  $p = 0.000$ ). Conclusion: Both ASR and self-leadership affect the quality of life of college students under COVID-19. Self-leadership can provide positive internal guidance to emotions and behaviors that may release ASR and improve QOL.

### (3156) Functional profiles of people with multiple sclerosis receiving inpatient rehabilitation in Canada

Kedar K.V. Mate, Mayo Clinic/McGill University Health Centre, Montreal, Quebec, Canada; Nikki Ow, OT, PhD candidate, McGill University, Montreal, Quebec, Canada; Stanley Hum, PhD., McGill University, Montreal, Quebec, Canada; Nancy Mayo, PT, PhD., McGill University, Montreal, Quebec, Canada

Aims: Multiple Sclerosis (MS) is a chronic neurodegenerative condition affecting young adults during the peak productive years of their lives. With progression of MS, quality of life (QOL) becomes more closely coupled with independence and hence timely access to rehabilitation is important for QOL. It is important to understand what profile of disability will benefit from rehabilitation to best allocate these scarce resources. The aim of this study was to estimate the extent to which functional profiles, using Functional Independence Measure (FIM), of adults with MS differ from admission to discharge. Methods: Data were obtained from the Canadian Institute of Health Information's National Rehabilitation Reporting System (2001-2017). The FIM comprises five domains: activities of daily living (ADL), transfers, locomotion, sphincter control, and social cognition rated as by degree of independence. Latent class analysis (LCA) identified profiles of FIM domains at admission and discharge. These domains were rated according to whether total, maximal, or moderate assistance was required (0), minimal assistance; (1) supervision (2); or no assistance (independent 3). Results: 3500 adults with MS spent on average 40 days in rehabilitation. A five-class model fit the data at admission and a four-class model fit at discharge. Figure 1 shows these profiles ordered from left to right by degree of disability. Of note is that all people at admission were dependent for locomotion (0: in red). At discharge 28% were independent in locomotion (Class A;  $n = 946$ ). At admission, 81% needed some form of assistance with ADLs (Classes C,D,E) but, at discharge only 43%, were so impaired (Classes D,E). There was a trend for profiles of less disability over time except that the profile of greatest disability (Class E) was larger at discharge ( $n = 771$ ; 23%) than admission ( $n = 523$ ; 15%). Conclusion: The profiles are clinically coherent and confirm that inpatient rehabilitation in Canada is reserved exclusively for the most disabled adults with MS, particularly those unable to walk without assistance. Substantial proportions of people made gains in ADLs, transfers, and locomotion, gains that would directly impact their QOL. These data support allocating rehabilitation services to this population.

**Figure 1. Latent Classes of FIM Domains at Admission and at Discharge for People with Multiple Sclerosis in In-Patient Rehabilitation in Canada**

	Class A n=208 (6%)	Class B n= 465 (13%)	Class C n=1205 (34%)	Class D n=1009 (29%)	Class E n=523 (15%)
Admission					
ADL	3	2	1	1	0
Transfers	3	0, 1, 2	1, 2	1	0
Locomotion	0	3	1	1	0
Bowel	3	3	3	3	0
Cognition	3	3	3	3	3
Discharge					
ADL	3	3	1	1	0
Transfers	3	2	1	1	0
Locomotion	3	3	1	1	0
Bowel	3	3	3	3	0
Cognition	3	3	3	3	3

KEY	
0	Total to moderate assistance
1	Minimal assistance
2	Supervision
3	Independence

### (3157) Functioning and disability in neurosarcoidosis from the patient perspective

Brandon Moss, MD, Cleveland Clinic Foundation, Cleveland, Ohio, United States; Deborah Miller, PhD, Cleveland Clinic Foundation, Cleveland, Ohio, United States; Daniel Culver, MD, Cleveland Clinic Foundation, Cleveland, Ohio, United States

**Aims:** NS represents a neglected, costly, and often disabling disease, characterized by its heterogeneity and variability. Its impact on disability has rarely been studied systematically and never from the patient perspective. We conducted a series of focus groups to characterize impairment and disability in neurosarcoidosis (NS) from the patient perspective. **Methods:** Subjects were recruited from our site and Foundation for Sarcoidosis Research patient conferences in Birmingham, AL, Durham, NC, and Iowa City, IA. Subjects were 18 or older and had probable or definite NS, defined by the NS Consortium Consensus Diagnostic Criteria. Focus groups were co-moderated by a trained PhD-level facilitator and a physician with experience in treating NS. A topic guide with pre-defined, open-ended questions was used. Focus groups continued until we reached saturation of themes. All discussions were transcribed verbatim. Thematic coding was performed based on grounded theory. **Results:** There were 18 subjects, mean age 49 (range 34–61), 15 (83%) female, 5 (28%) African American (the rest Caucasian). Average duration of diagnosis was 9 years (range 2–23). Variable levels of disability were reflected by employment status, 5 (28%) full-time, 1 (6%) student, 2 (11%) retired, 9 (50%) disabled, 1 (6%) other. Subjects stressed the importance of accounting for heterogeneity and variability of the disease and side effects of treatment. Upper extremity function, lower extremity function, vision, cognition, fatigue, pain, and bowel and bladder function were identified as key functional domains for inclusion in outcome measures. **Conclusion:** Patient-centered outcome measures in NS should account for heterogeneity and variability of the disease, assess treatment side effects, and include key functional domains important to patients' health-related quality of life.

### (3158) More than memory problems: the impact of multimorbidity on quality of life for memory clinic patients and their carers

Caroline Potter, PhD, Health Services Research Unit, University of Oxford, Oxford, United Kingdom; Ray Fitzpatrick, PhD, Health Services Research Unit, University of Oxford, Oxford, United Kingdom; Michele Peters, PhD, Health Services Research Unit, University of Oxford, Oxford, United Kingdom

**Aims:** Cognitive impairment including dementia is known to decrease quality of life, but the cumulative impact of multiple long-term conditions (LTCs) on this patient group is not well understood. People with cognitive impairment are often supported by informal carers, but it is not known how carers' own experiences of multimorbidity (presence of more than one LTC) in addition to their caring responsibilities impact quality of life. In this study we explored the prevalence of multimorbidity and impact on quality of life among patients and carers in the immediate post-diagnosis period. **Methods:** Participants were recruited through one of 14 memory clinics in South-East England, at a first assessment appointment leading to a diagnosis of mild cognitive impairment (MCI) or dementia. Shortly after diagnosis participants completed a survey at home that included demographics, self-reported health conditions, and two quality-of-life measures: the Long-Term Conditions Questionnaire (LTCQ) or its associated carer measure (LTCQ-Carer), and the EuroQol five-dimensional index with visual analogue scale (EQ-5D-5L with EQ-VAS). Descriptive statistics and analysis of variance were used to compare quality-of-life scores for participants with zero, one, and two or more comorbidities in addition to their primary recruitment status (cognitive impairment/caring responsibilities). **Results:** Comorbidity was high among both groups: 78% of patients (n = 105) and 57% of carers (n = 107) were living with at least one LTC in addition to the recruiting condition, with 25% of carers reporting multimorbidity. Hypertension, arthritis, and depression were the most commonly reported comorbidities for both groups. For MCI/dementia patients, a statistically significant decrease in quality-of-life scores was observed for all measures as the number of comorbidities increased. Decreased quality-of-life scores were observed for carers with multimorbidity, with statistically significant differences in EQ-5D-5L and EQ-VAS scores. **Conclusion:** At the start of the clinical pathway for cognitive impairment, patients and carers with multimorbidity already experience decreased quality of life. Outcomes of MCI/dementia/carer support services might vary in relation to multimorbidity status, but these data are not routinely collected. Large-scale analyses of multimorbidity patterns in people affected by cognitive impairment are needed to inform person-centered approaches to policy and health services for this group.

**Table: Quality of life scores by multimorbidity status in patients with cognitive impairment and their carers**

	N	Mean LTCQ	Mean EQ-5D-5L index		Mean EQ-VAS	
			One-way ANOVA	One-way ANOVA	One-way ANOVA	One-way ANOVA
Patients	No other LTCs	23	74.6	0.85	78.9	
	1 other LTC	40	77.0	0.84	76.9	F(2,99) = 5.61
	2+ other LTCs	42	63.1	0.78	66.2	p=0.002**
	Total	105	71.0	0.78	73.2	p=0.005**
Carers	No other LTCs	46	74.5	0.89	80.6	
	1 other LTC	34	74.4	0.84	77.4	F(2,103) = 5.15
	2+ other LTCs	27	66.9	0.74	65.9	p=0.217 (ns)
	Total	107	72.5	0.84	75.8	p=0.003**

LTCQ = Long-Term Conditions Questionnaire, score range 0 to 100, higher scores 'living well with LTCs'

LTCQ-Carer = Long-Term Conditions Questionnaire for Carers, score range 0 to 100, higher scores 'living well while caring for someone with LTCs'

EQ-5D-5L index = EuroQol five-dimensional five-level index, score range -0.285 to 1.0, higher scores 'better health-related quality of life'

EQ-VAS = EuroQol visual analogue scale, score range 0 to 100, higher scores 'better health-related quality of life'

ANOVA = one-way analysis of variance, \*\*statistical significance at p<0.01 level

### (3159) Is there seasonal variation in activity and participation post stroke?

Ruth Barclay, PhD, MHS, BMR(PT), College of Rehabilitation Sciences, University of Manitoba, Winnipeg, Manitoba, Canada; Sandra Webber, PhD, MSc, BMR(PT), College of Rehabilitation Sciences, University of Manitoba, Winnipeg, Manitoba, Canada;

Leanne Leclair, PhD, MSc, BHSc(OT), College of Rehabilitation Sciences, University of Manitoba, Winnipeg, Manitoba, Canada

**Aims:** People with stroke have identified weather conditions as barriers to activity and participation. This study aimed to: 1. Quantify and compare summer and winter activity and participation. 2. Explore how community dwelling people with stroke describe their feelings about and thoughts on their level of activity and participation in winter and summer. **Methods:** This embedded mixed-methods feasibility study occurred in a city with a mean temperature difference July–January of  $\sim 36$  °C. Qualitative results helped explain the quantitative. Participants were community dwelling individuals at least one year post stroke, walking 50 metres with or without a walking aide. Measurements occurred at participants' homes in summer and winter months: Reintegration to Normal living Index, Stroke Impact Scale recovery, Activities-specific Balance Confidence, the Timed-up-and-Go, and the interview Chedoke-McMaster Stroke Assessment Activity Inventory. They wore an Actigraph GT3X + activity monitor for 1 week each season, resulting in steps/day and cadence data. Analysis included descriptive statistics and paired t-tests from winter to summer. Interested participants were interviewed in their home following use of the Actigraph. An inductive approach to content analysis was taken. **Results:** There were no differences between winter-summer values of self-perceived, observed or activity monitor outcomes ( $n = 13$ , mean age 61.5 years, 6.2 years post-stroke, 62% females). Despite these results, participants ( $n = 8$ ) described many challenges in winter: participation and activity limitations, walking aide difficulties, fear of falling, inclement weather, decreased travel ability and difficulty with activities of daily living. However, participants described still being able to participate and stay active in winter for a number of reasons: doing things despite bad weather, being less independent, tasks taking longer and being more difficult; finding other ways to keep active; still participating due to social supports; and the ability for some to winter in warmer climates. **Conclusion:** We had expected participation scores and activity monitor data would be different winter to summer. The qualitative findings provided insights as to why this did not occur. Participants described many challenges with winter weather, but also ways they continued to participate and be active.

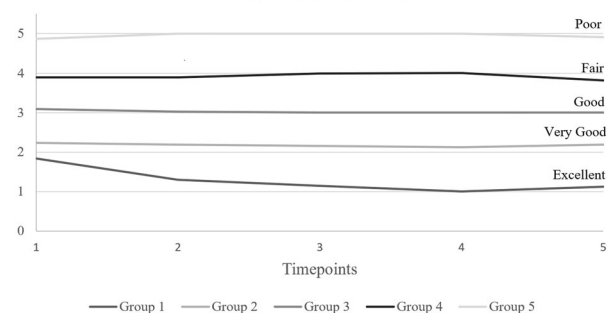
### (3160) Age and sex differences in the trajectories of general health perception of young people with MS

Nikki Ow, McGill University, Montreal, Quebec, Canada; Ayse Kuspinar, PhD, McMaster University, Hamilton, Ontario, Canada; Nancy E. Mayo, PhD, McGill University, Montreal, Quebec, Canada

**Aims:** Recent evidence has suggested the existence of a multiple sclerosis (MS) prodrome, the presence of MS symptoms before the first clinical event. Hence, some young adults with MS are very likely to have had symptoms in childhood or adolescence. It is, therefore, reasonable to assume that people aged under 25 years with MS might have had pediatric-onset. In contrast, young people aged between 26 to 35 are less likely to have had pediatric-onset but are not affected by age-related changes. Contrasting these two groups of people could lead to valuable information about the impact of MS over time. The purpose of this study is to characterize how general health perception in young people with MS changed over time and to estimate the extent to which general health perception differs between age groups (18 to 25 years and 26 to 35 years) and sex. **Methods:** This study utilized placebo arm data from the Multiple Sclerosis Outcome Assessment Consortium database. Responses to the RAND-36 General Health Perception item of 393 participants were included. Group-

based trajectory models were used to identify patterns of change over two years. Ordinal regression was performed to estimate whether these trajectory groups differed by age group, sex and relapse event. **Results:** Using group-based trajectory modeling, a five-group model was chosen (Fig. 1). Over time, all groups were stable except one progressing from a rating of “very good” to “excellent.” Posterior probabilities showed that 35% of people consistently rated their health “very good or excellent” and 2% consistently rated their health as “poor.” **Results:** of ordinal regression showed that health perceptions differed across age groups ( $b = 0.5$ , OR 1.7 CI 1.1, 2.6) but not sex ( $b = -0.1$ , OR 0.9 CI 0.6 1.3). Compared to the younger group, people aged 26 to 35 years are 1.7 times more likely to rate their health poorer. People with relapses are also 2.2 times more likely to rate their health poorer ( $b = 0.8$ , OR 2.2 CI 1.5, 3.2). **Conclusion:** General health perception of young people with MS is relatively stable, suggesting that having relapses does not change young people's health perception over time.

Graph of two-year trajectories on the SF-36 – General Health Perception (5 group model)



### Psychometrics & Statistics

#### (3161) The relevance of revealed preference in healthcare: a scoping review of the literature

Paul Williams, IQVIA, Paris, France; Laurie Batchelder, IQVIA, Reading, United Kingdom; Stephanie Philpott, IQVIA, Reading, United Kingdom; Ana Maria Rodriguez, IQVIA, Madrid, Spain

**Aims:** Stated preferences (SP), where choices are made in response to hypothetical scenarios, are commonly used in a variety of settings but may differ from the reality of clinical care. Revealed preferences (RP) are derived from actual choices made. They can result in more reliable preference estimates than SPs. The purpose of this scoping review was to determine the type of RP-focused studies in healthcare research. **Methods:** Online databases (PubMed and Google Scholar) were searched to identify published studies with specific mention of ‘revealed preference’ in the title or abstract. All articles were double screened for report of RP in original studies within healthcare fields and descriptive features of the studies were extracted. **Results:** Of 67 records, 17 were included. Studies investigated drugs ( $n = 6$ ), devices ( $n = 1$ ), public health interventions ( $n = 4$ ), medical system utilization ( $n = 4$ ), level of activity in elderly ( $n = 1$ ) and informal care ( $n = 1$ ). The majority of studies investigated available products or services ( $n = 15$ ). The most common objectives were identifying factors that predicted choice decisions ( $n = 6$ ) and combining RP with SP data for modeling ( $n = 4$ ) followed by providing external validity to SP models ( $n = 2$ ), eliciting cost information ( $n = 2$ ), descriptive only ( $n = 2$ ) and informing development of an SP instrument ( $n = 1$ ). Data were collected from primary ( $n = 11$ ) and secondary ( $n = 6$ )

data sources. Most studies were conducted in the healthy population ( $n = 7$ ), followed by patients ( $n = 6$ ), caregivers ( $n = 1$ ), patients and caregivers ( $n = 1$ ), and physicians ( $n = 1$ ). Regression was the most frequently used method to analyze RP data ( $n = 11$ ), followed by descriptive statistics ( $n = 4$ ) or other ( $n = 2$ ). Of the 4 studies combining RP and SP, 2 reported better model prediction as additional parameters were identified from RP data and 2 showed improved model fit. Conclusion: Although RP studies can provide more reliable preference estimates than SPs, findings from this review suggest that RP studies are not common. Where conducted, the objectives of RP studies are heterogeneous, and RP studies are often conducted for other purposes than addressing RP objectives, for example providing external validity to SPs. There is scope for more RP studies to be conducted with clear objectives around understanding preference decisions in healthcare.

**(3162) Should we calibrate PRO instruments in clinical trials? A comprehensive simulation study based on Rasch Measurement Theory**

Angély Loubert, PharmD, MSc, Modus Outcomes, Lyon, France; Antoine Regnault, PhD, Modus Outcomes, Lyon, France; Véronique Sébille, PhD, UMR INSERM 1246—SPHERE, Nantes, France; Stefan Cano, PhD, CPsychol, AFBPsS, Modus Outcomes, Letchworth, United Kingdom; Jean-Benoit Hardouin, PhD, UMR INSERM 1246—SPHERE, Nantes, France

Aims: Calibration allows measuring instruments to produce reliable and consistent measurements. While it is an established process for physical measures, its use for non-physical measures, such as patient-reported outcomes (PROs) in clinical trials, is not standard practice. Rasch measurement theory (RMT) is a psychometrics approach that guarantees algebraic separation of person and item parameter estimates, allowing formal calibration of measurement instruments. In the RMT framework, two approaches can be distinguished for the analysis of clinical trial data: (1) Calibrated approach, where a previous “validation” study is used to estimate item parameters, allowing for comparison of scores between studies and (2) Uncalibrated approach where item parameters are estimated from the current trial data, optimizing the estimation to the trial context. This simulation study aims at assessing the impact of calibration on the comparison of PRO results between trial arms, using different analysis methods. Methods: PRO results are simulated following a Rasch model, for a validation and a trial sample. Scenarios include varying sample sizes (50 to 1000), with instrument of varying number of items (4 to 20) and modalities (3 or 5). Different treatment effect sizes and distributions of the 2 samples are also explored. Comparison of treatment groups is performed using statistical testing of the person parameter estimates and modeling using random-effect Rasch model. Calibrated and uncalibrated approaches are compared based on bias, power, alpha risk and variance of the estimate for difference between arms. Results: The calibrated and uncalibrated approaches showed comparable power (26%, 76% and 100% for different effect sizes) and alpha risk (5%). The calibrated approach led to a slightly higher bias in the estimate of treatment effect, which increased when the samples distributions highly differed. Conclusion: Overall statistical performance of the calibrated an uncalibrated approach was very similar. The theoretical benefits of a calibrated approach thus justify its use in most cases. Calibration is more questionable when validation and trial samples highly differ. Validation studies should thus include a heterogenous sample of patients to cover a wide spectrum of the measured concept to create item parameter estimates that are valid in most contexts.

**(3163) Estimating measurement error of the Oswestry Disability Index with partially missing data: Guidance for clinical research and regulatory decision-making**

Richard Skolasky, Sc.D., Johns Hopkins University, Baltimore, Maryland, United States; Emmanuel McNeely, M.S., M.H.A., Johns Hopkins University, Baltimore, Maryland, United States; Brian Neuman, MD, Johns Hopkins University, Baltimore, Maryland, United States

Aims: Patient-reported outcomes are increasingly used for clinical and regulatory decision-making. The Oswestry Disability Index (ODI) is the most commonly used outcome measure for low back pain with scores from 0-100 indicating percent disabled and a disability level (minimal to bedridden). To assess the effectiveness of surgical implants on treatment of lumbar spine pathology, researchers and decision-makers must deal with missing data. While there are many ways to handle partially missing data (e.g., alternate scoring or multiple imputation), the impact on score is not well understood. We estimate the measurement error introduced by increasing missingness in the ODI. Complete data were available from a cohort of patients presenting to an academic orthopedic surgery department for lumbar spine surgery. Patients were mainly older (mean age 57 years), female (51%), and non-Hispanic (95%) White (79%). Each patient ( $n = 991$ ) completed the ODI (mean 45.3 standard deviation, 18.5) with 108 minimal, 278 moderate, 395 severe, 195 crippled, and 15 bedridden. Methods: The following steps were taken across the full range of missing items (1–9): We created missing data by randomly setting items to missing over 1,000 simulations; scored ODI using two methods (an alternate scoring based on Fairbanks, et al. or multiple imputation using STATA MI procedure); calculated absolute percentage error (APE) from true score and classification of level of disability using the simulated score; and compared the measurement error and misclassification of level of disability between the two scoring methods. Results: Using alternate scoring, APE increased with number of missing items [APE1 0.7%; APE5 3.1%, APE9 11.6%] but decreased as level of disability increased (Fig 1). Misclassification (MS) increased with missing data [MS1 46; MS5 82; MS9 579]. Using multiple imputation, APE increased with number of missing items [APE1 1.3%, APE5 7.9%; APE9 55.6%] but decreased as level of disability increased (Fig 2). Misclassification increased with missing data [MS1 46; MS5 71; MS9 577]. Conclusion: At each level of missing data, multiple imputation introduced more measurement error but similar misclassification. Researchers and policy-makers should use caution when interpreting results with missing data, as inferences may be unduly influenced by induced measurement error or misclassification of disease severity.

Missing	OVERALL APE			Level 1 APE			Level 2 APE			Level 3 APE			Level 4 APE			Level 5 APE		
	Mean	min	max	Mean	min	max	Mean	min	max	Mean	min	max	Mean	min	max	Mean	min	max
1	0.7%	0.2%	8.7%	1.04%	0.2%	10.2%	1.0%	0.8%	10.8%	0.5%	0.3%	9.9%	0.4%	0.3%	9.0%	0.3%	0.2%	8.0%
2	1.1%	0.5%	17.4%	2.3%	0.6%	22.0%	1.5%	0.8%	21.9%	0.8%	0.4%	14.9%	0.7%	0.5%	11.3%	0.2%	0.0%	7.0%
3	1.7%	0.4%	26.1%	3.6%	0.8%	38.2%	2.2%	0.4%	33.6%	1.2%	0.5%	23.5%	0.9%	0.4%	15.5%	0.3%	0.0%	9.3%
4	2.3%	0.8%	35.4%	5.1%	1.4%	57.5%	3.0%	0.5%	46.1%	1.6%	0.6%	30.6%	1.3%	0.5%	18.5%	0.4%	0.0%	11.0%
5	3.1%	0.7%	45.4%	7.4%	1.8%	81.3%	4.1%	0.6%	60.2%	2.2%	0.7%	37.5%	2.1%	0.6%	24.0%	0.5%	0.0%	13.2%
6	4.1%	1.0%	56.7%	9.8%	2.6%	112.7%	5.4%	0.8%	76.0%	2.9%	0.9%	46.2%	3.1%	0.7%	27.8%	0.5%	0.0%	15.3%
7	5.6%	1.1%	70.1%	13.6%	3.1%	158.6%	7.1%	0.9%	93.0%	3.9%	1.0%	55.0%	3.9%	0.8%	36.9%	0.5%	0.0%	16.1%
8	7.8%	0.9%	86.3%	19.2%	3.7%	224.0%	10.0%	0.7%	110.5%	5.4%	1.4%	64.4%	5.7%	0.8%	45.8%	0.6%	0.0%	16.1%
9	15.4%	4.7%	87.0%	30.0%	13.7%	455.5%	41.3%	1.7%	93.1%	11.1%	4.0%	68.8%	21.1%	3.0%	53.5%	0.7%	0.0%	20.1%

Figure 1: Average absolute percent errors (APEs) between true ODI mean scores and simulated mean, minimum, and maximum scores, overall and separately by ODI levels

### (3164) Performance of using the Best–Worst Scaling technique in eliciting health state utilities in mainland China

Ahmed M.Y. Osman, MSc, School of Pharmaceutical Science and Technology, Tianjin University and Center for Social Science Survey and Data, Tianjin University, Tianjin, China; Jing Wu, PhD, School of Pharmaceutical Science and Technology, Tianjin University and Center for Social Science Survey and Data, Tianjin University, Tianjin, China; Xiaoning He, PhD, School of Pharmaceutical Science and Technology, Tianjin University and Center for Social Science Survey and Data, Tianjin University, Tianjin, China; Gang Chen, PhD, Centre of Health Economics, Monash Business School, Monash University, Melbourne, Australia

**Aims:** The ordinal data collection methods, such as discrete choice experiments, have increasingly been used to develop tariffs (value sets) for preference-based instruments. The evidence for using the best–worst scaling (BWS) method in eliciting health state preference is still scarce. This paper aims to evaluate the performance of applying the BWS technique to develop a Chinese-specific tariff for the Short Form 6 Dimension version 2 (SF-6Dv2). **Methods:** An online survey was conducted among the general population of mainland China, recruited from an online panel company, through age, gender, geographical allocation, educational level, and household registration (HuKou) quota. Each participant was randomized to different valuation tasks to elicit their personal preference for several health states described by SF-6Dv2. The study sample completed eleven BWS tasks alongside other questions, including subjective face validity, sociodemographics and health situation. Conditional (Multinomial) Logistic Regression (clogit) was used to analyze the aggregate sample, and Hierarchical Bayes (HB) was used to explore heterogeneity among individuals and to generate disaggregate models. The tariff scores were calculated directly from anchoring each model to death. Anchored scores were evaluated based on monotonicity and model-fit statistics. **Results:** After excluding responses from repeated IP-Addresses and an indicator of logical choices, 463 responses were eligible for analysis. Most of the participants were male, Han (ethnicity), married and had no to little health problems. The majority stated ease of understanding and answering the tasks. Statistical evidence was against the pool-ability of the best, and worst data and models based on the best choices only were superior. Models based on HB performed better than those of clogit in both fitting the data (RLH 0.47 vs 0.26) and prediction (McFadden's rho-squared 58.2 vs 26.4). HB models showed a degree of preference difference between subgroups. Tariff scores for health states were all better than death and ranged, for the pits state, from 0.170 (clogit) to 0.197 (HB). HB had better differentiation ability for the light states. **Conclusion:** BWS can be applied alone to obtain health state utilities, as demonstrated by an online panel in mainland China. HB is a potential modeling technique that generates the flexibility of producing personalized preference values.

### (3165) Risk of Bias tool to assess the quality of studies on reliability and measurement error of outcome measurement instruments

Lidwine Mokkink, Amsterdam UMC, Amsterdam, Netherlands; Maarten Boers, prof, Amsterdam UMC, Amsterdam, the Netherlands; Cees van der Vleuten, prof, Maastricht University, Maastricht, the Netherlands; Lex Bouter, prof, Amsterdam UMC, Amsterdam, the Netherlands; Jordi Alonso, prof, IMIM-Institut Hospital del Mar d'Investigacions Mèdiques, Barcelona, Spain; Donald Patrick, prof, University of Washington, Barcelona, Washington, United States; Henrica de Vet, prof, Amsterdam UMC, Amsterdam, the Netherlands; Caroline Terwee, dr, Amsterdam UMC, Amsterdam, the Netherlands

**Aims:** Reliability and measurement error are important quality aspect of outcome measurement instruments, which should be taken into account when selecting an instrument. Studies on reliability (i.e., ability to distinguish between people) and measurement error (which refers to how close scores of repeated measurements in stable patients are) can be complex to perform and understand, as many different sources of variance can play a role in the design of these studies. The aim of this Delphi study was to develop the COSMIN Risk of Bias tool for assessing the quality of studies on reliability and measurement error. This tool will be developed for researchers and clinicians who may not be familiar with all aspects of reliability, but who need to understand reliability studies when selecting their outcome measurement instruments. Previously, we focused on patient-reported outcome measures (PROMs). In this study we focus on any type of instruments, such as performance-based tests, clinical scales, imaging modalities or laboratory values. **Methods:** We conducted a three round online Delphi study among international experts; consensus was set at 67% agreement on a 5-point Likert scale. Proposals (e.g., on risk-of-bias items) were based on a literature search, were in line with current COSMIN terminology and the Risk of Bias checklist for PROMs. Arguments for their ratings were asked for each proposal, enabling us to understand panelist, improve our proposals, and explain our decisions in a user manual. **Results:** We invited 175 experts, 45 completed Round 1. Round 2 is ongoing. We reached consensus on components of measurement instruments, on how to formulate a specific research question for reliability studies; and on five standards for design requirements for these studies. Three standards on preferred statistical methods for reliability studies, and two standards on preferred statistical methods for studies on measurement error are being discussed. **Conclusion:** The tool containing these standards can be used to assess the quality of an existing study on reliability or measurement error to understand whether the results of the study can be trusted. We aim to improve the quality of future reliability studies.

### (3166) Longitudinal evaluation of patient-reported outcomes: within-individual correlation and identification of outliers

Zachary Hopkins, MD, Broward Health, Fort Lauderdale, Florida, United States; Vanina Taliercio, MD, MS, University of Utah, Salt Lake City, Utah, United States; Ashley Snyder, MPH, University of Utah, Salt Lake City, Utah, United States; Gregory Stoddard, MPH, University of Utah, Salt Lake City, Utah, United States; Kristina Duffin, MD, MS, University of Utah, Salt Lake City, Utah, United States; Jacob Kean, PhD, University of Utah, Salt Lake City, Utah, United States; Rachel Hess, MD, MS, University of Utah, Salt Lake City, Utah, United States; Aaron Secrest, MD, PhD, University of Utah, Salt Lake City, Utah, United States

**Aims:** Clinicians can gain insight into disease burden from their patients' perspective through the regular clinical use of relevant patient-reported outcomes (PROs). Using psoriasis as a model, we explored intra-individual alignment of PRO (Skindex-16) scores over time with a psoriasis-specific clinician-reported outcome (CRO) called the overall Physician Global Assessment (OPGA). Our aim was to identify the intra-individual correlation of scores in psoriasis patients seen longitudinally. **Methods:** We assessed 619 visits from 248 patients. Intra-individual correlation over time was assessed using an adapted ANCOVA model (rmcorr R package). We calculated 10th and 90th percentiles of Skindex-16 scores at each OPGA score (0-4). For discordant case identification: Clear skin was defined as OPGA = 0 (10th global percentile OPGA) and severe skin was OPGA33 (90th percentile global OPGA). Two groups, "clear skin, poor QoL" (OPGA = 0; Skindex-16 > 90th percentile) and "severe skin, good QoL" (OPGA3/4; Skindex-16 < 10th percentile) were



flagged. Predictors of these two discordant situations were assessed using univariable logistic regression. Results: After comparing each case graphically, most slopes over time for the two scores were largely parallel; however, several discordant situations were observed. Intra-individual correlations between scores were higher, but still moderate, compared with cross-sectional between-individual values reported separately. These were overall ( $r = 0.61$ ), emotion (0.59), functioning (0.53), and symptoms ( $r = 0.56$ ) Skindex-16 ( $p < 0.001$  for all). “Clear skin, poor QoL” was seen in 17 (2.8%) cases for overall Skindex-16, and “Severe skin, good QoL” was seen in 14 visits (2.3%). Among these, one and two patients, respectively, had  $> 1$  discordant visit. All other patient visits showed discordance at only 1 visit. For overall score, no predictors of “Severe skin, good QoL” were observed. For “Clear skin, poor QoL,” female sex (OR 4.4, 95% CI 1.3–15.4), other dermatologic conditions (OR 3.0, 1.1–8.7) and number of other diagnoses included on visit encounter (OR 1.3, 1.0–1.7) all predicted this type of encounter. Age was not a significant predictor for either discordant visit type. Conclusion: Intra-individual Skindex-16 (a PRO) moderately correlates with OPGA (a CRO) over time in psoriasis patients, and both scores move concordantly in most patients. Although rare, visits with discordant scores can be a helpful prompt for more in-depth discussion.

### (3167) Describing patterns in longitudinal diary data

Brandon Foster, PhD, Endpoint Outcomes, Boston, Massachusetts, United States; Dave Andrae, PhD, Endpoint Outcomes, Austin, Texas, United States

**Aims:** The use of patient diaries of participants’ health-related quality of life (HRQoL) is ubiquitous. Frequently, only data from a few time points are used to assess efficacy in clinical trials. This study presents a novel descriptive framework to examine intensive longitudinal diary data, which can be used as a starting point to characterize the natural history of diseases and inform endpoint development. **Methods:** Data were simulated for 20 patients, with 48 days, randomized 1:1 to active or placebo. Ordinal response data were generated for a 7-category item using a longitudinal mixed-effects model, with a positive effect for the treatment group. Differences in response between the groups were generated using novel intra-patient heat maps. Case-wise longitudinal descriptive summaries were calculated for the statistical moments for categorical data for central tendency, dispersion, and asymmetry of the categorical time-series data. Specifically, dominance, Shannon’s entropy index, and the log-skew are utilized. **Results:** Novel intra-participant heat maps displayed differences in the responses for active and placebo. A dominance statistic established the proportion of observations in the modal rating category. Shannon’s entropy index provides insight about the evenness of the ratings split across categories. The log-skew statistic will show information for rare category endorsement. Contrasts were drawn between select active and placebo participants to visualize differences in these descriptive statistics and to lend interpretative value to these summaries. **Conclusions:** Results demonstrate that both the data visualizations and descriptive methods capture nuanced differences in the patterns of responses for placebo and treatment.

### (3168) Clustering of EORTC QLQ-C30 health-related quality of life scales across several disease sites

Abigirl Machingura, MSc, Biostatistics, EORTC, Brussels, Belgium; Jjambe Musoro, PhD: Clinical Epidemiology, Biostatistics and Bioinformatics, EORTC, Brussels, Belgium; Jolie Ringash, MD, Professor, Department of Radiation Oncology, University of Toronto,

Princess Margaret Cancer Centre and the University of Toronto, Toronto, Ontario, Canada; Mekdes Taye, MSc Biostatistics, EORTC, Brussels, Belgium; Madeline Pe, PhD, Quantitative Psychology, Personality and Social Psychology, EORTC, Brussels, Belgium; Corneel Coens, MSc Biostatistics, EORTC, Brussels, Belgium; Francesca Martinelli, EORTC, Brussels, Belgium; Dongsheng Tu, Queen’s University, Kingston, Ontario, Canada; Ethan Basch, Lineberger Comprehensive Cancer Center, North Carolina, United States; Yvonne Brandberg, Karolinska Institutet, Solna, Sweden; Mogens Grønvold, Karolinska Institutet, Solna, Sweden; Mogens Grønvold, University of Copenhagen, Copenhagen, Denmark; Alexander Eggermont, Princess Maxima Center, Utrecht, Netherlands; Fatima Cardoso, Champalimaud Clinical Center, Lisbon, Portugal; Jan Van Meerbeeck, Antwerp University Hospital, Edegem, Belgium; Winette Van der Graaf, Institute of Cancer Research, London, United Kingdom; Martin Taphoorn, Universiteit Leiden/McHaaglanden, Leiden, Netherlands; Jaap Reijnveld, VU University Medical Center and the Academic Medical Centre, Amsterdam, Netherlands; Riccardo Soffiatti, Università di Torino, Torino, Italy; Galina Velikova, Leeds Institute of Medical Research at St James’s, Leeds, United Kingdom; Henning Flechtner, Magdeburg Universitätsklinik, Magdeburg, Germany; Andrew Bottomley, PHD, Psychology, EORTC, Brussels, Belgium; submitted on behalf of The EORTC Quality of Life Group, Brain tumour, Breast Cancer, Melanoma Group, Lung Cancer, Soft Tissue and Bone Sarcoma, Radiation Oncology, Lymphoma, Gastrointestinal Tract Cancer, Head and Neck Cancer, Genito-Urinary Cancers, and Gynecological Cancer Groups

**Aims:** The EORTC QLQ-C30 questionnaire measures 15 health-related quality of life (HRQOL) domains relevant to cancer patients. A previous study by Martinelli et al. (2011) demonstrated that these domains could be collapsed into three main clusters: physical, psychological and gastro-intestinal. Given the current replication crisis in psychology and medical research, it is critical to evaluate the robustness and generalizability of those exploratory findings through replication. This study aims to perform a validation of Martinelli’s findings in an independent data set and evaluate whether these clusters are consistent across disease sites and patient characteristics. **Methods:** Published clinical trial data were obtained from EORTC and Project Data Sphere (public data platform) databases. Only baseline data from trials that assessed HRQOL using the EORTC QLQ-C30 version-3 were retained. Based on the results of Martinelli, pre-defined criteria for successful replication were set: three main clusters should emerge with a minimum R-squared(R<sup>2</sup>) value of 0.51. Following Martinelli’s methodology, cluster analysis was performed on the 15 QLQ-C30 HRQOL-domains. Subgroup analysis for each disease site and selected patient characteristics to examine the sensitivity of the results. **Results:** The pooled dataset consisted of 21 EORTC and 9 non-EORTC datasets, totalling 14,106 patients. Sixteen disease sites were included. Median age was 58 years with 44%  $> 60$  years, 41% had advanced disease and 66% had WHO performance-status 0-1. Overall, three main clusters were identified (R<sup>2</sup> = 0.616). The first cluster (physical) included role, physical and social functioning, fatigue, pain and global quality of life; the second (psychological) included emotional, cognitive functioning and insomnia; finally, the third (gastro-intestinal) included nausea/vomiting and appetite loss. The results were consistent across different levels of disease severity, socio-demographic and clinical characteristics with minor differences in some disease sites. **Conclusion:** This study successfully replicated Martinelli’s findings: the EORTC QLQ-C30 domains can be collapsed into three main clusters: physical, psychological and gastro-intestinal. These results allow us to have a better understanding of how these multidimensional HRQOL domains are related to each other. Knowing this may aid clinicians and cancer patients manage symptom burden as well as policy-makers in defining

social support plans and in choosing HRQOL domains for a clinical trial.

### (3169) Investigating the German response scale of the EORTC QLQ-C30 in a sample of cancer patients and in a norm data reference survey

Michael Koller, University Hospital Regensburg, Regensburg, Germany; Karolina Müller, University Hospital Regensburg, Regensburg, Germany; Sandra Nolte, Charité, Berlin, Germany; Heike Schmidt, Martin Luther University Halle-Wittenberg, Halle, Germany; Christina Harvey, St. Marienwörth Hospital, Bad Kreuznach, Germany; Ulrike Mölle, St. Georg Hospital Leipzig, Bad Kreuznach, Germany; Andreas Boehm, St. Georg Hospital Leipzig, Leipzig, Germany; Daniel Engeler, Cantonal Hospital St. Gallen, St. Gallen, Switzerland; Jürg Metzger, Cantonal Hospital Lucerne, Lucerne, Switzerland; Monika Sztankay, Medical University Innsbruck, Innsbruck, Austria; Bernhard Holzner, Medical University Innsbruck, Innsbruck, Austria; Bernhard Holzner, Medical University Innsbruck, Innsbruck, Austria; Mogens Groenvold, University of Copenhagen, Copenhagen, Denmark; Dagmara Kulis, EORTC QLQ, Brussels, Belgium; Andrew Bottomley, EORTC QLQ, Brussels, Belgium; submitted on behalf of EORTC

**Aims:** The EORTC QLQ-C30 is an internationally established questionnaire for assessing quality of life (QoL) in cancer patients, using four-point Likert scales labeled not at all (1), a little (2), quite a bit (3), and very much (4). The German translation of the response option quite a bit as *mäßig* has shown to violate interval scale assumptions. Previous research has shown that *ziemlich* may be a more suitable translation for quite a bit, located right in between a little and very much. The present studies investigated differences between the *mäßig* and *ziemlich* questionnaire versions. The studies were based on the hypothesis that the *mäßig* version yields higher symptom and lower functioning ratings than the *ziemlich* version, particularly in respondents with higher health burden because they are more likely to choose very much to indicate their symptom burden. **Methods:** The first study enrolled patients with different types of cancer from three German-speaking countries (Germany, Austria, Switzerland). Employing a balanced cross-over design, patients filled in the *mäßig* and *ziemlich* version of the questionnaire within one week. The second study was a representative survey in Germany including 2033 respondents. Half of the participants filled in the *mäßig*, the other half the *ziemlich* version of the questionnaire. The primary endpoint was the summary scale scored from 0 (low functioning) to 100 (high functioning). **Results:** As expected, the summary score was lower in the *mäßig* than in the *ziemlich* version,  $-4.5$  (95% CI  $-7.8$  to  $-1.3$ ),  $p < .006$ . This effect was pronounced in patients with higher health burden,  $-6.8$  (95% CI  $-12.2$  to  $-1.4$ ),  $p < .013$ . This effect was also seen in the survey study,  $-3.1$  (95% CI  $-4.6$  to  $-1.5$ ),  $p < .001$ ; respondents with health burden:  $-4.5$  (95% CI  $-7.3$  to  $-1.7$ ),  $p < .002$ . **Conclusion:** We found subtle but consistent differences between the original translated response format version and an optimized version that better meets the requirements of interval scaling of the EORTC QLQ-C30. The new translation, QLQ-C30 version 3.1 is therefore recommended for future use.

### (3170) Loss of health-related quality of life after an injury: estimation of minimal clinically important difference values

Erik Eriksson, Karolinska Institutet, Stockholm, Sweden; Marie Hasselberg, Professor, Karolinska Institutet, Stockholm, Sweden; Ritva Rissanen, PhD, Karolinska Institutet, Stockholm, Sweden

**Aims:** How patients or populations value any change in their health, or in their life, is important for clinicians and policymakers. By calculating Minimal Clinically Important Difference (MCID) values, it is possible to estimate whether such a change is of importance to the patient or population. MCID values must be estimated for different outcome instruments and different populations and currently no MCID values corresponding to the loss of Health-Related Quality of Life (HRQoL) instruments have previously been estimated for a population that has suffered injuries. Thus, the aim of the current study was to estimate the MCID values corresponding to the loss of HRQoL after an injury. **Methods:** Four distribution-based and four anchor-based methods were used to calculate the MCID values. As anchor, the perceived change question-item was used. **Results:** In a web-based questionnaire, 746 participants, at least 18 years of age, reported that they had an injury during the last 12 months. Participants reported their HRQoL before and after the injury, using the EQ-5D 3L instrument (both EQ Index and Visual Analogue Scale, EQ-VAS), as well as how they perceived the change in the HRQoL after the injury: worse, no change, or better. By comparing the results from different methods, a range of MCID values was obtained: 0.047–0.181 for EQ Index, and 2.23–10.68 for EQ-VAS. A receiver operating characteristic analysis indicated that lower or higher MCID values in the range could be utilized, depending on whether it is important to maximize sensitivity or specificity, respectively. **Conclusion:** These first estimations of MCID values corresponding to loss of HRQoL after an injury indicated an upper limit and a lower limit of an actual MCID value. Further calculations of MCID values for EQ-5D 3L should include an anchor-item that differentiates between additional levels of change, as well as different injury severities and injured body parts.

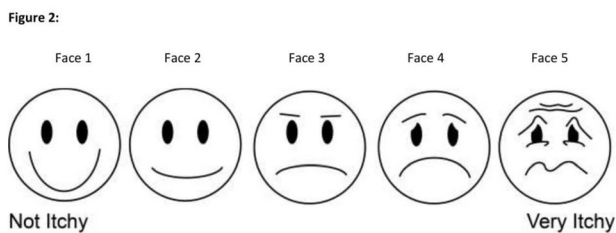
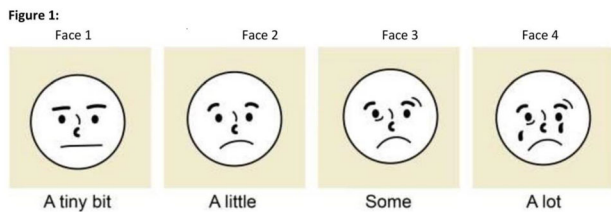
## Qualitative & Mixed Methods Research I

### (3171) Exploring the need for linguistic validation of smiley face response options in pediatric questionnaires

Rachna Kaul, MPA, RWS Life Sciences, East Hartford, Connecticut, United States; Barbara Brandt, MA, RWS Life Sciences, East Hartford, Connecticut, United States; Tim Poepsel, PhD, RWS Life Sciences, Chicago, Illinois, United States; Elizabeth Yohe-Moore, MPH, RWS Life Sciences, Chicago, Illinois, United States; Elizabeth McCullough, MA, RWS Life Sciences, East Hartford, Connecticut, United States; Shawn McKown, MA, RWS Life Sciences, East Hartford, Connecticut, United States

**Aims:** Questionnaires for pediatric populations commonly use smiley faces as response options, due to lack of cognitive sophistication and advanced language skills, to improve accuracy of data. Pediatric patients' understanding of smiley face response options may vary across countries and cultures and responses may be affected by social desirability (1) or cultural and emotional (2) biases. Linguistic Validation (LV) of such questionnaires is necessary to pinpoint these issues and obtain quality data in subsequent clinical trials. **Methods:** We compiled a convenience sample of three recent cognitive debriefing projects on pain and symptom severity questionnaires with smiley faces as response options. Patient feedback on response options was extracted. Overall, debriefing results from 60 language-country pairs were analyzed, each containing 5 or 6 patient interviews ( $n = 241$ ; 6–17 years). **Results:** Our results indicate a lack of distinction between smiley faces depicting conceptually similar response options of "a little" and "some" symptom severity (faces 2 and 3 in Fig. 1) noted across 3 out of 23 languages by both younger (6–11 years) and older (12–17 years) patients. A face representing "very itchy" (face 5 in Fig. 2) was thought to be "scary" and "ugly" by a younger (6–11 year) patient, leading to avoidance of the "very

itchy” answer option as they thought it to be a representation of their own face. They instead chose more desirable “smiling faces” (faces 1 and 2 in Fig. 2) regardless of whether these faces actually represented their symptom severity. Conclusion: Our data suggest that developers should be cautious when designing smiley faces depicting symptom severity. Specifically, RWS Life Sciences recommends ensuring sufficient distinction between smiley face response options to help pediatric patients select an appropriate answer without confusion (see the similarity between faces 2 and 3 in Fig. 1), and without triggering avoidance of faces depicting a strongly negative emotional state. In the overall context of studies on pediatric patients from diverse cultures, using LV to render pictographic response options in culturally acceptable and understandable ways is highly recommended. This allows for collection of quality data free from desirability and culturally specific bias.



### (3172) Illustrations developed as tools for use in pediatric qualitative interviews

Shayna Egan, MPH, Endpoint Outcomes, Long Beach, California, United States; Aisling Ryan, Endpoint Outcomes, Boston, Massachusetts, United States; Christina Graham, Endpoint Outcomes, Boston, Massachusetts, United States; Siobhan McDonold, Endpoint Outcomes, Boston, Massachusetts, United States; Maggie Paulich, Endpoint Outcomes, Long Beach, Massachusetts, United States; Kayla Scippa, Endpoint Outcomes, Long Beach, Massachusetts, United States; Julie Whyte, MPH, Endpoint Outcomes, Boston, Massachusetts, United States; Christopher Evans, PhD, Endpoint Outcomes, Boston, Massachusetts, United States

**Aims:** There is increased demand for patient-reported outcome (PRO) development in illnesses that occur in children. Qualitative interviewing—a key portion of the PRO development process—is often difficult to conduct with pediatric participants. Since research has shown that drawings can help children talk about their health, Endpoint Outcomes developed illustrations to assist children in describing their health status and well-being. **Methods:** The Illustrated Kid Cards are novel cartoon depictions of common symptoms and impacts associated with pediatric illnesses; these illustrations were developed based on concepts in the Pediatric Quality of Life Inventory v4.0. They are intended to be both gender- and race-neutral. After ethics approval was obtained, healthy children ages 5–12 were recruited for in-person interviews. Participants were presented with 40 illustrations and asked to describe the meaning of each illustration, report on its relevance and clarity, and provide input on changes to the illustration. If participants did not understand the illustration, interviewers

described what the image’s intended meaning was and subsequently collected clarity information. Participant data were analyzed to determine whether the illustrations conveyed the intended concepts. **Results:** Sixteen children aged 5–11 (average = 8.5, standard deviation = 1.8) from the Boston area participated in interviews. The majority of illustrations ( $n = 29$  of 40, 72.5%) were interpreted as intended by  $\geq 70\%$  of participants providing data. Of the illustrations that were misinterpreted, some ( $n = 4$  of 11, 36.4%) were reported to be clear by  $> 70.0\%$  of participants after interviewers described the illustrations’ intended meanings. The remaining illustrations—tired, headache, joint pain, difficulty in school, feeling bad about appearance, difficulty walking, and worried—were found to be unclear by  $> 30\%$  of participants providing data, even after the interviewer described the intended meaning of the illustration. **Conclusion:** The illustrations were generally well understood by participants. Revisions are planned for illustrations found to be unclear by  $> 30\%$  of participants. Future research objectives include interviews with children in various disease areas for illustrations with a disease-specific context. The Illustrated Kid Cards are available in the public domain for use by all researchers.

Figure 1. Example illustration



### (3174) From item to interface: updating the EORTC Item Library using results from a qualitative content analysis

Claire Piccinin, MSc, EORTC, Brussels, Belgium; Dagmara Kuliś, MA, EORTC, Brussels, Belgium; Andrew Bottomley, PhD, EORTC, Brussels, Belgium; Francesca Martinelli, MSc, EORTC, Brussels, Belgium; Madeline Pe, PhD, EORTC, Brussels, Belgium; Mogens Groenvold, MD, PhD, DSci, University of Copenhagen, Copenhagen, Denmark

**Aims:** The EORTC Item Library is a dynamic platform that allows users to search for relevant PRO questionnaires and items while providing interactive features to create customized item lists for assessment in cancer patients. However, there is a need to optimize searchability by allowing users to more easily access relevant content for their PRO assessment needs. As part of a previous qualitative content analysis, we identified 43 main codes to classify the Item Library’s 952 unique items. The work presented here aimed to further refine the coding and classification by creating more specific second-level descriptive codes and implementing them as a search function within the Item Library, allowing users to search based on these hierarchical codes capturing a range of symptoms, functioning, and satisfaction with care. **Methods:** All items were assessed with their main codes, and secondary codes were inductively assigned based on the content of each item and main code, where relevant. Secondary

codes were reviewed by a second rater, with any discrepancies discussed until a consensus was reached. Main and secondary code mapping results were then summarized and evaluated by the authors, who mapped out and tested the best approach for integration into the online platform. Results: Of the 43 main codes, 27 qualified for further specification with secondary codes, representing broader domains that could be divided into subcategories. In total, 109 secondary codes were created (1 to 10 per main code). General physical health and physical functioning main codes received the highest number of secondary codes. Following evaluation from the study team, main and secondary code mapping results were used to create a tree view classification in the Item Library, allowing users to search for items based on specific categories comprised of collapsible main and secondary codes, under which linked items are listed. Conclusion: The newly implemented online classification provides a new search function for selecting items within the Item Library using a hierarchical system. These results highlight the feasibility of integrating findings from other methods of classification, which will facilitate the future use of alternative frameworks (e.g., CTCAE, WHO-ICF) to increase the accessibility of content and promote comparison and collaboration.

**(3175) Integration of emerging good practices for Translatability Assessment (TA) of patient-reported outcome (PRO) measures: a discussion of the benefits to PRO translations**

Barbara Brandt, RWS Life Sciences, East Hartford, CT, Connecticut, United States; Tim Poepsel, PhD, RWS Life Sciences, Chicago, Illinois, United States; Rachna Kaul, MPA, RWS Life Sciences, East Hartford, Connecticut, United States; Elizabeth Yohe Moore, MPH, RWS Life Sciences, Chicago, Illinois, United States; Elizabeth McCullough, MA, RWS Life Sciences, East Hartford, Connecticut, United States; Shawn McKown, MA, RWS Life Sciences, East Hartford, Connecticut, United States

Aims: A Translatability Assessment (TA) is conducted by linguists to identify concepts, phrases, or components of an instrument that are difficult to translate or require localization. In 2018, the ISOQOL Translation and Cultural Adaptation Special Interest Group (TCA-SIG) published emerging TA good practices, with the goal of guiding future use in PRO development (Acquadro, et al., 2018). RWS Life Sciences integrated these recommendations in 2019 with its existing methodology, to achieve greater specificity in reporting severity of translation difficulty and improved guidance to questionnaire developers for source text revision. This research compared TA outputs before and after integration of TCA-SIG guidance. Methods: A convenience sample of three recent TA projects employing the ISOQOL SIG good practices and three TA projects using the prior methodology was compiled. Linguist feedback per language and per project, representing 42 unique languages, was compared and the degree of severity of potential translation issues and resulting actions were analyzed. Results: TA projects where ISOQOL SIG good practices were integrated rated source text issues on a scale of translation difficulty, ranging from Level 1 (No Difficulty) to Level 4 (Extreme Difficulty). Recommended actions included no change to original wording, no change to original wording but provision of conceptually equivalent alternative wording to address translation issues, a change to original wording to address issues threatening conceptual understanding in the target language, or omission of wording because of extreme translation difficulty. On average, PROs using the previous TA methodology yielded 4 total translation and source text revisions per instrument, compared to 10 revisions per instrument with the new ISOQOL SIG methodology. The new methodology also increased developer involvement in discussions regarding intent of each item and translation revisions. Conclusion:

Integration of ISOQOL TCA-SIG emerging good practices for TA of PROs has resulted in greater specificity and consistency in the guidance provided to questionnaire developers when rating severity of translatability problems, yielding an increased number of revisions made to the translations and source text. This enhanced methodology supports improved data pooling across translated PROs within patient trials and a more efficient translation process.

**(3176) Examination of a modified approach to the Delphi Technique in health outcomes research**

Nathan Johnson, Endpoint Outcomes, Long Beach, California, United States; Ted Wells, Endpoint Outcomes, Boston, Massachusetts, United States

Aims: The Nominal Group Technique (NGT) and Delphi Technique (DT) are two consensus methods that are used to inform the development of guidelines, standards, or assessments in areas where research-based evidence is absent or inconclusive. While NGT and DT are increasing in outcomes research, methodological standards remain lacking. A modified version of the DT that includes key components of the NGT has been developed and implemented in a number of health outcomes studies. This study presents modified DT, examines its strengths and weaknesses, and discusses potential future implications. Methods: Existing literature was reviewed to identify key considerations and best practice methods of DT/NGT in healthcare research. Our study team tested this modified DT in six studies to gain consensus from clinical experts on: item selection for screening questionnaires; new treatment paradigms; treatment expectations and ideal treatment responders; and adverse event management. Areas of study included ophthalmology, gynecology, neurology, and dermatology. Results: Key components were identified from the existing literature in regards to the development and methodology of a practical but well-grounded modified DT: size and composition of panels, the number of Delphi rounds, and analysis of Delphi panel results. From this comparative analysis, design features that were found to be effective in our studies included: methods to generate round 1 questions (literature review and expert review); the number of Delphi panelists (8-10); the number of Delphi rounds (2-3); consensus thresholds (70–80%); the use of technology during rounds of voting (web-based survey, voter response systems); and the incorporation of in-person panels or videoconferencing during the final round of voting. Conclusion: The modified DT in the six studies was found to be pragmatic and methodologically robust. Benefits included the ability to complete the process within 3-6 months, and the use of electronic voting and videoconferencing to minimize in-person interactions and travel. Limitations included a loss of meaningful consensus if the number of voting rounds were cutoff prematurely. Since the DT is mentioned in the patient-focused FDA guidance as a potential avenue for concept elicitation, our methodology could be employed to help gain consensus on what is important to patients for drug development.

**(3177) Examining the Content Validity of Generic Preference-Based Measures in Chronic Obstructive Pulmonary Disease**

Ava Mehdipour, BSc, McMaster University, Hamilton, Ontario, Canada; Sachi O'Hoski, PT, McMaster University, Hamilton, Ontario, Canada; Marla Beauchamp, PT, PhD, McMaster University, Hamilton, Ontario, Canada; Joshua Wald, MD, St. Joseph's Healthcare, Hamilton, Ontario, Canada; Ayse Kuspinar, PT, PhD, McMaster University, Hamilton, Ontario, Canada

Aims: Generic preference-based measures (GPBMs) are health-related quality of life (HRQL) measures that can be used to assess the

cost-utility of different interventions. However, their content validity has not yet been evaluated in individuals with chronic obstructive pulmonary disease (COPD). The primary objective of this study was to assess the content validity of GPBMs in COPD. The secondary objective was to examine the convergent validity of an individualized HRQL measure (the Patient-Generated Index (PGI)) against a GPBM (the Six-Dimensional Short Form Survey (SF-6D)) in COPD. Methods: The PGI and the RAND-36 (to derive SF-6D scores) were administered to adults with a physician-diagnosis of COPD. The PGI allowed participants to nominate up to five areas of their life affected by their COPD. These areas were coded independently by two researchers using the International Classification of Functioning, Disability and Health and mapped onto well-recognized GPBMs. If normality assumptions were met, a Pearson's correlation coefficient of at least 0.50 was hypothesized between the PGI and SF-6D. Results: The sample ( $n = 60$ ) had a mean age of 69 (Standard Deviation 8) years, an equal distribution of males and females, and a mean FEV1 of 43% predicted. Thirteen HRQL domains important to people with COPD were identified using the PGI, with the top three being 'mobility (26%)', 'recreation and leisure' (25%) and 'domestic life' (19%). The Quality of Well-Being Scale covered the highest number of domains important to individuals with COPD (85%). The remaining GPBMs covered between 15 and 62% of domains. A moderate positive correlation ( $r = 0.32$ ) was found between the PGI and the SF-6D. Conclusion: To our knowledge, this was the first study to evaluate the content validity of GPBMs in COPD. While content coverage varied between the measures, most well-recognized GPBMs did not cover an adequate amount of life areas important to those living COPD. This may affect cost-utility decisions in healthcare as HRQL scores provided by these measures may inadequately reflect the HRQL of individuals living with COPD.

### (3178) Evaluating intercoder reliability in qualitative research analysis

Brooke Witherspoon, Endpoint Outcomes, Boston, Massachusetts, United States; Shayna Egan, MPH, Endpoint Outcomes, Long Beach, California, United States; Martha Gauthier, MA, Endpoint Outcomes, Boston, Massachusetts, United States

Aims: Evaluating consistency in coding across multiple coders is an important component of qualitative research quality assurance, without which the validity of data and its interpretation cannot reliably be established. This study examines intercoder reliability (ICR) methods and results across qualitative research studies involving multiple coders to illuminate methods and explore potential contributing factors for achieving acceptable levels of ICR. Methods: ICR results from qualitative interview studies that aimed to spontaneously elicit concepts (e.g., symptoms, impacts) were reviewed. Coding of the interviews was based on established qualitative research methods, including grounded theory and constant comparison method. The extent to which 3-4 independent coders were concordant in coding was evaluated using percent agreement for each study. Circumstances across studies were evaluated to explore if they potentially contributed to the coding team's ability to reach at least 90% agreement on the first transcript coded. Results: Across 13 studies,  $\geq 90\%$  agreement was met on the first transcript for 10 studies. The remaining studies required an additional 1-3 transcripts to be coded to reach  $\geq 90\%$  agreement. For studies for which ICR was achieved after coding one transcript, coders averaged approximately 5 months more experience than studies requiring more than one transcript to reach ICR (Table 1). Among the three studies where multiple transcripts were needed to achieve ICR, two studies were among the least experienced coder group. In general, the percentage of studies that achieved ICR after coding one transcript decreased as

the number of codes increased: 100% of studies with  $\leq 50$  codes achieved ICR on the first transcript whereas 0% of the studies with 151–200 codes achieved ICR on the first transcript (Table 2). A greater percentage of studies with homogeneous disease characteristics achieved ICR after coding one transcript (80%) than studies with heterogeneous disease characteristics (67%). The number of coders did not make a notable difference in whether ICR was achieved after coding one transcript. Conclusion: Achieving acceptable levels of ICR after coding the first transcript was more common among more experienced coding teams, studies with shorter code lists, and studies investigating diseases with more homogeneous disease characteristics.

Table 1. ICR achievement by average coder experience (months)

	Minimum	Maximum	Average
All coders (n=43)	1.1	66.2	29.6
Coders achieving ICR after 1 transcript (n=33)	5.5	64.8	30.8
Coders requiring > 1 transcript to achieve ICR (n=10)	1.1	66.2	25.7

Table 2. ICR achievement by number of codes

Number of codes	Number of studies N	Coders achieving ICR after 1 transcript n (%)	Coders requiring > 1 transcript to achieve ICR n (%)
1-50	5	5 (100.0%)	0 (0.0%)
51-100	4	3 (75.0%)	1 (25.0%)
101-150	2	1 (50.0%)	1 (50.0%)
151-200	1	0 (0.0%)	1 (100.0%)
201-250	1	1 (100.0%)	0 (0.0%)

### (3179) Subjective experiences in telemedicine and the call for setting-related quality of life: A qualitative study in telemedical professionals as well as patients with mental disorders and chronic conditions

Klara Greffin, Dipl.-Psych., University of Greifswald, Greifswald, Germany; Silke Schmidt, Prof. Dr., University of Greifswald, Greifswald, Germany; Neeltje van den Berg, PD Dr., University Medicine Greifswald, Greifswald, Germany; Wolfagng Hoffmann, Prof. Dr., University Medicine Greifswald, Greifswald, Germany; Oliver Ritter, Prof. Dr., Brandenburg City Hospital, Medical University Brandenburg, Brandenburg an der Havel, Germany; Michael Oeff, Prof. Dr., Brandenburg City Hospital, Brandenburg an der Havel, Germany; Georg Schomerus, Prof. Dr., University Medicine Leipzig, Leipzig, Germany; Holger Muehlan, Dr., University of Greifswald, Greifswald, Germany

Aims: Telemedicine (TM) is applied to improve health care management of patients with, e.g., mental disorders or chronic conditions. Yet, reviews show inconsistent results with regard to the effectiveness of TM on patient-reported outcomes (PRO) like quality of life (QoL). We assume that PRO measures may lack sensitivity to assess the intended results of TM applications. Our study aimed to explore the experiences of TM and its impact on QoL from complimentary perspectives of patients with major depression or heart failure as well as from TM professionals. Methods: Overall, 63 semi-structured single interviews and 15 focus groups ( $n = 68$  participants) have been conducted between July 2018 and February 2019. Participants were patients with heart failure or major depression with or without TM supported health care management as well as TM professionals. Mayring's content analysis approach was used to encode the qualitative data material using MAXQDA software. Results: Patients and professionals highlighted advantages of TM as compared to

conventional care such as enlarged accessibility, fast availability of medical advice or continuity of care. Our category system comprised four additional domains with 35 attributes related to QoL issues in TM contexts, that are specific for this type of health care provision and not covered by already established approaches, e.g., patient safety, health care-related empowerment and needs-based health care from a patient's or health care professional's perspective. TM can be an improvement in managing a disease and has a positive impact on QoL when it is tailored to the patients' needs. Conclusion: Interviewing patients and health care professionals brought forth specific aspects of QoL evolving in TM contexts. These results reinforce the assumption that existing QoL measurements lack sensitivity to assess the intended results of TM applications. We will address this deficiency by a TM-related re-conceptualization of the assessment of QoL and the development of a suitable instrument based on the resulting category system of this study. The project is aimed at improving PRO's for TM care as well as optimizing patient-orientation in this innovative and rapidly emerging type of health care provision. Furthermore, it contributes to increase patient participation in health services research.

### **(3180) Towards adjunct setting-related quality of life assessment in telemedical care: concept elicitation, item generation and initial testing**

Klara Greffin, Dipl.-Psych., University of Greifswald, Greifswald, Germany; Silke Schmidt, Prof. Dr., University of Greifswald, Greifswald, Germany; Neeltje van den Berg, PD Dr., University Medicine Greifswald, Greifswald, Germany; Wolfagn Hoffmann, Prof. Dr., University Medicine Greifswald, Greifswald, Germany; Oliver Ritter, Prof. Dr., Brandenburg City Hospital, Medical University Brandenburg, Brandenburg an der Havel, Germany; Michael Oeff, Prof. Dr., Brandenburg City Hospital, Brandenburg an der Havel, Germany; Georg Schomerus, Prof. Dr., University Medicine Leipzig, Leipzig, Germany; Holger Muehlan, Dr., University of Greifswald, Greifswald, Germany

Aims: Reviews about the impact of telemedical care (TM) on quality of life (QoL) show inconsistent results. Among other reasons, existing instruments may not be sensitive enough to assess setting-specific aspects of QoL in TM contexts. Therefore, we aimed to explore and conceptualize QoL in TM as well as to develop and test a content-valid "add-on" assessment to measure the specific aspects of QoL in TM contexts, which are not sufficiently covered by established instruments yet. Methods: For concept elicitation, we underwent a review of criteria and instruments for patient-reported outcomes used in TM studies ( $n = 482$  studies included). Second, we conducted interviews ( $n = 63$ ) and focus groups ( $n = 68$ ) with chronic physically or mentally ill patients and TM professionals. This material was analyzed using a content analytic approach. The resulting category system served as basis for item generation. We used cognitive debriefings to test how relevant, plausible and comprehensible the items were for the patients ( $n = 32$ ). Additionally, an online survey among TM professionals was conducted ( $n = 15$ ) to assess the relevance, applicability and scope of the item pool. Finally, the initial questionnaire was applied to patients with depression or heart failure, with or without TM care ( $n = 200$ ), to explore dimensionality of the item pool and analyze the psychometric performance on item and scale level. Results: The category system comprised four additional domains with 35 attributes related to specific QoL issues in TM contexts. The initial item pool with 227 items, derived from the qualitative material, was further refined by cognitive debriefings, excluding 122 items. In the expert survey, 105 items of the provisional instrument were rated and an average of about 20 items assessed to be an optimal scope. Initial psychometric analysis of the

pilot study data confirmed the multidimensional structure of the item pool with a hierarchical model including a dominant higher-order factor. Conclusion: Our results confirm that QoL assessment in TM-contexts should be complemented by a setting-related domain, covering needs-based health care, health care-related empowerment, and patient-experienced safety. Therefore, an instrument is needed to provide a more sensitive detection of the intended effects of TM on QoL.

### **3181 Development of a new patient-reported outcome measure for use in patient's diagnosis with headaches**

Joanna O'Gorman, Brigham and Women's Hospital, Boston, Massachusetts, United States; Elena Tsangaris, PhD, Brigham and Women's Hospital, Boston, Massachusetts, United States; Lisa Gfrerer, Md, PhD, Massachusetts General Hospital, Boston, Massachusetts, United States; Anne Klassen, DPhil, McMaster University, Hamilton, Ontario, Canada; Andrea Pusic, MD, Brigham and Women's Hospital, Boston, Massachusetts, United States

Aims: Headaches impose a tremendous burden on one's quality of life, including their physical and cognitive function. Traditionally, treatment for headaches has involved preventative and abortive medications, opioids, injectables, and medical devices. More recently, surgical techniques have evolved to decompress the peripheral sensory nerves around the skull to improve the frequency, severity, and duration of headaches. Trigger site deactivation surgery for headaches is available for patients who have failed conservative treatment and who meet other anatomic and clinical characteristics. To understand the effectiveness of headache surgery and its impact on one's quality of life, there is a need for a new patient-reported outcome measure (PROMs). Currently available PROMs are not tailored towards headache patients undergoing surgery. Methods: Participants are being recruited from Massachusetts General Hospital and Beth Israel Deaconess Medical Center in Boston, MA. Semi-structured qualitative interviews are ongoing with a heterogeneous group of patients who are over the age of 18 years, have a confirmed diagnosis of migraine, headaches, or occipital neuralgia, and who have or have not had surgery for treatment of their pain. Interviews are audio-recorded, transcribed, and coded using the constant comparison approach. The Interpretive description qualitative approach is being used for data analysis, aimed at identifying key themes and subthemes to develop an item pool for use in scale development. Results: Twelve participant interviews have been conducted. Among these patients, most were Caucasian ( $n = 11$ ), working full time ( $n = 7$ ), and had surgery ( $n = 11$ ). These patients provided codes that covered a range of headache-specific concerns including, social function (e.g., support from family and friends), symptoms (e.g., throbbing, numb), physical function (e.g., bathing, driving, child care), and experience of care (e.g., information, establishing/obtaining care). Conclusion: These preliminary findings will be used to inform scale development for a new PROM. We plan to conduct additional patient interviews and involve experts to develop and refine the PROM, which will then be field-tested internationally.

### **(3183) "How do I ask that?" Recommendations for using patient-appropriate language during qualitative interviews to develop fit-for-purpose clinical outcome assessments (COAs)**

Hannah C Pegram, BSc, DRG (part of Clarivate), Manchester, United Kingdom; Sarah Knight, MSc, DRG (part of Clarivate), Bicester, United Kingdom; Chris Marshall, MSc, DRG (part of Clarivate), Manchester, United Kingdom; Jake Macey, MSc, DRG (part of

Clarivate), Bicester, United Kingdom; Natalie V J Aldhouse, MSc, DRG (part of Clarivate), Manchester, United Kingdom; Chloe Johnson, MSc, DRG (part of Clarivate), Manchester, United Kingdom; Tamara Al-Zubeidi, BSc, DRG (part of Clarivate), Bicester, United Kingdom; Maile Hunter, MBA, DRG (part of Clarivate), Nashville, Tennessee, United States; Helen Kitchen, MSc, DRG (part of Clarivate), Manchester, United Kingdom

**Aims:** To identify key considerations when discussing complex conditions and concepts in qualitative patient interviews for development of clinical outcome assessments (COA). **Methods:** A retrospective review of research learnings from qualitative interview studies was conducted. Researchers extracted examples of how complex questions, concepts or subject matters had been approached in the interview guide, patient facing materials and during interviews. Following extraction, recommendations for the development of interview materials and interviewer guidance were collated. **Results:** Several successful approaches to complex conditions/concepts were identified in studies across several therapy areas. To help patients feel at ease and empowered to talk about their experiences in their own way both before and during the interview, it is recommended that researchers identify terminology patients use through qualitative literature, social media reviews and/or consulting subject matter experts. This terminology should be used throughout the interview guide and patient-facing materials, especially for sensitive topics and in place of complicated terminology (e.g., medical terms). Additionally, use of indirect and funnel questioning (beginning broadly and progressively narrowing the discussion to the specific concept of interest; e.g., general symptom experience > flare experience > impact of flares) can help patients understand and articulate complex topics. Piloting materials to test patient understanding allows methods/questioning to be refined prior to interviews to improve data quality. Considerations/adaptations for specific patient groups (e.g., children, the cognitively impaired) is advised, e.g., simplification of high-level terms and constructs (e.g., using ‘memory problems’ in addition to ‘Alzheimer’s disease’). Patient understanding can also be aided by creative tasks or illustrative examples including diagrams and visual aids (e.g., use of directional arrows when exploring meaningful change in COA scores). Discussion of each patient’s preferred terminology during the interview is recommended to aid consistent patient-researcher understanding of the focus of the interview and complex topics. **Conclusion:** By implementing patient-centric approaches when designing interview materials, COA researchers can communicate efficiently with patients and aid patient interpretation and understanding during interviews exploring complex conditions and concepts. Ultimately, this will generate stronger evidence for COA selection/development to support endpoints in clinical trials.

## Qualitative & Mixed Methods Research II

### (3184) Understanding the Motivational Factors Behind Patient Engagement

Tamara McCarron, University of Calgary, Calgary, Alberta, Canada; Tamara L. McCarron, PhD, University of Calgary, Calgary, Alberta, Canada; Thomas Noseworthy, MD, University of Calgary, Calgary, Alberta, Canada; Karen Moffat, Patient Co-investigator, Calgary, Alberta, Canada; Gloria Wilkinson, None, Patient Co-investigator, Calgary, Alberta, Canada; Sandra Zelinsky, None, Patient Co-investigator, Calgary, Alberta, Canada; Deborah White, PhD, University of Calgary, Calgary, Alberta, Canada; Derek Hassay, PhD, University of Calgary, Calgary, Alberta, Canada; Diane Lorezzenti,

PhD, University of Calgary, Calgary, Alberta, Canada; Nancy Marlett, PhD, University of Calgary, Calgary, Alberta, Canada; PhD, University of Calgary, Calgary, Alberta, Canada

**Aims:** As the role of patients as partners in health research continues to evolve, understanding the motivations of why these individuals engage with health systems is an important factor to the success of engagement initiatives. This study reports on how three patient co-investigators and a researcher co-designed a study to understand the motivational factors of individuals who engage as partners in health care. **Methods:** Key informant interviews with patient and family members were conducted, and results were themed using a constant comparative approach to inform the development of the survey tool. The survey was administered to patients and family members who are actively involved in engagement activities in Alberta. Survey data were analyzed using descriptive statistics and exploratory factor analysis. Cronbach’s alpha determined reliability of the identified motivations. **Results:** One thousand, four hundred and forty-nine individuals participated in the provincial survey. All returned surveys were analyzed. The majority of participants were female, retired, well-educated, and lived in an urban centre. After factor analysis, seven motivations were identified. Analysis of internal consistency revealed acceptable reliability for the 7 motivations. These motivations were named by considering the variables within the resulting dimensions. The identified motivations were named as follows: Self-fulfillment, Improving Healthcare, Compensation, Influence, Learning New Things, Conditional and Perks. **Conclusion:** The results of this research describe a sample of patient and family members currently involved in various roles such as patient and family advisors with health care organizations. We identified seven motivational factors underlying their engagement. A deeper knowledge of these motivations will not only create meaningful engagement opportunities for patients, but also enable health organizations to gain from the experience of these individuals, thereby enhancing the quality and sustainability of patient engagement programs.

### (3185) The validity of proxy responses on patient-reported outcome measures: Are proxies a reliable alternative to stroke patients’ self-report ?

Brittany Lapin, PhD MPH, Cleveland Clinic, Cleveland, Ohio, United States; Nicolas Thompson, MS, Cleveland Clinic, Cleveland, Ohio, United States; Andrew Schuster, BA, Cleveland Clinic, Cleveland, Ohio, United States; Irene Katzan, MD, Cleveland Clinic, Cleveland, Ohio, United States

**Aims:** As many as 25% of stroke patients may be unable to report their health status due to language or cognitive impairments, with caregivers, or proxies, completing patient-reported outcome measures (PROMs) on the patient’s behalf. Research indicates that proxies overestimate patients’ symptoms, however it is unclear if proxy responses may be more reliable than patient responses, especially in the acute setting. The objective of our study was to assess the validity, reliability, and responsiveness of proxy responses as compared to patient self-responses across multiple domains of health. **Methods:** Stroke patients and their proxies were recruited to complete PROMs between 7/2018-11/2019. PROMs included PHQ-9, Neuro-QoL cognitive function, PROMIS physical function, social role satisfaction, anxiety, fatigue, pain interference, sleep disturbance, and PROMIS Global Health. Internal reliability and convergent validity were compared between patient- and proxy-reported measures. Known-groups validity was assessed across increasing stroke disability. Within a subset of patients who

attended rehabilitation following acute stroke, internal responsiveness was evaluated using paired t-tests, and external responsiveness assessed the ability of the PROM to detect clinically important change, based on clinician-reported stroke disability, through linear regression models of the change in PROM. Results: This cross-sectional study included 200 stroke patients (age  $62.2 \pm 13.3$ , 41.5% female) and their proxies (age  $56.5 \pm 13.9$ , 70% female, 72% spouses). Internal reliability was  $> 0.90$  for the majority of PROMs and all PROMs were significantly correlated in both patients and proxies. Self- and proxy-reported measures worsened significantly with increasing stroke disability. For 34 (17%) patients who attended rehabilitation, significant improvement over 30 days was self-reported for 5 domains whereas proxies reported no improvement. Eighteen (53%) patients had clinician-reported improvement, yet it was not predictive of either self- or proxy-reported change across any domains. Conclusion: Both self- and proxy-reported domains of health demonstrated high internal reliability, convergent validity, and known-groups validity. Only self-reported PROMs were responsive to change, however neither self- nor proxy-reported change were associated with clinician-reported improvement. These findings justify the utilization of proxy responses in stroke patients at a single time point but further research is necessary to determine the ability of proxies to identify change.

#### **(3186) Cultural adaptation and content validation of the Palliative Care Quiz for Nursing (PCQN) for Brazilian nurses**

Aline de Mello, RN, Master's Degree Student, University of Guarulhos, Guarulhos, Brazil; Noélie Freitas, RN, PhD, University of Guarulhos, Guarulhos, Brazil

**Aims:** To adapt the Palliative Care Quiz for Nursing (PCQN) into Brazilian Portuguese language and to perform content validation of the adapted version. **Methods:** Methodological study. PCQN is an instrument used to assess the knowledge of nurses about palliative care and it contains 20 items evaluated by means of a scale of answer options: correct, incorrect or do not know. The steps of cultural adaptation of the PCQN were carried out according to the recommendations of the literature. Data collection at the validation stage was performed at a hospital located in the city of Guarulhos, State of São Paulo, Brazil. Content validity were assessed in the expert committee step. All participants signed the Informed Consent Form. The content validity was assessed by means of content validity index (CVI) and the Content Validity Coefficient per item (CVCi). The results from the pre-test step with nurses were analyzed by means of descriptive statistics PCQN adapted version. Data were analyzed using the SPSS® Program version 21.0. **Results:** The PCQN was submitted to the process of translation and back-translation, changes were made during the process. Five individuals participated in the expert committee, who compared the Consensual Portuguese Version 1 of the PCQN with the original version. It was found a CVI value for the total score of 0.87. Most items had a Content Validity Coefficient per item (CVCi) above 0.80. However, items 12, 13 and 17 presented CVCi below 0.70, respectively, 0.70, 0.70 and 0.60. 30 nurses participated in the validation step, 53 (72.6%) were women, average age 40.3 years (Standard deviation [SD] = 7.44) and 20 (66.7%) had no experience in palliative care. The nurses had no difficulties in understanding the adapted version of the PCQN. Most nurses selected wrong answers from the PCQN (10 of the 20 items). **Conclusion:** The adapted version of the PCQN for Brazil had its content validity confirmed. The nurses who participated in the pre-test did not have any difficulties in understanding the adapted version of the PCQN, however, most of them missed the answers. Future study will carry out the validation of the adapted version of the PCQN.

#### **(3187) A standard set of value-based patient-centered outcomes for hepatic carcinoma: an international Delphi survey**

Zineb Cherkaoui, Hôpitaux Universitaires de Strasbourg, Strasbourg, France; Emanuel Felli, Hôpitaux Universitaires de Strasbourg, Strasbourg, France; Elodie Léost, Integrated Health Solutions (IHS), Medtronic France, Boulogne-Billancourt, 92100, France, Strasbourg, France; Bastien Delattre, Integrated Health Solutions (IHS), Medtronic France, Boulogne-Billancourt, 92100, France, Strasbourg, France; Gregoire Gessain, Chair of Innovation & Value in Health, University of Paris School of Medicine, Versailles, France; Véronique Christian, Hôpitaux Universitaires de Strasbourg, Versailles, France; Stephanie Mathis, Hôpitaux Universitaires de Strasbourg, Strasbourg, France; Christiane Bruck, Hôpitaux Universitaires de Strasbourg, Strasbourg, France; Elisa Martinis, Hôpitaux Universitaires de Strasbourg, Strasbourg, France; Elena Bangoura, Hôpitaux Universitaires de Strasbourg, Strasbourg, France; Sophie Blanes, Hôpitaux Universitaires de Strasbourg, Strasbourg, France; Sophie Blanes, Hôpitaux Universitaires de Strasbourg, Strasbourg, France; Benoit Gallix, MD, PhD, Institut Hospitalo-Universitaire de Strasbourg, Strasbourg, France; Didier Mutter, MD, PhD, Hôpitaux Universitaires de Strasbourg, Strasbourg, France; Constance Lugiez, Integrated Health Solutions (IHS), Medtronic France, Boulogne-Billancourt, 92100, France, Strasbourg, France; Patrick Pessaux, MD, PhD, Hôpitaux Universitaires de Strasbourg, Strasbourg, France

**Aims:** Due to the rise of costs and the limitations of the resources, the global health care system is looking for new models based on value (value-based health care—VBHC) to improve patient outcomes. To define the value, it is necessary to identify standardized outcome measurements that allow the optimal monitoring and comparison of the quality of care. These outcomes include medical results (CROMs, Clinical Reported Outcomes Measurements) and quality of life components that matter most to patients (PROMs, Patient-Reported Outcomes Measurements), particularly important in case of short life expectancy. The aim of this work was to create standardized metrics that could be used for outcomes analysis of primary or secondary liver cancer. **Methods:** A multidisciplinary working group (WG) was assembled. A systematic review was conducted to collect the most common outcomes in liver cancer clinical studies. A total of 377 clinical studies were reviewed, 1539 outcomes were listed, including CROMs and PROMs. After workshops the WG reached a consensus on the definition of the main outcomes for patients with hepatic cancer, identified existing questionnaires which could be used for PROMs collection and set the timeline for data collection. To refine and validate the final outcomes set, an international external committee completed a Delphi process (two rounds both for CROMs and PROMs). **Results:** After the systematic literature review, the WG selected 141 outcomes (121 CROMs/20 PROMs) to submit to the international Delphi vote committee. In total 96 variables (86 CROMs/10 PROMs) were retained. PROMs were added to monitor quality of life, functioning, and symptoms, through the EORTC QLQ-C30 and EORTC QLQ-HCC18 questionnaires. **Conclusion:** A standardized set of outcome measures that need to be validated through an international health outcome comparisons and quality-of-care assessments was built.

#### **(3188) Translation and cross-cultural adaptation of ten pediatric PROMIS® item banks into Swedish**

John Chaplin, PhD, Sahlgrenska Academy at University of Gothenburg, Gothenburg, Sweden; Evalill Nilsson, PhD, Linneaus University, Kalmar, Sweden; Ida Blomqvist, Psychologist, Umeå University, Umeå, Sweden; Martin Stackelberg, Psychologist, Örebro



University, Örebro, Sweden; Mats Ericsson, PhD., Professor, Örebro University, Örebro, Sweden; Eva Henje, PhD., Assoc. Professor, Umeå University, Örebro, Sweden; Rikard Wicksell, Psychologist, PhD, Karolinska Institute, Stockholm, Sweden; Karin Blomberg, PhD, Örebro University, Örebro, Sweden; Inga Dennhag, Psychologist, PhD, Umeå University, Umeå, Sweden

**Aims:** We describe the first large-scale translation and cross-cultural adaptation into Swedish of ten pediatric PROMIS item banks: Anger (9-items), Anxiety (15-items), Depressive symptoms (14-items), Cognitive function (43-items), Family relationships (47-items), Fatigue (25-items), Pain interference (20-items), Peer relationships (15-items), Physical activity (10-items), Positive affect (38-items) and Profile-25. **Methods:** ISPOR recommendations were followed. The review was carried out in multi-professional small groups of twelve health-related quality-of-life researchers, pediatric clinicians and linguists from different geographical regions in a two-day workshop. Cross-cultural comparisons were made to identify problems and to produce a consensus-derived version, which was then back translated, evaluated, and revised where necessary. Prior translations and review of version 1 of the item-banks were utilized (4 item-banks). The reviewed item-banks were presented in four sessions of cognitive debriefings over 2 years with twenty-two young people (8-18 years) in three dialectically different regions of Sweden before finalization. **Results:** One hundred and seventy-five items were translated. All item banks had translation issues to be resolved. Sixty-six items (38%) needed resolution at the cognitive debriefing stage, the majority of issues were in the area of unclear definitions in the English items (35 items), followed by language and cultural differences (16 issues) and age appropriate language (14 items). The cultural issues identified were 1) identifying suitable word alternative to match the English where Swedish lacked the volume of words to choose from; 2) adjectival agreement on intensity levels of the concept to be translated; 3) culturally specific idiomatic phrases; 4) use of linguistically specific homonyms in English that did not match Swedish word usage; 5) cultural differences in describing members of the family unit and the family unit itself. **Conclusion:** The Swedish translations of ten PROMIS Pediatric item banks and Profile-25 were rigorously translated using internationally standardized methods. Close consideration of the translations, and multiple translations helped to ensure conceptual equivalence and comprehensibility. The banks are culturally adapted and appropriate for the age range 8 to 18 years. They can be used for clinical trials and routine pediatric health care.

### **(3189) Patient-Reported Outcome Qualitative Analysis Review and Documentation (PRO-QARD): an approach to analyzing qualitative data for measure development and content validation**

Karen Kaiser, PhD, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States; Sara Shaunfield, PhD, Northwestern University Feinberg School of Medicine, Chicago, Illinois, United States

**Aims:** The United States Food and Drug Administration and other health organizations have emphasized the importance of qualitative data from key stakeholders for the development and content validation of patient-reported outcome (PRO) measures of symptoms, side effects, and quality of life. While there is some guidance about the types of qualitative data needed for PRO measure development and content validation, we lack best practices for the analysis of these data. This abstract describes the PRO Qualitative Analysis Review and Documentation (PRO-QARD) process we have developed and used to analyze qualitative data for PRO measure development and content validation. **Methods:** The PRO-QARD process for measure

development consists of 6 steps: (1) extract all data for each code; (2) review the data code by code; (3) enter key summary information about each code into a spreadsheet, including number of patients who reported the concept, patient language used to describe the concept, and any diverging meanings within the code; (4) prioritize codes by relevance to the study aims and by the frequency and/or impact of each code; (5) draft items for each high-priority code; and (6) review draft items and select items for measure inclusion. For content validity studies, step 6 is dropped and step 5 becomes: map content onto existing items to inform content validity assessment. **Results:** The PRO-QARD process provides a systematic method for analyzing qualitative data for measure development or content validity assessment. Additionally, the resulting PRO-QARD spreadsheet of summarized data provides excellent, detailed documentation to guide measure development and content validity decisions. **Conclusion:** The PRO-QARD approach provides a blueprint for high-quality data analysis for PRO measure development and content validation. Moreover, by thoroughly reviewing the coded data code-by-code, the PRO-QARD process provides an opportunity to flag coding errors as an additional validity check or identify previously missed themes within the coded data.

### **(3190) Is the treatment satisfaction questionnaire for medication fit for purpose?: A mixed-methods approach**

Ana Maria Rodriguez, PhD MSc PT, IQVIA, Patient-Centered Endpoints, Montreal, Quebec, Canada; Paul Williams, MPH, IQVIA, Patient-Centered Endpoints, Paris, France; Alexandra Palmer, MA, IQVIA, Boston, Massachusetts, United States

**Aims:** The Treatment Satisfaction Questionnaire for Medication (TSQM) is the most commonly used generic Patient-Reported Outcome (PRO) measuring treatment satisfaction. Its psychometric properties have been assessed in numerous intervention trials using Classical Test Theory (CTT). Given the heterogeneous treatment experience of people taking different medicines in routine health care, the psychometric properties of the TSQM may be challenged in non-interventional or patient preference research. The objectives were to determine the extent to which the psychometric properties of the TSQM are supported across various conditions in a real-world context. **Methods:** Electronic versions of both the TSQM v1.4 and vII were administered to subjects of an online patient community managed by IQVIA (MediGuard) in the US, in the UK, in Spain, in France, and in Australia. CTT and Rasch measurement theory were used to determine, compare, and contrast the cross-sectional psychometric properties of the TSQM. **Results:** The survey was completed by 108 and 354 patients in the UK and in the US, respectively. Most patients who responded had university-level education (67%), were female (68%) and had a mean of 50 ( $\pm$  18.4) years of age. CTT analyses confirmed that the TSQM had excellent internal consistency and construct validity. High floor effects were seen for several items of the TSQM in some disease-level subgroups of patients (e.g., hypothyroidism). Rasch measurement theory (RMT) analyses showed that items were consistently ordered on an interval scale from low to high levels of treatment satisfaction. The Differential Item Functioning and Person-Item Threshold distribution suggests a potential alternate scoring structure that may be more generalizable across therapeutic areas in routine health care. **Conclusion:** The TSQM demonstrates generally high psychometric evidence for its use in a heterogeneous patient population in routine health care. However, the TSQM scoring could be revisited to improve its fit-for-purpose applicability.

### (3191) Enabling cross-cultural data pooling in trials: Linguistic validation of head and neck cancer measures for Indian patients

Chindhu Shunmuga Sundaram, The University of Sydney, Sydney, Australia; Haryana Dhillon, PhD, The University of Sydney, Sydney, Australia; Phyllis Butow, PhD, The University of Sydney, Sydney, Australia; Puma Sundaresan, PhD, The University of Sydney, Sydney, Australia; Claudia Rutherford, PhD, The University of Sydney, Sydney, Australia

**Aims:** Head and neck cancers (HNC) and their treatments cause dysfunction and distress. Ongoing psychological assessment using disease-specific patient-reported outcome measures (PROMs) may optimize clinical decision-making, facilitate interventions to reduce psychosocial burden. Most measures are developed in English, disadvantaging non-English speaking patients. HNCs are highly prevalent in developing countries such as India due to high rates of tobacco smoking and chewing. Also, with changing global migration patterns, more Indian immigrants are diagnosed with HNC around the globe. This study translated measures (Body Image Scale (BIS), Patient Concerns Inventory (PCI), Zung's Self-Rating Anxiety (SAS) and Patient Health Questionnaire-9 (PHQ-9)) suitable for use in HNC populations into three Indian languages (Hindi, Tamil and Telugu) and linguistically validated them. **Methods:** Translation followed EORTC/MAPI guidelines on linguistic validation. Process involved two independent forward translations, reconciliation, two independent backward translations by bilingual experts, and cognitive debriefing interviews with healthcare professionals (HCPs) and HNC patients. Analysis included translation report which detailed issues arising during each step and their solutions. The percentage of responses to item difficulty, ambiguity, sensitivity and comprehension were also calculated. Translated versions were compared with the original versions for semantic, cultural and conceptual equivalence. **Results:** Overall, 17 Hindi items, 19 Tamil items and 13 Telugu items were identified as having semantic, cultural, and/or conceptual issues. These were resolved to achieve equivalence with the original measures. Interviews with nine HCPs indicated equivalent terms for words such as anxiety, panicky, sexuality, and self-conscious might be difficult to understand. Interviews with 29 patients indicated all items were understandable, easy, sensitive, unambiguous and relevant. Hence, no further revisions were made as the overall comprehension rates were high. **Conclusion:** The translated Hindi, Tamil and Telugu versions of the Body Image Scale, Patient Concerns Inventory, Zung's self-rating Anxiety Scale and Patient Health Questionnaire-9 measures are conceptually and linguistically validated and equivalent with the original English versions. Psychometric validation of these measures with relevant patient populations is needed.

### (3192) Patient perceptions of outcomes of immunotherapies for cancer: a qualitative study

Rebecca Wiltshire, BSc Hons, RN, University of Oxford, Oxford, United Kingdom; Mark Middleton, FRCP, DPhil, University of Oxford, Oxford, United Kingdom; Michele Peters, PhD, University of Oxford, Oxford, United Kingdom

**Aims:** Immunotherapy is changing the landscape of cancer care. It is not clear whether current patient-reported outcome measures (PROMs) are valid in this population in this modality. This study aimed to explore patients' quality of life when receiving immunotherapy for cancer, with a view to identify items for inclusion in a PROM. **Methods:** Qualitative interviews were conducted with 30 patients who had received or were receiving checkpoint inhibitor immunotherapy for cancer. Participants were recruited at Oxford

University NHS Foundation Trust through their treating oncologist. Interviews were audio-recorded, transcribed verbatim and analyzed using a thematic approach. Candidate items were developed to represent the identified themes. **Results:** Nineteen men and eleven women participated, with a mean age of 65.2 years. Their cancer diagnoses were melanoma ( $n = 14$ ), lung ( $n = 9$ ) and oesophageal ( $n = 6$ ). Eight themes were identified: (1) side effects and symptoms, (2) roles and responsibilities, (3) leisure and social activities, (4) dependency on others, (5) stigma, (6) uncertainty, (7) finances, and (8) burden of care. These themes, in particular the type of side effects and burden of care, differ from those identified for patients receiving chemotherapy or radiotherapy. Additionally, the higher intensity, longer duration and more frequent nature of immunotherapy treatment impacted participants differently on their roles and responsibilities. A proportion of participants reported feeling well enough to continue with their usual life (such as work), but were not necessarily able to because of the time spent accessing cancer treatment. **Conclusion:** Patients receiving immunotherapy for cancer seem to experience of quality of life differently than patients receiving conventional systematic anti-cancer therapies. This means that measures such as the EORTC QLQ C30, the most widely used PROM for cancer, may not be valid in people receiving immunotherapy. The development of an immunotherapy specific PROM is important to measure quality of life in this population. In the next steps of this research, candidate items are being identified and pre-tested to inform a new measure to assess the impact of receiving immunotherapy for cancer.

### (3193) Evaluating the readability of the EORTC Quality of Life Questionnaire QLQ-C30

Dagmara Kulis, MA, EORTC, Brussels, Belgium; Anne Stone, EORTC, Brussels, Belgium; Anne-Sophie Darlington, University of Southampton, Southampton, United Kingdom; Andrew Bottomley, EORTC, Brussels, Belgium; submitted on behalf of the EORTC Quality of Life Group

**Aims:** The European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Core Questionnaire (QLQ-C30) is a widely used questionnaire to assess quality of life of cancer patients. This abstract describes the first phase of our study, aiming to determine whether the QLQ-C30 is easy to read and understand in its current form. **Methods:** The English QLQ-C30 was analyzed with the Flesch-Kincaid Readability test, a standard functionality of Microsoft Word, widely used for readability assessment. We reformatted QLQ-C30 to remove question numbers, response scales, headers and footers, and added the time frame ('during the past week') to each applicable item (questions 6-28), as this preamble forms part of the question. **Results:** The reformatted QLQ-C30 scored 5.9 Flesch-Kincaid Grade, indicating readability at the level of 6th grade in the American education system. Item scores ranged 0.0–12.7 (median 4.7; mean 5.29). The C30 comprises of 36 sentences with an average of 11 words/sentence, 53 characters/sentence, and 5 characters/word. There is one passive sentence (grade 5.2). 25% ( $n = 9$ ) of sentences were above grade 8. Analysis of these showed all contained words with  $3 \geq$  syllables (33% had two such words;  $67\% \geq 3$ ) and scored above average in words/sentence, characters/sentence and characters/word. On average instructions scored slightly lower (5.2) than items (5.3). The lowest graded item (0.6) contains below average number of words ( $n = 7$ ), characters ( $n = 31$ ), average number of characters/word ( $n = 4.1$ ) and no words with  $3 \geq$  syllables. Words with  $3 \geq$  syllables were both commonplace and less frequently used words: 'information,' 'confidential,' 'activities,' 'limited,' 'pursuing,' 'strenuous,' 'interfered,' 'concentrating,' 'newspaper,' 'television,' 'remembering,' 'condition,' 'medical.' **Conclusion:** Based on the

results, the QLQ-C30 can be considered easily readable. However, given the shortcomings of standard readability formulas (focus on quantitative aspects of language, no account of semantics or syntax, development for and validation in other types of texts), future steps of research are planned and include cognitive debriefing with patients from a wide range of ages and education levels to collect qualitative data on the English version and its translations.

#### **(3194) Thematic analysis in integrated health technology assessment (HTA)**

Bethany English, MSc Economics and Health Economics, Analyst, Milton Keynes, United Kingdom; Andrew Mumford, BSc Chemistry, Principal Consultant—Market Access, Northampton, United Kingdom; Lorraine Munetsi, MSc Economics and Health Economics, Health Economist, Northampton, United Kingdom

**Aims:** The purpose of this research is intended to explore how thematic analysis may be used and add value within the framework of an integrated health technology assessment (HTA) submission. This research will assess how current practices for thematic analysis may be developed in order to produce uniform guidelines for the use of qualitative research. **Methods:** In the context of technology appraisals, qualitative research can provide an in-depth understanding of areas, such as attitudes, viewpoints and patient experience. Thematic analysis is a method of synthesizing qualitative data from several sources, which allows us to make conclusions that go beyond the findings of individual primary studies. We first review the current methods used, by outlining the main steps in conducting a thematic analysis, illustrating this process using a case study. Thematic analysis is separated into three stages. Text is first coded in order to identify themes, resulting in the identification of various descriptive themes, which reflect the outcomes of the primary study. This further leads to the development of analytical themes, underpinning each individual primary study. We then explore how uniform guidelines may be developed for the integration of qualitative analysis into HTAs. **Results:** Conclusion: Thematic analysis is a robust method for the analysis of qualitative data. However, there are currently no set guidelines in place for how the results of such analysis should be incorporated into HTA submissions. The implementation of uniform guidelines would be beneficial for the integration of qualitative data into HTAs. This research is a contribution to the development of how best to incorporate qualitative data into HTA submissions.

#### **(3195) Recommendations to enable patients to discuss physical functioning concepts in a way they can understand during qualitative interviews**

Hannah C Pegram, BSc, DRG (part of Clarivate), Manchester, United Kingdom; Chloe Johnson, MSc, DRG (part of Clarivate), Manchester, United Kingdom; Maile Hunter, MBA, DRG (part of Clarivate), Nashville, Tennessee, United States; Sarah Knight, MSc, DRG (part of Clarivate), Bicester, United Kingdom; Jake Macey, MSc, DRG (part of Clarivate), Bicester, United Kingdom; Natalie V J Aldhouse, MSc, DRG (part of Clarivate), Bicester, United Kingdom; Tamara Al-Zubeidi, BSc, DRG (part of Clarivate), Bicester, United Kingdom; Helen Kitchen, MSc, DRG (part of Clarivate), Manchester, United Kingdom; Chris Marshall, MSc, DRG (part of Clarivate), Manchester, United Kingdom

**Aims:** To present best practices for exploring the concept of physical functioning (PF) in qualitative patient interviews. **Methods:** Researchers conducted a retrospective review of qualitative interview studies exploring PF to appraise: how well patients understood the

concept; terms patients used related to PF; and techniques implemented in interview materials to aid patient understanding. **Results:** PF is commonly assessed in clinical trials using patient-reported outcome measures, particularly when treatment is expected to directly impact patients' movement and mobility. While 'PF' may be a familiar concept to the research team, patients can find PF difficult to conceptualize; thus, patient interpretation of PF may not always be consistent with the researcher's intended meaning of the concept. Several techniques were identified during the review as helping patients understand the concept of PF, particularly when linked to their ability to perform activities of daily living (ADLs). Before the interviews, researchers should consider how the interview guide is structured to support discussion of this topic in a manner patients can understand. Relating PF to patients' abilities to perform ADLs at work, home, and during leisure activities, may facilitate patients in identifying limitations in PF. During interviews, patient understanding of PF appeared to be strongest when questions explored: ADLs that required physical function (e.g., 'climbing stairs' rather than 'flexion/extension'), level of difficulty with PF (e.g., number of stairs climbed, distances walked), need for aids (e.g., walking stick/handrails), and frequency of impact. Interviewers used patient-friendly language (e.g., 'ability to walk' rather than 'ambulation') and probing questions to explore specific PF movements/tasks after initial open-ended questions. Asking participants to provide a diary/event timeline can be a useful method to explore PF impacts in daily life. Following the interviews, analysis should focus on the physical aspect of the task, not the task itself, and ADLs affected by physical impairment rather than other experiences (e.g., embarrassment, fatigue). **Conclusion:** It is important to explore understanding of PF using patient-centric language during interviews. With careful design of qualitative research, COA researchers can ensure patient understanding of the intended concept, relevant discussion, and support the content validity of PRO measures in PF.

#### **(3196) Evidence-based practice questionnaire for physical therapists: Portuguese translation, adaptation, validity, and reliability**

Ricardo Maia Ferreira, Physical Education and Sports Department, N2i, Polytechnic Institute of Maia, Castelo da Maia, Porto, Portugal; Pedro Lopes Ferreira, Faculty of Economy, University of Coimbra, Coimbra, Portugal; Luís Cavalheiro, Coimbra Health School, Polytechnic Institute of Coimbra, Coimbra, Portugal; José Alberto Duarte, Faculty of Sport, CIAFEL, University of Porto, Porto, Portugal; Rui Soles Gonçalves, Coimbra Health School, Polytechnic Institute of Coimbra, Coimbra, Portugal

**Aims:** Evidence-Based Practice (EBP) is becoming increasingly important in Physical Therapy (PT). For proper designing, implementing, disseminating and evaluating EBP in PT, a valid and reliable questionnaire measuring attitudes, knowledge, behavior, prerequisites, and barriers related to EBP and guidelines is needed. One questionnaire that could be used to collect this information is the "EBP Questionnaire," developed by Jette et al. 2003. However, to our knowledge, no Portuguese version or published study with the Portuguese PT population was performed using this questionnaire. The aim of this study was to translate, cross-culturally adapt and validate the "EBP Questionnaire" to European Portuguese and for the PT population. **Methods:** A draft version was pilot tested for content validity ( $n = 17$ ), and a revised version was tested for test-retest reliability ( $n = 72$ ). The percentage of agreement and the Kappa coefficients between the 2 tests were analyzed. Additionally, the internal consistency was calculated. **Results:** The preliminary final version of the European Portuguese EBP questionnaire was well accepted (only the items 22, 23, 45-51 needed to be reviewed). The

mean average percentage of agreement was 82% (ranged 58–97%), and the Kappa coefficients were 0.658 (ranged 0.336–0.844). In the internal consistency, the mean average Cronbach's  $\alpha$  coefficients were 0.665 (ranged 0.365–0.879). Conclusion: The results suggested that this questionnaire can be a useful instrument for measuring self-reported beliefs, attitudes, knowledge, and behaviors related to EBP in the Portuguese PT's context.

## Research in Caregivers, Relatives & Vulnerable Populations

### (3197) Health-Related Quality of Life of French night-shift hospital staff in Paris public hospitals

Adriano Cannafarina, ECEVE UMR 1123, Inserm, Université de Paris, AULNAY SOUS BOIS, France; Pauline Chaussareau, APHP—Service d'Addictologie, Hôpital Paul Brousse, Paris, France; Lorraine Cousin, ECEVE UMR 1123, Inserm, Université de Paris, Paris, France; Olivia Rousset-Torrente, ECEVE UMR 1123, Inserm, Université de Paris, Paris, France; Laurence Baumann, Unité de médecine générale Université de Paris, Paris, France; Isabelle Chavignaud, Mission FIDES—Hôpital Paul Brousse, Paris, France; Sarah Coscas, APHP—Service d'addictologie, Hôpital Paul Brousse, Paris, France; Patrizia Carrieri, SESSTIM U1252, Inserm, IRDES, Université Aix Marseille, Paris, France; Olivier Chassany, ECEVE UMR 1123, Inserm, Université de Paris, Paris, France; Martin Duracinsky, ECEVE UMR 1123, Inserm, Université de Paris, Paris, France; ECEVE UMR 1123, Inserm, Université de Paris, Paris, France

Aims: Studies revolving around consequences of working at night defined several symptoms, yet health-related quality of life (HRQL) has not been approached in a qualitative way. Thus we aimed to explore different HRQL aspects of night-shift hospital staff and to know if they desired interventions targeting to improve their HRQL. This is the first of two qualitative studies and a part of a larger mixed-method research project on HRQL. Methods: Volunteer professionals working night-shift (except physicians) were interviewed in order to reach data saturation. They were from different departments, age, and positions. Interviews were semi-directive and were afterwards analyzed using grounded theory by at least 3 different researchers to minimize subjective biases. Themes addressed during the interviews include socio-demographic data, personal life habits, health, professional identity and experience about their work. Results: Data saturation was reached at the 15th interview, 4 more were done after that to confirm it. There were 13 women and 6 men, with a median age of 40 years (min–max: 23–61). 10 were nurses, 6 caregivers, 1 executive, 1 auxiliary and 1 midwife. The median duration of night work was 12 h. The main health issues related to night work were sleep problems, weight gain, fatigue, irritability and social isolation. We found two main profiles: younger professionals who are interested in learning and said that they were not correctly informed, if at all, of the health consequences of working night-shifts for a long period of time. The second one includes older professionals who developed their way to cope with the symptoms but did not feel that their work valued due to an underestimation and prejudice stereotype. Overall, night-shift hospital staff declared enjoying the night's ambiance at work, its feeling of freedom, peacefulness. In small night teams, teamwork was also considered necessary in order to work correctly, along with responsibility and autonomy. Conclusion: Night-shift hospital staff declared to have, in overall, a good quality of life which relies on strong professional identity and teamwork. However, they expressed a need for an intervention regarding the consequences of working at night and to feel recognized from peers.

### (3198) A more holistic approach to quality of life in patients with psoriasis: a qualitative study

Ashley Snyder, MPH, University of Utah, Salt Lake City, Utah, United States; Vanina Taliercio, MD, MS, University of Utah, Salt Lake City, Utah, United States; Abram Beshay, MD, University of Utah, Salt Lake City, Utah, United States; Dominik Ose, DrPH, University of Utah, Salt Lake City, Utah, United States; Rachel Hess, MD, MS, University of Utah, Salt Lake City, Utah, United States; Jamie Rhoads, MD, MS, University of Utah, Salt Lake City, Utah, United States; Aaron Secrest, MD, PhD, University of Utah, Salt Lake City, Utah, United States

Aims: Psoriasis is a chronic inflammatory skin condition characterized by red, scaly plaques ranging from a few itchy spots to full-body distribution. Psoriasis can have significant impact on patients' lives. However, dermatologists often focus their assessments on the characteristics of the plaques and body surface(s) involved with less focus on other aspects of the disease that can impact quality of life. In this study, we assessed how psoriasis impairs various aspects of quality of life for psoriasis patients and their family members. Methods: We conducted five focus groups and ten semi-structured interviews with psoriasis patients ( $n = 25$ ) and their family members ( $n = 11$ ) seeking to understand their daily struggles with psoriasis. All groups were moderated by a trained facilitator using a semi-structured interview guide. Two researchers independently coded narratives and identified major themes using the grounded-theory approach. NVivo 12 software was used in managing codes. Results: Psoriasis patients' median age was 46 years (range 22–82); for family members, median age was 38 years (range 28–72). About 47% of participants were female, and about 89% were white. From our thematic analysis, several overarching themes appeared that were important to patients and their family members: ((1) Symptoms<sup>3/4</sup>the demanding continuous state of itchiness and the presence of intermittent, but often severe, pain; (2) Social hardships—tension in family and friend relationships, sexual relations, and interactions with strangers; (3) Emotional consequences—psychological impacts, effects on self-image, and tendency for depression and negative moods; (4) Daily activities<sup>3/4</sup>impairment was common, including but not limited to work, leisure, and sleep. Conclusion: From our results, it seems apparent that psoriasis is a significant factor for quality of life based on the degree to which social, emotional, and daily activity aspects were affected. Rather than focusing on symptoms and appearance, we encourage dermatologists to inquire more holistically into how psoriasis impacts their patients' quality of life or use a dermatology-specific quality of life measure to best understand the true impact psoriasis is having on their patients.

### (3199) Patient perceptions and understanding of a housing status measure for use in HIV care

Bridget M. Whitney, PhD, MPH, University of Washington, Seattle, Washington, United States; Emma M. Fitzsimmons, University of Washington, Seattle, Washington, United States; Matthew K. Jackson, University of Washington, Seattle, Washington, United States; Sharon M. Brown, University of Washington, Seattle, Washington, United States; John Nguyen, University of Washington, Seattle, Washington, United States; Robin M. Nance, MS, University of Washington, Seattle, Washington, United States; Stephanie Ruderman, University of Washington, Seattle, Washington, United States; Joseph C. Delaney, PhD, University of Manitoba, Winnipeg, Manitoba, Canada; Heidi M. Crane, MD, MPH, University of Washington, Seattle, Washington, United States; Rob J. Frederickson, PhD, MPH, University of Washington, Seattle, Washington, United States

States; PhD, MPH, University of Washington, Seattle, Washington, United States

**Aims:** Assess patient comprehension and interpretation of items in a patient-reported assessment of current housing status for use in routine HIV care. **Methods:** We conducted cognitive interviews ~ 30 min in length with primarily unstably housed patients living with HIV prior to routine care visits in a US hospital-based outpatient HIV clinic. We showed patients a proposed measure adapted from several sources assessing current housing status comprised of two items: one with ten response options querying places stayed at least one night in the past month (e.g., “shelter,” “someone else’s home or apartment”), and a self-description of one’s housing status (e.g., “stable,” “unstable,” “homeless”). We asked patients to interpret items and response options in their own words, and assessed perceptions of the recall period. Patients were remunerated \$20 for participation. We transcribed patient responses into an Excel matrix for thematic analysis of interpretations of each item/response option. Two trained qualitative analysts independently summarized themes and facilitated discussion to reconcile interpretive differences. **Results:** Patients ( $n = 9$ : mean age 47; 8 white, 1 Asian-American; 4 male, 4 cisgender female, 1 transgender female) consistently showed uniform interpretation of most items and response options, with notable exceptions: “long-term care facility,” “boarding house,” and “shelter.” Patients found the concept of “unstable” housing vague and in need of supportive clarifying text. The concept of “couch surfing” was well-understood but regarded as duplicative of another item (“someone else’s home or apartment”); some found this term to imply voluntary or fun activity exclusive to younger patients. Patients had mixed views on appropriate recall period, either preferring the existing 30-day window or suggesting a longer period (typically 3 months) to capture patterns. Several patients, particularly women, found it more important to query perceptions of personal safety in living situations rather than solely asking about types of physical spaces. **Conclusion:** We created a housing measure in which most items were uniformly understood. Patients identified a need to clarify context for some response options, to eliminate repetitiveness for one item, and to adapt the measure to focus on personal safety in various living situations rather than solely on types of physical spaces.

### **(3200) Assessment of generic and dialysis-related fatigue—identification of individual trajectories and investigation of ecological validity**

Kathrin Fischer, Charité Universitätsmedizin Berlin, Berlin, Germany; Peter J. Blankestijn, Prof. Dr., UMC Utrecht, Utrecht, Netherlands; Krister Cromm, Fresenius Medical Care Deutschland GmbH, Bad Homburg, Germany; Michiel L. Bots, Prof. Dr., UMC Utrecht, Utrecht, Netherlands; Claudia Barth, Prof. Dr., B. Braun Avitum AG, Melsungen, Germany; Bernard Canaud, Prof. Dr., Fresenius Medical Care Deutschland GmbH, Melsungen, Germany; Jörgen Hegbrant, Dr., Lund University, Lund, Sweden; Fernando Macário, Dr., Diaverum Renal Services Group, Lund, Sweden; Belén Marrón, Dr., Diaverum Renal Services Group, Lund, Sweden; Marietta Török, Dr., Diaverum Renal Services Group, Lund, Sweden; Andrew Davenport, Dr., Diaverum Renal Services Group, Lund, Sweden; Andrew Davenport, Prof. Dr., University College London, Royal Free Hospital, London, United Kingdom; Mei-Man Lee, Dr., The George Institute for Global Health, University of Oxford, Oxford, United Kingdom; Giovanni FM Strippoli, Prof. Dr., University of Bari Aldo Moro, Bari, Italy; Luisa Wohn, B. Braun Avitum AG, Melsungen, Germany; Mark Woodward, Prof. Dr., The George

Institute for Global Health, University of Oxford, Oxford, United Kingdom; Matthias Rose, Prof. Dr., Charité - Universitätsmedizin Berlin, Berlin, Germany; submitted on behalf of the CONVINC Consortium

**Aims:** Fatigue is considered one of the most important patient-reported outcomes (PROs) in the field of end-stage kidney disease (ESKD). However, it is not only related to the disease itself, but also a consequence of the continuous renal replacement therapy. Thus, different aspects of fatigue in ESKD patients undergoing dialysis need to be addressed, such as generic fatigue as well as dialysis-related fatigue. This study investigates the assessment of fatigue in patients with ESKD including trajectory patterns related to hemodialysis sessions, and explores ecological validity of the fatigue instrument from the Patient-Reported Outcomes Measurement Information System (PROMIS®). **Methods:** PRO data of the CONVINC trial (Trial NL 6942) is used. While recruitment is ongoing, currently  $n = 423$  hemodialysis patients have completed the screening assessment of the trial, using an electronic data capturing tool. Generic fatigue is measured by a six-item customized short form of PROMIS®. Trajectories of dialysis-related fatigue (i.e., before, during, and three different time points after dialysis) are captured using a retrospective time point-related measurement approach. Dialysis-related fatigue was analyzed as quasi longitudinal data. Ecological validity was evaluated by correlating the PROMIS® fatigue score with the dialysis-related measure. Growth curve mixture models were fitted to explore latent fatigue trajectory patterns (i.e., latent classes of individual fatigue trajectories) among the hemodialysis patients. **Results:** Mean PROMIS® fatigue T-score of hemodialysis patients was 49.66 (SD = 9.24). Generic fatigue scores and dialysis-related fatigue items correlated moderately between  $r = .49$  and  $r = .61$ . A fatigue trajectory curve including two independent fatigue peaks (during dialysis and the evening after dialysis) followed by a recovery phase with decreasing fatigue levels approximated the data well. Three distinct classes with different fatigue trajectories (low fatigue, high fatigue, peak fatigue) were identified. **Conclusion:** The retrospective PROMIS fatigue short-form does only partly reflect fatigue levels before, during and after dialysis. The identification of subgroups of fatigue trajectories during dialysis could help to improve individual treatment in hemodialysis patients.

### **(3201) “I tried melatonin and some plants, but now, I’m on zopiclone”: How mental health and substances consumption affect quality of life in hospital night shift workers**

Lorraine Cousin, PhD, U1123 ECEVE and URC eco, Inserm, APHP, Université de Paris, Paris, France; Olivia Rousset Torrente, Master Degree (MD), U1123 ECEVE and URC eco, Inserm, APHP, Université de Paris, Paris, France; Laurence Baumann, PhD, MD, Département de médecine générale, Université de Paris, Paris, France; Adriano Cannafarina, Master Degree, URC eco, APHP, Paris, France; Isabelle Chavignaud, FIDES Mission APHP, Paris, France; Sarah Coscas, MD, FIDES Mission APHP, Paris, France; Olivier Chassany, Professor, PhD, MD, U1123 ECEVE and URC eco, Inserm, APHP, Université de Paris, Paris, France; Patrizia Carrieri, PhD, U1252 SESSTIM, Inserm, IRDES, Université Aix Marseille, Marseille, France; Martin Duracinsky, PhD, MD, U1123 ECEVE and URC eco, Inserm, APHP, Université de Paris, Paris, France

**Aims:** Night shift healthcare workers (NHW) are exposed to several risk factors (e.g., alteration of circadian cycles and associated sleep disorders, anxiety and stress, limited social interaction) potentially

leading to impaired quality of life (QoL) and addictive behaviors. We aimed to explore how night shift affected QoL and health of NHW and to what extent addictive behaviors mediate this relationship. Methods: We conducted semi directive interviews ( $n = 18$ ) with 6 nurses, 2 laboratory and 2 radio technicians, 1 midwife, 4 care givers, and 3 night managers working at Public Hospitals of Paris. Interviews were analyzed relying on the grounded theory. Results: Age ranged between 22 and 56, 61% were men. Years of work experience varied between 2 and 30, workplace included intensive care/emergency units and medical wards. Participants' QoL was affected by the number of years of night work, the impact of work on sleep, their family status and their relations with hierarchy and colleagues. The feeling of not being synchronized with their social relations and the impact of night shift on their sleep, causing sleep disturbances and facilitating addictive behaviors were the most reported issues. Seventeen reported consuming at least one substance: tobacco, alcohol, cannabis or sleeping pills. Work influenced their consumption habits which increased during stressful periods or decreased due to lack of time for a work break. Thus, the work context could be either a protective or a risk factor for specific addictive behaviors. By affecting sleep and mental health, shift work decrease the NHW's QoL. Cannabis and sleeping pills can be perceived as a solution to reduce the harmful consequences of night work on their sleep, while tobacco and alcohol appear to be an outlet for overwork. These substances also enable them to adjust to the "daytime rhythm" necessary to maintain social and family life. Conclusion: The type of substance use is involved in the relationship between how night shift affect sleep disorders, social relationship or stress management in NHW. Quantifying the prevalence, the way substance affect QoL and which interventions are needed for this under-diagnosed population are to be considered in our research agenda.

### **(3202) Effects of supported exercise training on health-related quality of life in patients with type 2 diabetes**

Eren Timurtaş, PhD, Marmara University, Istanbul, Turkey; Mehmet Inceer, PhD (c), McGill University, Montreal, Quebec, Canada; Nancy Mayo, Prof., McGill University, Montreal, Quebec, Canada; Yasar Sertbas, Assoc. Prof., Fatih Sultan Mehmet Education and Research Hospital, Istanbul, Turkey; Mine Gulden Polat, Prof., Marmara University, Istanbul, Turkey

Aims: The management of diabetes includes pharmacological interventions and structured self-management practices to incorporate diet and exercise. Ongoing self-management of diabetes include exercise education and lifestyle changing interventions are quite common in the literature. However, studies that focus on the health-related quality of life (HRQL) of the patients are limited. The purpose of this study is to demonstrate the change in the quality of life, for Type-2 Diabetes (T2D) patients after receiving 12 weeks exercise program. Methods: The study is an analysis of data obtained from a randomized controlled trial designed to test the comparative effectiveness of structured exercise intervention delivered through smartphone and smartwatch for glycemic control in individuals with T2D. Health-related Quality of life measured with EQ-5D-5L was compared for the cohort before and after the trial. Results: A total of 84 people (53Women;31Men) completed the study with mean age 51 years. At baseline, the most commonly endorsed items were slight problems in Mobility, Pain/Discomfort, and Usual Activities (38%, 42.9%, and 56%); and no problem in Self-care and Anxiety (58.3%, 60.7%). While mobility improved from a slight problem to no problem (58.3%), the other four items remained the same after the intervention. In terms of the improvements in the dimensions, mobility increased the most ( $n = 52;62%$ ), pain followed ( $n = 39;46%$ ), then

usual activities ( $n = 31;37%$ ), self-care ( $n = 17;20%$ ), and anxiety ( $< 5;5%$ ). VAS scores ranged from 0 to 100 with a mean 54.1 and SD 24.9 at baseline, and it improved at time two to mean 67.8 and SD 19.5. Conclusion: The EQ-5D-5L index scores varied among T2D patients. Structured exercise programs are effective at increasing mobility as an aspect of quality of life in T2D. Including HRQL measures in clinical care and trials are highly important to understand the subgroups that are at risk. Targeted approaches to reduce these problems should be considered in T2D care and research.

### **(3203) The moderating role of spirituality on quality of life and caregiver burden among the parents of adolescents with Spina Bifida**

Asma Taha, PhD, Oregon Health & Science University, Portland, Oregon, United States; Aaron M. Eisen, BS, National University of Natural Medicine, Helfgott Research Institute, Portland, Oregon, United States; Hana Abdul-Rahman, Portland State University, Portland, Oregon, United States

Aims: To investigate the relationships between caregiver burden, depressive symptoms, spirituality, and quality of life among parental caregivers of adolescents diagnosed with spina bifida while considering demographic factors. Methods: In this exploratory cross-sectional study, fifty-eight caregivers of adolescents with spina bifida in southern California were recruited during routine visits to a multidisciplinary clinic at a healthcare university from January 2016 to January 2017. Each parent completed a series of self-report scales including the Patient Health Questionnaire, Zarit Burden Interview, System of Belief Inventory, and the Caregiver Quality of Life Index Revised. Results: The mediation-moderation analysis showed that caregiver burden partially mediated the relationship between depressive symptoms and quality of life ( $B = 0.08(0.03)$ , CI 95% [0.03–0.15]), and spirituality moderated the relationship between caregiver burden and quality of life ( $\beta = 0.396$ ,  $p < .01$ ). Depressive symptoms did not mediate the relationship between caregiver burden and quality of life ( $B = 0.08(0.01)$ , CI: 95% [– 0.01, 0.3]). Conclusion: Parents with higher levels of caregiver burden and depressive symptoms had a lower quality of life and parents who were more spiritual had a higher quality of life. Caregivers with greater levels of spirituality had a higher quality of life when caregiver burden was moderate/high, but no differences were noted when caregiver burden was low. Caregiver burden appeared to have a more profound effect on quality of life compared to depressive symptoms. Accordingly, we recommend that health care professionals actively screen for caregiver burden in parental caregivers of adolescents with SB.

### **(3204) The Observer-Reported Communication Ability (ORCA) measure: Validation for use in clinical trials for individuals with Angelman syndrome**

Christina Zigler, PhD, MEd, Duke University School of Medicine, Durham, North Carolina, United States; Li Lin, Duke University School of Medicine, Durham, North Carolina, United States; Nicole Lucas, Duke University School of Medicine, Durham, North Carolina, United States; Molly McFatrach, Duke University School of Medicine, Durham, North Carolina, United States; Allyson Berent, Foundation for Angelman Syndrome Therapeutics, Downers Grove, Illinois, United States; Paula Evans, Foundation for Angelman Syndrome Therapeutics, Downers Grove, Illinois, United States; Kelly Gordon, Duke University, Durham, North Carolina, United States; Harrison Jones, Duke University, Durham, North Carolina,

United States; Jennifer Panagoulas, Foundation for Angelman Syndrome Therapeutics, Downers Grove, Illinois, United States; Bryce Reeve, Duke University School of Medicine, Durham, North Carolina, United States; Duke University School of Medicine, Durham, North Carolina, United States

**Aims:** The Observer-Reported Communication Ability (ORCA) measure was developed to assess caregiver perceptions of communication ability for their children with Angelman syndrome (AS). This survey was created specifically for use in clinical trials, as no existing tools were appropriate for this context. Prior development work included detailed concept elicitation interviews with caregivers and speech language pathologists, and two rounds of cognitive testing. The aim of this study was to evaluate the psychometric properties of the ORCA measure. **Methods:** The ORCA measure was administered to adult caregivers of children with AS (aged > 2 years) via a REDCap survey. Caregivers also reported demographic information (including AS subtype of their child) and completed other observer-reported measures including PROMIS® Parent-Proxy measures of Mobility and Sleep Disturbance, and the Communication and Symbolic Behaviors Scale (CSBS). A subset of caregivers completed the ORCA measure again within 5–12 days of the first administration. We calculated overall scores representing communication ability at the ‘mastery’ level using item response theory. We evaluated measurement properties such as structural validity, reliability, and construct validity (convergent and known-groups). **Results:** The final ORCA measure included 73 items nested within 22 concepts, and was administered to 295 caregivers (Table 1). A one-factor confirmatory factor analysis model found evidence for good model fit; CFI = 0.955; TLI = 0.950; RMSEA = 0.064, 90% CI 0.055–0.073. Internal consistency and reproducibility were excellent (Cronbach’s alpha = .90; intraclass correlation = .90, respectfully). Individuals with Deletion Positive AS genotype had communication ability scores that were significantly lower than the other genotypes ( $p < .001$ ). Total scores on the ORCA measure (Fig. 1) were strongly correlated with the CSBS total score ( $r = .83$ ), moderately correlated with mobility ( $r = .54$ ), and not correlated with sleep disturbance ( $r = -.09$ ). **Conclusion:** The ORCA measure is the first tool designed specifically from the caregiver perspective to assess the communication ability of individuals with AS. These results provide strong evidence for the content validity, construct validity, and reliability of the ORCA measure. Future work will incorporate the ORCA measure in upcoming trials as an exploratory endpoint and evaluate responsiveness of the measure over time.

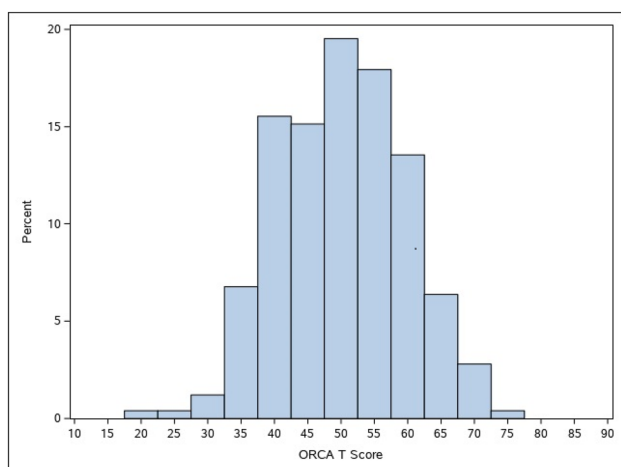


Table 1. Basic demographic information for 295 caregivers who participated in the study and the individual with AS that they care for.

Caregivers	n(%)
Non-white	20 (7.2)
Hispanic or Latino	22 (7.9)
Female	247 (88.2)
Age (Mean/STD)	41.6/8.6
Individuals with AS	n(%)
Subtype	
Deletion Positive	199 (67.5)
Mutation/UBE3A	56 (19.0)
Imprinting (ICD)	14 (4.7)
Uniparental Disomy (UPD)	26 (8.8)
Female	125 (44.6)
Non-white	36 (12.9)
Hispanic-Latino	29 (10.6)
Age (Mean/STD)	10.5/7.8

### (3205) HRQOL in adolescence parents

Gudrun Rohde, Faculty of Health and Sport Sciences, University of Agder, and Department of Clinical Research, Sorlandet Hospital, - Kristiansand, Norway; Sølvi Helseth, PhD, Department of Nursing and Health Promotion, Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway; Hilde Timenes Mikkelsen, PhD-student, Department of Health and Nursing, Faculty of Health and Sport Sciences, University of Agder, Kristiansand, Norway; Siv Skarstein, PhD, Department of Nursing and Health Promotion, Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway; Milada Småstuen, PhD, Department of Nursing and Health Promotion, Faculty of Health Sciences, Oslo Metropolitan University, Oslo, Norway; Kristin Haraldstad, PhD, Department of Health and Nursing, Faculty of Health and Sport Sciences, University of Agder, Oslo, Norway

**Aims:** Adolescence is a unique and complex developmental phase characterized not only by significant physical and cognitive changes, but also critical psychosocial challenges, related to self-identity, peer relationships, development of autonomy, and sexuality. Being a parent and role model for adolescent might be demanding, stressful and influence on HRQOL **Aim:** To explore association between demographic and psychosocial variables, pain and HRQOL in parents of adolescents **Methods:** A cross-sectional study was performed among 561 parents. Data on sociodemographics, self-efficacy, self-esteem, pain, sleep, loneliness, and stress were collected. All variables were measured with well-validated instruments. HRQOL was assessed using SF-36. Data were analyzed using Chi-square tests, independent samples *t*-tests, linear—and hierarchical regression analyses. **Results:** Among the 561 parents, 436 (78%) were women and 125 (22%) men, mean age 45 (SD = 5) years. Eighty-one percent were married/cohabiting, 74% worked full-time and 50% had university education of more than 4 years. Almost one-third reported daily or weekly pain, and more than half (58%) reported using pain analgesic during the previous 4 weeks. Women reported lower scores on self-efficacy, self-esteem, had worse sleep quality and experienced more stress than men. However, there were no statistically significant differences between genders regarding loneliness and pain. Women reported

significantly lower scores for all SF-36 domains, including the physical component summary (PCS) and mental component summary (MCS) scores. When adjusted for demographic and psychological variables and pain in the final model, and when interpreted in terms of effect sizes (standardized B), the covariates not being in a paid job ( $B = -0.33$ ), pain 1–3 months ( $B = 0.14$ ), pain more than 3 months ( $B = -0.38$ ) and stress ( $B = -0.14$ ) revealed the strongest association with SF-36-PCS. Self-esteem ( $B = 0.11$ ) and stress ( $B = -0.61$ ) revealed the strongest association with SF-36-MCS. In the final models, the explained variance of demographic and psychosocial variables and pain for SF-36-PCS was 36.5% and 57.2% for SF-36-MCS. Conclusion: One-third of adolescent parents reported daily or weekly pain. Mothers reported worse psychosocial status and lower HRQOL than fathers. Not being in a paid job, pain, stress and low self-esteem are strongly associated with low HRQOL among parents of adolescents.

### Research in Children and Adolescents I

#### (3206) Yoga training as an effective approach for improving the executive abilities in children with ADHD

Eleonora Mirzajonova, Fergana State University, Fergana, Uzbekistan; Sergey Kiselev, Ph.D., Ural Federal University, Ekaterinburg, Russia

Aims: It is known that children with attention deficit/hyperactivity disorder (ADHD) have deficit in executive abilities. It is very important to develop trainings for children with ADHD to improve their executive abilities and attention. The goal of this study was to reveal effect of yoga training on executive abilities in 8-9 years of age children with this disorder. We compared the efficacy of two methods of training (yoga training vs. conventional motor exercises) in a randomized controlled pilot study. Methods: 18 boys with ADHD at the age of 8-9 years ( $M = 8.41$  years,  $SD = 0.95$ ) were included and randomly assigned to treatment conditions according to a  $2 \times 2$  crossover design. Both groups of children have participated in 12 weeks of training (body-oriented training vs. conventional motor exercises). A total of 36 training sessions lasting 30 min were performed. Yoga training included body-oriented activity and breathing exercises. To assess the executive functions we used 3 subtests from NEPSY (Auditory Attention and Response Set, Visual Attention, Statue). Effects of training were analyzed by means of an ANOVA for repeated measurements. We have also performed qualitative neuropsychological assessment based on Luria's syndrome analysis. Results: The ANOVA has revealed ( $p < .05$ ) that for all subtests (Auditory Attention and Response Set, Visual Attention, Statue) the yoga training was superior to the conventional motor training, with effect sizes in the medium-to-high range (0.42-0.86). Besides, we have found a decrease in distractibility in children from experiment group. In particular, these children showed a decrease in sensitivity to various distracting sounds and environmental events. Luria's syndrome analysis has revealed the improving in third functional unit of the brain which is responsible for voluntary attention and executive abilities according to Luria's approach [Luria, 1973]. Conclusion: The findings from this pilot study suggest that yoga training have positive effect on executive abilities in children with ADHD. It influences predominantly the selective and sustained attention, inhibition, monitoring, and self-regulation. However, it is necessary to do further research for revealing the impact of yoga exercises on the prevention and treatment of attention deficit/hyperactivity disorder in children.

#### (3207) Long-term effect of visuospatial training on the language abilities and visuospatial functions in children with specific language impairments

Sergey Kiselev, Ph.D., Ural Federal University, Ekaterinburg, Russia

Aims: It is known that children with specific language impairments (SLI) have deficit not only in language abilities but also in other cognitive abilities, including visuospatial functions. We have revealed that visuospatial training has positive effect on language abilities and visuospatial functions in 7–9 years of age children with SLI. The goal of this study was to reveal long-term effect of visuospatial training on cognitive abilities in children with SLI. Methods: The participants were 22 children aged 7–9 years ( $M = 8.34$  years,  $SD = 1.79$ , 19 boys and 3 girls) with SLI. Children were randomly assigned to the intervention and comparison group. Children from intervention group participated in 36 weeks of visuospatial training. This training trains the child to do different visuospatial exercises on motor level. This training is built on the conceptual framework derived from the work of Luria's theory of restoration of neurocognitive functions (Luria, 1963, 1974). We used the subtests from Luria's child neuropsychological assessment battery to assess language abilities and visuospatial functions in children one year after training. Effects of training on five language subtest and 4 visuospatial subtests were analyzed by means of an ANOVA for repeated measurements. Results: The ANOVA has revealed that there were significant ( $p < .05$ ) group differences for one language subtest which assesses understanding prepositions that describe the spatial relations between objects and for one visuospatial subtest which assesses the ability for orientation in the body scheme (Head subtest). Posttest mean for the intervention group were significantly ( $p < .05$ ) greater than for the control group. Conclusion: The findings from this pilot study suggest that visuospatial training has not only immediate but also long-term positive effect on cognitive abilities in 7–9 years of age children with SLI. It can be assumed that visuospatial training can be used as a promising way to help children with SLI to overcome deficit in understanding sentences with spatial prepositions and weakness in orientation in body scheme. However, we need to do further research for revealing the impact of visuospatial training on cognitive abilities in children with SLI in the framework of longitudinal research.

#### (3208) Measurement of health-related quality of life in a mainland China adolescent population: feasibility, reliability and validity of the Mandarin Chinese version of the KIDSCREEN-27 and KIDSCREEN-10 index

Yuhang Zhu, Ph. D. Candidate, School of Public Health, Weifang Medical University, Weifang, China; Juan Li, PhD Candidate, School of Public Health, Weifang Medical University, Weifang, China; Gaopei Zhu, School of Public Health, Weifang Medical University, Weifang, China; Anne Wüstner, Department of Child and Adolescent Psychiatry, Psychotherapy, and Psychosomatics, Center for Psychosocial Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Adekunle Adedeji, Department of Child and Adolescent Psychiatry, Psychotherapy, and Psychosomatics, Center for Psychosocial Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Christiane Otto, Department of Child and Adolescent Psychiatry, Psychotherapy, and Psychosomatics, Center for Psychosocial Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; Fuyan Shi, School of Public Health, Weifang Medical University, Weifang, China; Suzhen Wang, School of Public Health, Weifang Medical University, Weifang, China; Ulrike Ravens-Sieberer, Department of Child and Adolescent Psychiatry, Psychotherapy, and



Psychosomatics, Center for Psychosocial Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

**Aims:** This study aimed to evaluate the feasibility, reliability and validity of the Mandarin Chinese versions of the KIDSCREEN-27 and KIDSCREEN-10 index, which are shorter versions of the KIDSCREEN-52 health-related quality of life (HRQoL) questionnaire. **Methods:** We used the data from the Mandarin Chinese version of the KIDSCREEN-52 validation study conducted in Weifang, China (October–November 2016). Adolescents aged 11–17 years ( $n = 4385$ ) were surveyed, and subsamples ( $n = 841$ ) were re-tested. Statistical analyses on the feasibility, item and dimension properties, reliability, and validity of both versions were conducted. **Results:** The response rates of both surveys were more than 90%. In both versions, missing values and ceiling/floor effects were negligible; the measurement properties of items were also satisfactory. Internal consistency was robust with inter-item and item-total correlations (0.173–0.873; 0.422–0.786), Cronbach's alpha (0.847–0.881), Guttman's lambda-2 and -6 (0.807–0.889; 0.829–0.896), and McDonald's omegas (0.797–0.883). At both item and dimension levels, test–retest reliability was excellent, the intraclass correlation coefficients (ICCs) were (0.590–0.696; 0.785–0.842); the standard error of measurement (SEM) were (0.352–0.635; 1.018–1.949). Confirmatory factor analysis (CFA) confirmed the five and single-dimensional structure of both versions, and supported its configural and metric invariance across gender and age groups. Using KIDSCREEN-52 and PedsQL™4.0 as the verification instruments respectively, the correlation coefficients between the two short KIDSCREEN versions and their similar domains were strong which revealed a reasonably robust concurrency, convergence and discriminate validity. The known groups validity was supported by mainly medium to large effect sizes for differences according to the socioeconomic (Cohen's  $d = 0.554$ – $1.029$ ) and mental health status (Cohen's  $d = 0.732$ – $1.830$ ). **Conclusion:** The Mandarin Chinese versions of the KIDSCREEN-27 and KIDSCREEN-10 index demonstrated excellent psychometric properties. Both versions have good potential for measuring HRQoL in Chinese children and adolescents.

### (3209) Bullying and Health-Related Quality of Life in Colombian children and adolescents

Martha J Rodríguez, MSc, Universidad Santo Tomás, Bucaramanga, Colombia; Diana M Camargo, MSc, Universidad Industrial de Santander, Bucaramanga, Colombia; Sonia C Concha, PhD, Universidad Santo Tomás, Bucaramanga, Colombia; Ethman A Torres, MBA, Universidad Santo Tomás, Bucaramanga, Colombia

**Aims:** To establish the prevalence of bullying and its associated factors among school children and adolescents aged 8 to 18 years old in the city of Bucaramanga (Colombia). **Methods:** An observational cross-sectional study was conducted with 1332 children and adolescents who filled out the Colombian version of the Kidscreen-52. Both institutions and children were randomly selected, first by cluster sampling in thirty public and private schools and second, by simple random sampling. Being bullied was assessed by three items on the Kidscreen-52 “social acceptance” dimension. We took into account the following variables: sex, age groups (8–11 and 12–18 years), socio-economic conditions (low and high), functional limitation, and three Kidscreen-52 dimensions (“physical well-being,” “psychological well-being,” and “moods and emotions”). Item scores were summed up and converted into Rasch person parameter estimates that were transformed to T values with a mean of 50 and a Standard Deviation (SD) of 10. The resulting measures were stratified: students scoring one SD below the mean (score < 40) were defined as victims of bullying. Two logistic regression models were conducted

(female/male), the variables included were those with a  $p < 0.10$  in the bivariate analysis. A  $p < 0.05$  was considered statistically significant. Informed consents were obtained from the parents/caregivers of the participants. **Results:** The mean age was  $12.4 \pm 2.7$  years, 54.8% were female, 88.5% of the sample belonged to public schools, and 22.3% had a functional limitation. The prevalence of bullying in children (8–11 years) was 20.6%, and in adolescents (12–18 years) was 9.0% ( $p < 0.0001$ ). Having low scores on the “psychological well-being” dimension and presenting a functional limitation were associated with being bullied in males [OR 3.24 95% CI 1.70–6.15 and OR: 2.43 95% CI: 1.49–3.96, respectively]. Being 12 to 18 years was considered a protective factor for males and females [OR: 0.43 95% CI: 0.27–0.68 and OR: 0.22 95% CI 0.13–0.37, respectively]. **Conclusion:** The percentage of participants being bullied was 13.7% with variations between sex and age. It is necessary to take action to strengthen the reduction and prevention of bullying in Colombian schools.

### (3210) Sex and age differences in Health-Related Quality of Life among child and adolescent students from Bucaramanga (Colombia): a multilevel analysis

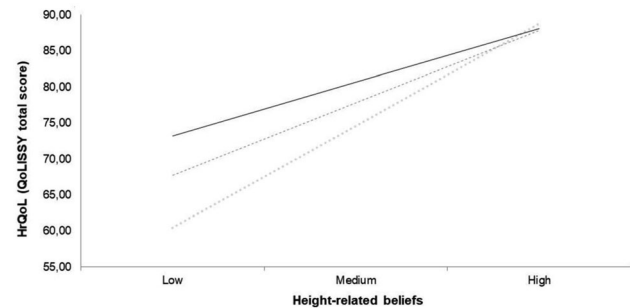
Martha J. Rodríguez, MSc, Universidad Santo Tomás, Bucaramanga, Colombia; Sonia C Concha, PhD, Universidad Santo Tomás, Bucaramanga, Colombia; Diana M Camargo, MSc, Universidad Industrial de Santander, Bucaramanga, Colombia

**Aims:** To assess sex and age differences in Health-Related Quality of Life (HRQL) in child and adolescent students using the Colombian version of the Kidscreen-52. **Methods:** A cross-sectional study was conducted with a population of 1334 children and adolescents from third to high school graduation who were selected by cluster sampling from thirty public and private schools from Bucaramanga (Colombia). All of them filled out the Colombian version of Kidscreen-52. General HRQOL and ten dimensions of Kidscreen-52 were evaluated: “physical well-being,” “physiological well-being,” “moods and emotions,” “self-perception,” “autonomy,” “financial resources,” “parent relations and home life,” “peers and social support,” “school environment,” and “social acceptance.” A multilevel linear regression analysis was used taking into account two levels: (1) children and adolescents (sex, age, functional limitation, and socio-economic condition), and (2) schools and neighborhoods. Only variables with a  $p \leq 0.10$  were included in the model. Informed consents were obtained from the parents/caregivers of the participants. **Results:** The mean age was  $12.3 \pm 2.7$  years, 54.8% were women, 88.5% of the sample belonged to public schools, and 22.3% had a functional limitation. When comparing sex and age groups (8–11 vs. 12–18 years), we found that boys and children (8–11 years) exhibited better HRQOL scores ( $p < 0.0001$ ). Fixed multilevel model showed that general HRQL and five dimensions (“physical well-being,” “psychological well-being,” “moods and emotions,” “self-perception,” and “parent relations and home life”) were statistically significantly associated with sex, age and functional limitation. In addition, we observed that the HRQL score registered 7.7 points more in men, 2.7 points less for each year of life, and about 13.0 points less in students who reported some functional limitation. In contrast, four dimensions (“autonomy,” “financial resources,” “peers and social support,” and “social acceptance”) were statistically significantly associated with sex (increasing its score) and functional limitation (decreasing its score). The “school environment” dimension was statistically significantly associated with age and functional limitation increasing its score. **Conclusion:** Sex and age differences occurred for almost all Kidscreen-52 dimensions; these findings are consistent with others found in the literature. It is necessary to implement strategies to improve health and well-being of children and adolescents.

### (3211) Body height or height-related cognitive beliefs: What matters the most for the health-related quality of life of children and adolescents with short stature?

Beatriz Mergulhão, M.Sc., Center for Research in Neuropsychology and Cognitive and Behavioral Intervention (CINEICC), Faculty of Psychology and Education Sciences of the University of Coimbra, Coimbra, Portugal; José Paulo Almeida, Ph.D., Pediatric Endocrinology Unit; Integrated Pediatric Hospital, São João Hospital Center, Oporto, Portugal; Helena Moreira, Ph.D., Center for Research in Neuropsychology and Cognitive and Behavioral Intervention (CINEICC), Faculty of Psychology and Education Sciences of the University of Coimbra, Coimbra, Portugal; Cíntia Castro-Correia, Ph.D., Pediatric Endocrinology Unit; Integrated Pediatric Hospital, São João Hospital Center, Oporto, Portugal; Monika Bullinger, Ph.D., Department of Medical Psychology, Center for Psychosocial Medicine, University Medical Center Hamburg-Eppendorf (UKE), Hamburg, Germany; Maria Cristina Canavaro, Ph.D., Center for Research in Neuropsychology and Cognitive and Behavioral Intervention (CINEICC), Faculty of Psychology and Education Sciences of the University of Coimbra, Hamburg, Portugal; Neuza da Silva, Ph.D., Center for Research in Neuropsychology and Cognitive and Behavioral Intervention (CINEICC), Faculty of Psychology and Education Sciences of the University of Coimbra, Coimbra, Portugal

**Aims:** This study aimed to (1) compare coping, height-related beliefs, social support, and health-related quality of life (HrQoL) between children/adolescents with short stature (SS) across different clinical characteristics (i.e., diagnosis, current height-deviation, and growth hormone [GH] treatment); (2) examine the associations between psychosocial variables (coping, height-related beliefs, and social support) and the pediatric HrQoL; and (3) test the moderating role of clinical variables on these associations. **Methods:** The participants were 114 Portuguese children/adolescents with SS (height-deviation  $\geq -2SD$  from population norms), aged 8 to 18 years ( $M \pm SD = 12.85 \pm 2.56$ ; 62.3% male), recruited by convenience at the Pediatric Hospital of the São João Hospital Center. The participants completed the Quality of Life in Short Stature Youth (QoLISSY) questionnaire and the Satisfaction with Social Support Scale. Clinical data were collected from the physicians and height deviation was computed using the Growth Charts UK-WHO application. **Results:** 46.5% of the children/adolescents were diagnosed with idiopathic short stature (ISS), 28.9% with growth hormone deficiency (GHD), and 24.6% had other diagnoses including genetic syndromes; 64.9% were not treated with GH; and 48.2% had reached normal height by the time of assessment. Univariate analyses of covariance controlling for sex and age showed no significant differences across diagnoses or treatment status in HrQoL, coping, beliefs or social support. However, children/adolescents who had reached normal height presented better HrQoL ( $F = 8.89, p = 0.004$ ) and lower coping scores ( $F = 6.09, p = 0.015$ ) than those with current SS. Regression analysis explained 54.7% of variance in HrQoL, with significant main effects of current height deviation ( $\beta = -0.28; t = -2.88, p = 0.005$ ) and height-related beliefs ( $\beta = 0.69; t = 9.45, p < 0.001$ ). Additionally, moderation analyses revealed a significant interaction effect between beliefs and diagnoses, with a stronger association for children/adolescents with ISS (Fig 1). **Conclusion:** The results suggest that reaching normal height and positive height-related beliefs play a key role in the HrQoL of children/adolescents with SS. Psychological interventions focused on the modification of height-related beliefs may even be more important to improve the HrQoL of children with ISS, as the unclear etiology of SS may hinder medical treatments.



#### Legend:

ISS:  $b \pm SE = 0.59 \pm 0.08; t = 7.77, p < 0.001$

GHD:  $b \pm SE = 0.42 \pm 0.09; t = 4.82, p < 0.001$

Other diagnoses:  $b \pm SE = 0.31 \pm 0.10; t = 3.17, p = 0.002$

### (3213) Reliability of change for the AQOL-MHS in the context of disaster recovery

Ligia Chavez, Ph.D., University of Puerto Rico, San Juan, Puerto Rico; Patrick Shrout, Ph.D., New York University, New York, New York, United States; Pedro Garcia, M.A., University of Puerto Rico, San Juan, Puerto Rico; Keilyn Vale, M.S., University of Puerto Rico, San Juan, Puerto Rico

**Aims:** The current study furthers psychometric work to track changes over time in HRQOL during the course of treatment in clinical samples of Latino adolescents aged 12-18 years, for the Adolescent Quality of Life Mental Health Scale (AQOL-MHS). This study's data collection started early 2018; a few months after a devastating hurricane caused considerable destruction to the Island of Puerto Rico. These data are unique because they capture the longitudinal impact and eventual recovery after a natural disaster in a mental health services sample. **Methods:** Our work tracks changes over 4 waves of data, 3 months apart for 227 adolescents. Data collection spans over a 2 year period. All participants were receiving services at baseline assessment and were tracked for follow-up appointments regardless of treatment status. We analyze conventional reliability statistics for individual differences (e.g., Cronbach's alpha) and we conducted a variance decomposition analysis to estimate the reliability of change. **Results:** Psychometric analyses from prior work were replicated with comparable results. A Generalizability Theory (GT) analysis revealed that the AQOL-MHS domains had moderate reliability estimates that varied from .54 to .66. Although there was reliable change at the individual level, on the average the AQOL-MHS means decreased slightly over time. **Conclusion:** The reliability of change for all three AQOL-MHS scales (Emotional Regulation, Self-Concept and Social Context) was acceptable. Recovery post-hurricane was unevenly distributed within our sampled population, with participants conceivably experiencing greater fluctuations in QOL than usual. The AQOL-MHS successfully differentiates between person changes, but the exceptional circumstances during the aftermath of the disaster could have affected the reliability of within person change.

### 3214 Offspring benefited from maternal mindfulness training during pregnancy

Natalia Kiseleva, Ural Federal University, Ekaterinburg, Russia; Sergey Kiselev, Ph.D., Ural Federal University, Ekaterinburg, Russia

**Aims:** There is evidence that maternal anxiety during pregnancy affects child outcomes. However, there is lack of studies that have evaluated the effects of maternal psychosocial factors during pregnancy on child cognitive outcomes. Can the maternal mindfulness

training during pregnancy influence neurocognitive development of offspring? This study evaluates the effect of the maternal mindfulness training during pregnancy on cognitive development in 5 years old children. Methods: In the current study we included 16 women who participated in the maternal mindfulness training during pregnancy. Women were between 14 and 20 weeks gestation. Women were eligible to participate if they were willing and able to attend the six-week mindfulness course. Participants were trained in the practice of mindfulness meditation and its applications to daily life through participation in instructor-led group meditations, lectures about mindfulness practices and discussions. The control group included 16 women who did not participate in this training during pregnancy. When the offspring of the target pregnancies were 5 years of age ( $M = 5.12$  years,  $SD = 0.34$ , 10 boys and 6 girls), their cognitive development was assessed by Lurias's child neuropsychological assessment battery that enables to assess five functional domains, including executive abilities, language, memory, sensorimotor, and visuospatial abilities. Results: One-way ANOVA was used to reveal group differences in performing subtests from five functional domains of Luria's battery. We have revealed that children from the experimental group performed significantly ( $p \leq 0.05$ ) better on subtests from executive functional domains. Conclusion: We have shown that preschool children whose mothers participated in mindfulness training during pregnancy had better level of executive abilities in comparison to children from control group. These results suggest that maternal mindfulness training during pregnancy may have positive effect on neurocognitive development of children, particularly on the development of executive abilities during preschool age period. However, we need to do further research for revealing the effect mindfulness training on neurocognitive development of children. Particularly, we are going to continue the investigation of children from our experimental and control group at the 6 years of age.

#### **(3215) Children with computer game benefited from visuospatial training**

Natalia Kiseleva, Ural Federal University, Ekaterinburg, Russia;  
Sergey Kiselev, Ph.D., Ural Federal University, Ekaterinburg, Russia

Aims: Children with computer game addiction have a risk for delay in development of cognitive functions. There are evidences that "digital environment" has negative effect on the development of visuospatial functions. The goal of this study was to reveal the effect of visuospatial training on visuospatial functions in children with computer game addiction. Methods: We used questionnaire for parents to reveal children with computer game addiction. The participants were 24 children at the age of 7 years ( $M = 7.12$  years,  $SD = 0.25$ , 20 boys and 4 girls) with computer game addiction. Children were included and randomly assigned to training conditions according to a  $2 \times 2$  cross-over design. We compared the efficacy of two methods of training (visuospatial training vs. conventional motor exercises). Children from intervention group participated in 36 weeks of visuospatial training. Training included different visuospatial exercises both on motor and cognitive level. We used subtests from NEPSY which are designed for assessing visuospatial functions (Arrows, Block Construction, Design Copying, Route Finding). Effects of training were analyzed by means of an ANOVA for repeated measurements. Results: The ANOVA has revealed ( $p < .05$ ) that for two subtests (Block Construction, Design Copying) the visuospatial training was superior to the conventional motor exercises, with effect sizes in the medium-to-high range (0.61–0.80). Conclusion: The findings from this study suggest that visuospatial training has positive effect on visuospatial functions in children with computer game addiction. However, we need to do further investigations to prove the effectiveness of visuospatial training for children with this addiction.

We are going to do longitudinal research for revealing the long-term effect of visuospatial training on visuospatial functions in children with computer game addiction.

#### **(3216) Motor sequencing training has positive effect on executive functions in a child with cerebral palsy**

Khilola Mashrabava, Fergana State University, Fergana, Uzbekistan;  
Sergey Kiselev, Ph.D., Ural Federal University, Ekaterinburg, Russia

Aims: It is known that cerebral palsy has negative effect on the development of neurocognitive abilities and executive functions in children. It is important to receive the evidence for efficiency of different trainings that are aimed to help children with cerebral palsy. The goal of this study was to assess the impact of 12 weeks of motor sequencing training on the executive abilities in a child with cerebral palsy. Methods: The child was 8-year-old boy. We used NEPSY and qualitative neuropsychological assessments in the framework of Luria's syndrome analysis for assessing executive abilities. The neuropsychological assessment of child has revealed the deficit of executive abilities. He participated in 12 weeks of motor sequencing training. A total of 36 therapy sessions lasting 50 min were performed. This therapy trained the child to plan, sequence and process information more effectively through repetition of goal-directed movements. This training is built on the conceptual framework derived from the work of Luria's theory of restoration of neurocognitive functions (Luria, 1963, 1974). Results: Neuropsychological assessment (NEPSY) of child after the intervention period has revealed apparent progress in performance of 4 subtests which are designed to assess executive abilities and attention (Tower, Auditory Attention and Response Set, Visual Attention, Statue). The qualitative neuropsychological assessment in the framework of Luria's syndrome analysis revealed the improve in the third functional unit of the brain. Particularly, child has demonstrated the decline in the impulsivity and distractibility, improvement in the sustained attention. Conclusion: According to result of this case report it can be assumed that motor sequencing training can be used as a prospective treatment approach for development of executive abilities in children with cerebral palsy. We are going to prove the effectiveness of this approach performing the research for group of children with cerebral palsy.

### **Research in Children and Adolescents II**

#### **(3217) The use of generic and disease-specific patient-reported outcome measures in daily clinical practice of a pediatric nephrology unit**

Lorynn Teela, Psychosocial Department Emma Children's Hospital, Amsterdam UMC, Amsterdam, Netherlands; Floor Veltkamp, MD, Department of Pediatric Nephrology Emma Children's Hospital, Amsterdam UMC, Amsterdam, Netherlands; Antonia Bouts, PhD, Department of Pediatric Nephrology Emma Children's Hospital, Amsterdam UMC, Amsterdam, Netherlands; Lotte Haverman, PhD, Psychosocial Department Emma Children's Hospital, Amsterdam UMC, Amsterdam, Netherlands

Aims: Research showed that Health-Related Quality of Life (HRQoL) is decreased in pediatric patients with chronic kidney disease (CKD), even in patients who received kidney transplantation. To identify, monitor and discuss HRQoL of patients in daily care, the KLIK PROM portal was implemented in the Emma Children's Hospital in 2011. Patients and/or their parents complete patient-reported outcome measures (PROMs) prior to the outpatient consultation. The answers are converted into a KLIK PROFILE, which is discussed by the

pediatrician during consultation. The aim of this study is to demonstrate the implementation of the KLIK PROM portal in the department of pediatric nephrology and to study (1) KLIK HRQoL data of CKD patients compared to the general population and (2) KLIK HRQoL data of kidney transplant recipients before and after kidney transplantation. Methods: CKD patients and their parents were invited to complete PROMs via the KLIK PROM portal prior to the outpatient consultation. Generic HRQoL was measured with the Pediatric Quality of Life Inventory for Children (PedsQL) or TNO-AZL Preschool children Quality Of Life (TAPQOL). Differences in (1) HRQoL scores of the first completed PedsQL were compared to the general population using ANCOVA and (2) HRQoL differences before and 1 year after transplantation were calculated using paired sample T-tests. Results: A total of 138 patients were invited to complete PROMs, of which 104 (75%) patients completed at least one PROM. Data from 73 patients (70%) who gave informed consent were used for analysis. Overall, CKD patients scored significantly lower than the general population on HRQoL (scales: total, physical, and school functioning). After transplantation, patients ( $n = 11$ ) scored better on the overall PedsQL score ( $p = 0.02$ ) compared to before transplantation, but not for the overall TAPQOL score ( $p = 0.07$ ). Conclusion: The KLIK PROM portal was successfully implemented in daily clinical care: 75% of invited patients completed at least one PROM before consultation. Results show that HRQoL in children with CKD is lower than in the general population. Improvement in HRQoL is shown after kidney transplantation, however the number of patients was small and differences in improvement based on different PROMs need to be studied further.

### (3218) Translation and Linguistic Validation of the University of Washington Caregiver Stress and Benefit Scales

Dagmar Amtmann, PhD, University of Washington, Seattle, Washington, United States; Alyssa Bamer, MPH, University of Washington, Seattle, Washington, United States; Rana Salem, MA, University of Washington, Seattle, Washington, United States; Dana Wilkie, University of Washington, Seattle, Washington, United States; Arnold Gammaioni, PhD, Zogenix, Inc., Emeryville, California, United States; Braley Galer, MD, Zogenix, Inc., Emeryville, California, United States; Mark Jensen, PhD, University of Washington, Seattle, Washington, United States

Aims: The University of Washington Caregiver Stress (UW-CSS) and Benefit (UW-CBS) Scales were developed in the United States (US) to measure impact on caregivers of caring for a child or children, including children with health conditions. The scales were translated to German, Spanish, Italian and French. This study examined whether translated versions functioned similarly to the English version. Methods: Cognitive interviews were completed with caregivers of children < 18 years with a health condition. The translated versions were also administered to at least 100 caregivers in each of the four countries. The US development sample of 722 caregivers was used as a comparison population for differential item functioning (DIF) analyses. DIF was assessed by each country individually (e.g., US vs Spain) as well as by the combined sample (i.e., US vs Europe) using lordif with an R2 criterion of 0.02. DIF adjusted scores were calculated to determine impact of DIF. Results: Interviews were completed with 45 caregivers (German  $n = 12$ ; Spanish  $n = 10$ ; French  $n = 13$ ; Italian  $n = 10$ ). UW-CSS and UW-CBS were administered to 456 (Germany  $n = 117$ , Spain  $n = 114$ , France  $n = 115$ , Italy  $n = 110$ ) caregivers of children with or without specific health care needs. All stress items functioned well in cognitive interviews and three of the 19 exhibited statistically significant DIF by multiple countries and in the overall sample, requiring minimal modifications. Four of the 13 benefit items required modifications based on cognitive interview

feedback and six items displayed DIF in one or more countries or in the combined sample. Average differences between DIF adjusted and non-adjusted scores were minimal, < 1 point on the T-score metric for both scales and for all comparisons. Conclusion: Published short forms were modified to minimize the impact of DIF on the UW-CSS and UW-CBS T-scores and to reflect feedback from cognitive interviews. Version 2 short forms function well in all four of the translated versions. All language versions are available at <https://uwcorr.washington.edu/measures/>.

### (3219) Patient-reported outcome measures in modern cystic fibrosis population

Rasa Ruseckaite, PhD, Monash University, Melbourne, Australia; Irushi Ratnayake, Monash University, Melbourne, Australia; Susannah Ahern, Monash University, Melbourne, Australia

Aims: The Australian Cystic Fibrosis Data Registry (ACFDR) collects clinical data of > 3500 patients diagnosed with cystic fibrosis (CF) attending specialist clinics; however, it does not capture health-related quality of life (HRQOL). Measuring HRQOL using patient-reported outcome measures (PROMs) integrated into the ACFDR would reinforce the patient voice in data collection and also enable researchers and clinicians to explore overall health and wellbeing of individuals with CF. The aim of this study was to determine suitability and acceptability of the existing CF-specific PROMs for incorporation into the ACFDR. Methods: Semi-structured qualitative interviews were conducted with patients or caregivers of children diagnosed with CF and their managing clinicians. Prior to the interviews, participants were emailed copies of two most frequently used CF-specific instruments: Cystic Fibrosis Questionnaire-Revised (CFQ-R) and Cystic Fibrosis Quality of Life (CFQoL) Questionnaire. Interview topics covered content and face validity, appropriateness, and acceptability to determine whether the instruments were suitable and useful, and whether they could be incorporated in the registry. Results: The number of participants included five adult patients, seven caregivers and thirteen clinicians. The majority of participants from all groups indicated that both instruments were comprehensive, clear and “easy to read and understand.” Although patients and caregivers felt the length of the instruments were acceptable, some clinicians expressed that both instruments were long and would result in a higher administrative burden. Caregivers and pediatric clinicians preferred the CFQ-R, as it was more appropriate for children and more acceptable and non-confronting. All participants felt, in comparison to the CFQoL, that this instrument would be appropriate for “people of varying competency levels.” Therefore, CFQ-R was chosen as the preferred instrument for a pilot study in the ACFDR. The study is currently under-way. Conclusion: Integration of PROMs into the ACFDR is necessary, as patients’ experiences of everyday functioning are not captured by physiological parameters and clinician-observed outcomes. PROMs in the ACFDR have the potential to be used in economic evaluations, to guide health policy decisions and to inform quality improvement for clinicians and health services.

### (3220) Visuospatial training improved the visuospatial abilities in child with cerebral palsy

Khilola Mashrabaeva, Fergana State University, Fergana, Uzbekistan; Sergey Kiselev, Ph.D., Ural Federal University, Ekaterinburg, Russia

Aims: It is known that children with cerebral palsy have delay in the development of neurocognitive abilities. It is important to receive the evidence for efficiency of different trainings that are aimed to help

children with cerebral palsy. The goal of this study was to assess the impact of 16 weeks of visuospatial training on the visuospatial abilities in 6 years of age child with cerebral palsy. Methods: The child was 6 year old boy with cerebral palsy. We used NEPSY and qualitative neuropsychological assessments in the framework of Luria's syndrome analysis for assessing visuospatial abilities. The neuropsychological assessment of child has revealed the deficit in visuospatial abilities. Child participated in 16 weeks of visuospatial training. A total of 42 therapy sessions lasting 50 min were performed. This training trained the child to do different visuospatial exercises both on motor and cognitive level. This training is built on the conceptual framework derived from the work of Luria's theory of restoration of neurocognitive functions (Luria, 1963, 1974). Results: Neuropsychological assessment of child after the intervention period has revealed apparent progress in performance of 4 subtests from NEPSY which are designed to assess visuospatial functions (Arrows, Block Construction, Design Copying, Route Finding). The qualitative neuropsychological assessment in the framework of Luria's syndrome analysis has revealed the improvement in the brain mechanism responsible for visuospatial processing. Particularly, child has demonstrated the decline in the amount of mirror and topological mistakes in the Mental Rotation subtest, and improving in the ability for copying 3-dimensional figure. Conclusion: According to result of this case report it can be assumed that visuospatial training can be used as a prospective approach for development of visuospatial functions in children with cerebral palsy. However, we need to prove this result using this training for group of children with cerebral palsy who have deficit in visuospatial abilities.

#### **(3221) Early pubertal onset impacting quality of life in children and adolescents**

Yang Pu, School of Public Health and Management, Chongqing Medical University, Chongqing, China; Qiuling Shi, School of Public Health and Management, Chongqing Medical University, Chongqing, China; Hong Wang, School of Public Health and Management, Chongqing Medical University, Chongqing, China

Aims: To determine the onset of puberty its' impact on young children and adolescents. And to provide a basis for the intervention of the quality of life in puberty children in the future, in order to promote the good adaptation and healthy body and mind development of children. Methods: A stratified cluster sampling method was used to conduct a survey in a district in December 2017. The five physiological change items of Puberty Development Scale(PDS) were used to assess students' onset of puberty. Students scoring higher than the 75th percentile of individual development scores were defined as the early pubertal onset group. A 39-item Quality of Life Scale for Children in Puberty was used to evaluate the quality of life of the respondents. Results: Among 7234 students, 3762 (52.0%) were boys and 3472 (48.0%) were girls. The prevalence of children in early pubertal onset was 20.1%. The total score of quality of life on children in early pubertal onset ( $134.29 \pm 18.05$ ) was significantly lower than non-early pubertal onset children ( $143.36 \pm 18.41$ ) ( $p < 0.001$ ). Multiple linear regression model showed that the early pubertal onset was a risk factor for quality of life ( $\beta = -0.080$ ,  $p < 0.001$ ). Conclusion: The early pubertal onset negatively affected quality of life in children and adolescents. The early pubertal onset children's quality of life is lower than that of non-advanced children. In the intervention of the quality of life of children during adolescence development, attention should be focused on children with early pubertal onset.

#### **(3222) Evaluation of the health-related quality of life of Uruguayan children, with life-threatening conditions, treated in a palliative care unit**

María de los Ángeles Dallo Campos, Degree in Psychology, Department of Medical Psychology. Faculty of Medicine. University of the Republic of Uruguay, Montevideo, Uruguay; Dr. Juan José Dapuetto Prieto, MD, PhD., Professor of Medical Psychology, Faculty of Medicine. University of the Republic of Uruguay. UdelaR., Montevideo, Uruguay

Aims: Introduction: Little is known in Uruguay, and globally, about health-related quality of life (HRQOL) of infants with life-threatening conditions (LTC) undergoing palliative care and how to evaluate it. Aims: To evaluate the reliability of the PedsQL Infant Scales and the HRQOL of infants with LTC treated in a palliative care unit. Methods: This is a descriptive and cross-sectional study in a sample of infants with LTC liable to receive palliative care admitted in the pediatric hospital between January and July 2019. The PedsQL Infant Scales, culturally and linguistically adapted for Argentina, were filled out by the primary caregiver. Permission to use the instruments was granted by the Mapi Research Trust. The reliability of the instruments was analyzed through Cronbach's alpha coefficient. The questionnaires were processed according to the scoring manuals and descriptive statistics was used to study the mean, median and standard deviation (SD). Institutional approval and informed consents were obtained. Results: The sample consisted of 29 children; of them, 21 (72%) were 1 to 12 months old, and 8 (28%) were 13 to 24 months old. The Infant Scales for 1 month to 12 months showed a very good internal consistency, with Cronbach alpha coefficients over 0.8 in most of the subscales and the total scale. According to the mean scores the most affected dimensions were Social Functioning (mean = 62.7; SD: 36.3) and Physical Functioning (mean = 65.7; SD: 26.3). The best rated dimension was Emotional Functioning (mean = 76.5; SD: 24.0). The 13 to 24 months scale obtained a Cronbach's alpha coefficient over 0.7. The most affected dimension was Cognitive Functioning (mean = 48.; SD: 31.3) and the best rated was Social Functioning (mean = 80.; SD: 23.1). Conclusion: The Peds QL Infant Scales showed good internal consistency when rated by primary caregivers of children in palliative care. The most affected dimensions were compatible with the impairments expected in this group of patients. In spite of the study limitations (small sample, only center study) we may conclude that it is feasible to continue using this instrument to assess interventions in children with LTC in palliative care.

#### **(3223) Modification of the Clinical Global Impression (CGI) Scale for Utilization in a Clinical Trial of Individuals with Angelman Syndrome**

Jennifer Panagoulas, RAC, GeneTx Biotherapeutics, Sarasota, Florida, United States; Allyson Berent, DVM, GeneTx Biotherapeutics, Sarasota, Florida, United States; Jill Mayhew, Ultragenyx Pharmaceutical, Inc., Novato, California, United States; Scott Stromatt, MD, GeneTx Biotherapeutics, Sarasota, Florida, United States

Aims: The CGI scale is a standardized assessment tool developed in 1976 for use in clinical trials to provide an assessment of an experienced clinician's interpretation of a patient's global functioning prior to and after initiating a study medication, including the impact of symptoms on the patient's function. The brevity and simplicity of

the tool has led to its widespread use in clinical trials for in many indications, including depression, anxiety, bipolar disorder, schizophrenia, autism, ADHD, Alzheimer's disease and, more recently, Angelman syndrome (AS). Methods: The CGI generally is made of 2 components. The first is the CGI-severity of illness (CGI-S) assessment of the current severity of the patient's symptoms rated on a 7-point scale that spans from "not at all impaired" to "among the most severely impaired." This is followed by the CGI-improvement (CGI-I), which rates change from the initiation of treatment. To increase the consistency and reliability of the CGI for evaluating patients with AS, disease-specific anchors were developed for rating severity and improvement on the 7-point scale. Assessment of the 5 most important functional domains, based on an AS specific disease concept model, were added. Results: A CGI-S-AS and CGI-I-AS were developed to allow global assessment of clinically meaningful domains; i.e., sleep, behavior, communication, gross motor function, and fine motor function. Specific questions in each of the 5 functional domains were added to obtain further detail on each domain to assess change over time. Inter-rater and test-retest reliability of this modified instrument will be determined and input from the FDA will be obtained. Conclusion: Modifications to the CGI to enhance its applicability for evaluation of baseline disease severity and improvement following treatment in individuals with AS have led to the use of this disease-specific tool in a first-in-human clinical trial in children with AS.

#### **(3224) Child with hemorrhagic stroke in the right fronto-parieto-temporal area benefited from visuospatial therapy**

Eleonora Mirzajonova, Fergana State University, Fergana, Uzbekistan; Sergey Kiselev, Ph.D., Ural Federal University, Ekaterinburg, Russia

Aims: It is known that stroke has a devastating power for the future of child. The aim of this study is to describe the visuospatial therapy findings of a child with hemorrhagic stroke in the right fronto-parieto-temporal area, showing the progress after 8 months of therapy. It is known that stroke has a devastating power for the future of child. The aim of this study is to describe the visuospatial therapy findings of a child with hemorrhagic stroke in the right fronto-parieto-temporal area, showing the progress after 8 months of therapy initiated early after acquired neurological injury. Methods: Boy of 6 years and 5 months suffered a sudden illness and was referred to the emergency hospital and diagnosed with hemorrhagic stroke in the right fronto-parieto-temporal area. Surgical procedures were performed. At the time of hospital discharge, there was guidance about the need for therapy care. Neuropsychological assessment revealed the severe deficit in visuospatial abilities in this child. A total of 60 visuospatial therapy sessions lasting 30 min were performed for 8 months. This therapy trained the child to do different visuospatial exercises both on motor and cognitive level. To assess the visuospatial abilities in child we used NEPSY and qualitative neuropsychological assessment in the framework of Luria's syndrome analysis. Results: Neuropsychological assessment of child after the intervention period has revealed apparent progress in performance of 4 subtests from NEPSY which are designed to assess visuospatial functions (Arrows, Block Construction, Design Copying, Route Finding). The qualitative neuropsychological assessment in the framework of Luria's syndrome analysis has also revealed the improve in the brain mechanism responsible for visuospatial processing. Conclusion: According to result of this case report it can be assumed that visuospatial therapy can be used as a prospective treatment approach for children with stroke in the right fronto-parieto-temporal area. However, we need to do further research for revealing the impact of visuospatial therapy on

children with stroke, specifically in the right fronto-parieto-temporal area.

#### **(3225) Child with Rolandic epilepsy benefited from motor sequencing training**

Shoxista Mamajanova, Fergana State University, Fergana, Uzbekistan; Sergey Kiselev, Ph.D., Ural Federal University, Ekaterinburg, Russia

Aims: It is known that children with epilepsy can have deficit in neurocognitive abilities. It is of great significance to receive the evidence for efficiency of different treatments that are aimed to help children with epilepsy. The goal of this study was to assess the impact of 12 weeks of motor sequencing training on a child with Rolandic epilepsy who had deficit in the executive abilities. Methods: We used NEPSY and qualitative neuropsychological assessments in the framework of Luria's syndrome analysis to assess executive abilities. The neuropsychological assessment of child has revealed the mild deficit in executive abilities. Child participated in 12 weeks of motor sequencing training. A total of 36 therapy sessions lasting 30 min were performed. This therapy trained the child to plan, sequence and process information more effectively through repetition of goal-directed movements. This training is built on the conceptual framework derived from the work of Luria's theory of restoration of neurocognitive functions (Luria, 1963, 1974). Results: After intervention period NEPSY has revealed apparent progress in performance of 4 subtests which are designed to assess executive abilities and attention (Tower, Auditory Attention and Response Set, Visual Attention, Statue). The qualitative neuropsychological assessment in the framework of Luria's syndrome analysis revealed the improvement in the third functional unit of the brain which is responsible for executive abilities according to the Luria approach (Luria, 1974). We have revealed also the decline in the impulsivity and distractibility in child after intervention period. Conclusion: According to result of this case report it can be assumed that motor sequencing training can be used as a prospective treatment approach for improvement of executive abilities in children with Rolandic epilepsy. To prove the efficiency of this approach we are going to do research using this therapy for group of children with Rolandic epilepsy.

#### **Theory & Policy I**

#### **(3226) Clinical Outcome Assessments (COA) measuring physical functioning and quality of life at the forefront by regulatory agencies in drug labels**

Céline Desvignes-Gleizes, MS, PhD, Mapi Research Trust, Lyon, France; Delphine Jurado, MSc, Mapi Research Trust, Lyon, France; Sandra Desruol-Allardin, MSc, Mapi Research Trust, Lyon, France; Nadine Kraft, MSc, Mapi Research Trust, Lyon, France; Sonia Bothorel, MSc, Exec MBA, Mapi Research Trust, Lyon, France; Laure-Lou Perrier, MSc, Mapi Research Trust, Lyon, France

Aims: The release of an updated version of the COA compendium in August 2019 by FDA shows the growing recognition of the importance of these tools in the drug development process. In order to support COA endpoint strategies in clinical trial context, we aimed at identifying the COA and their related concepts, receiving a claim issued by EMA or FDA for the first time, over the last 3 years. Methods: The ePROVIDE™ platform hosts 3 COA-focused databases including PROLABELS™ a database reporting COA claims for drugs approved by FDA and EMA. A search was performed to retrieve COA with a claim for the last 3 years, then refined to those

receiving a claim for the first time. Results: 84 COA receiving a claim for the first time were identified: 73 granted by FDA and 51 by EMA. Out of the 84 COA, 61% were Patient-Reported Outcomes (PRO) and 26% Clinical-Reported Outcomes. The main concepts covered by the identified COA were: Signs and symptoms: 20 COA for FDA and 17 for EMA (including Psoriasis Signs and Symptoms Diary, COPD Assessment Test for both agencies)Physical/motor functioning: 10 COA for FDA (including Foot Function Index, Migraine Physical Function Impact Diary) and 8 COA for EMA (including Inhibitor-Specific QOL with Aspects of Caregiver Burden-physical health subscale)Quality of life (QOL): 14 COA for FDA (including Haemophilia QOL Questionnaire for Adults, QOL in Epilepsy Inventory-31) and 7 COA for EMA (including FACT—Melanoma, Individualized Neuromuscular QOL Questionnaire)Drugs developed in 15 rare diseases (including hemophilia A, amyloidosis, hypophosphatemic rickets and Cutaneous T-cell lymphoma) received COA claims, mainly focusing on the measure of physical functioning (ex: Revised Upper Limb Module, Hammersmith Functional Motor Scale Expanded, Quantitative Myasthenia Gravis Score, Haemophilia QOL Questionnaire for Children-physical health subscale). Conclusion: These results confirm that FDA and EMA granted an increasing number of COA with claims over the years and particularly PRO. More interestingly, the impact of disease on patient's physical functioning and QOL tends toward to be acknowledged by regulatory agencies as important concepts to measure in the development of drugs.

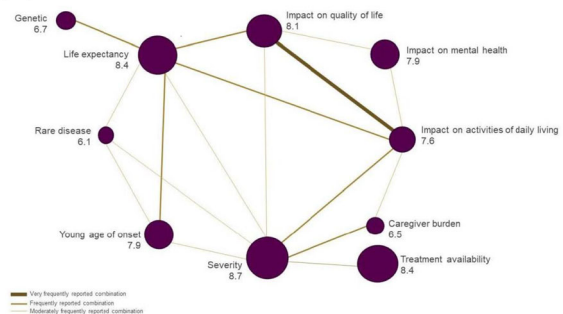
### (3227) Societal perspectives on the importance of disease and treatment attributes: A qualitative study from the United States

Ivana F Audhya, Sarepta Therapeutics, Cambridge, Massachusetts, United States; Shelagh M Szabo, Broadstreet Health Economics and Outcomes Research, Vancouver, British Columbia, Canada; David Feeny, McMaster University, Hamilton, Ontario, Canada; Daniel C Malone, The University of Utah, Salt Lake City, Utah, United States; Peter Neumann, Tufts School of Medicine, Boston, Massachusetts, United States; Meagan Harwood, Broadstreet Health Economics and Outcomes Rese, Boston, British Columbia, Canada; Katherine L Gooch, Sarepta Therapeutics, Cambridge, Massachusetts, United States

**Aims:** All disease and treatment attributes that society deems important should be considered within value frameworks evaluating costs and benefits of new therapies. Therefore, societal views on these attributes should guide health research priorities. However, little is known about the views of the general public about which attributes matter most. The aim of this study was to investigate the importance of attributes beyond health gain to the patient and cost to the healthcare system, from the perspective of the general public. **Methods:** Potentially important attributes to characterize health conditions were identified based on the ISPOR Special Task Force on Value Assessment, a literature review, and feedback from a convenience sample of eight members of the general public. A qualitative interview guide with visualizations, to elicit feedback on attributes, was developed and pilot tested. A sample of general public participants was recruited from Seattle, San Francisco, and Dallas, to reflect a balanced distribution in age, sex, and number of children living at home. Participants ranked attributes on a scale of 1 (not important) to 10 (very important) in terms of importance for future research; and commented on the perceived relationship between attributes. Interview transcripts were coded using NVivo for thematic analysis. **Results:** Thirty-three participants were included (mean [range] age, 49.8 [26–71] years, 48.5% male, 33.3% with children < 18 years at home). Of the attributes evaluated, disease severity (mean ranking,

8.7); treatment availability (8.7); impact on life expectancy (8.4), quality-of-life (8.1), or mental health (7.9); and young age of onset (7.9) were ranked most highly (Figure). Some novel elements were also identified, including whether diseases were externally visible, lifetime burden/disability, and impact on the family. **Conclusion:** Attributes including disease severity, impact on life expectancy, and treatment availability were all highly ranked by members of the general public in terms of their importance for guiding research into diseases and their treatment. Findings from this study uncover attributes which may be useful to more explicitly consider within evolving frameworks for assessing the costs and benefits of new therapies.

**Figure:** Reported importance of and relationship between attributes, based on feedback from qualitative interview participants (n=33)



The size of the node indicates attribute importance (mean ranking provided in node label), and lines between nodes indicate that 25% of respondents reported a relationship between two attributes. The length of the line does not provide details on strength of relationships.

**NOTE:** Only pre-specified disease and treatment attributes are represented in the figure; data on novel attributes (e.g. a disease being externally visible, lifetime burden/disability, and impact on the family) were collected but not displayed.

### (3228) A revised operational model of response shift: examples from patients' perspectives

Nikki Ow, McGill University, Montreal, Quebec, Canada; Antoine Vanier, PhD, University of Nantes - University of Tours, Nantes, France; Frans Oort, PhD, University of Amsterdam, Amsterdam, Netherlands; Leah McClimans, PhD, University of South Carolina, Columbia, South Carolina, United States; Jan Böhnke, PhD, University of Dundee, Dundee, United Kingdom; Bernice G Gulek, University of Washington, Dundee, Washington, United States; Mirjam Sprangers, PhD, Amsterdam UMC, Amsterdam, Netherlands; Véronique Sebillé, PhD, University of Nantes—University of Tours, Nantes, France; Nancy E Mayo, PhD, McGill University, Montreal, Quebec, Canada; submitted on behalf of The Response Shift in Sync Group

**Aims:** Response shift (RS) occurs in the measurement of patient-reported outcomes (PRO) when circumstances arise over time that make people change their evaluation of the underlying construct. A revised operational model of RS has been proposed by Vanier et al. (2020). This model depicts the relationships between characteristics of the person and their environment (antecedents), the catalytic experience (catalyst) induces the person to adapt, cope, or learn new ways of being (mechanisms), with the target construct and its operational measurement. The purpose of this qualitative synthesis was to provide 'real-life' examples for the presence of these processes as expressed by people living through them. **Methods:** A systematic search of three databases was carried combining relevant RS like recovery, adaptation and adjustment, and qualitative keywords. For studies to be included, the content had to involve a change in the perception of living with a health condition. Quotes were mapped onto the different processes of the RS model. **Results:** 33 studies were included. Table 1 shows the process in the model with accompanying quotes. The quotes illustrated that RS is not compartmentalized and sequential but consists of a

myriad of processes occurring simultaneously (Table 1). As illustrated by the quotes, the catalyst is often the consequence of a health condition, side effects of treatment, or unexpected functional deterioration. Catalysts have a known effect on the target construct and its score. Its effects can be modified positively or negatively through mechanisms of (mal)adaptation, growth, coping and learning. People can experience more than one mechanism over time because of their new circumstances. To describe how RS occurs, one must be asked about the situation at Time 2 with respect to Time 1. People clearly invoke the 3Rs (recalibration, reprioritization, and reconceptualization) when narrating their experiences over time. The quotes also reflected the need for homeostasis and to not deviate too much from personal set point, supporting theories of why RS occurs. Conclusion: When interpreting PROs, it is important to recognize these RS processes are real and can account for measurement invariance over time.

Quotes	
	<b>Antecedents</b>
<b>Personal factors</b>	I always faced challenges head on. And so I would face this one the same way. Determination has always been part of who I am, and adversity seemed somehow to make me stronger <sup>1</sup>
	<b>Catalyst</b>
<b>Catalyst</b>	I think that's the only, probably the only real loss that I couldn't do anything about immediately because I had already experienced a loss that made it real to me in February because I look both my breasts off. "I never thought I'd look so bad. My eyes became red. My hair fell out. I'd grab pieces of hair. When all your hair goes, you lose your masculinity." <sup>2</sup>
	<b>Adaptation</b>
<b>Adaptation - positive</b>	I may just say that you simply have to readjust, and live in a completely different way, you have to constantly keep in mind that you can't do too much and you always have to adapt yourself in relation to time. You simply have to readjust and rethink in life. <sup>3</sup>
<b>Adaptation - maladaptive</b>	The frustration – and I've always done these jobs and now I can't. I do get irritable sometimes, I'm too ambitious. I'd like to be my old self. I get dependent when you see other people doing things you used to. <sup>4</sup>
<b>Adaptation - non-uniform</b>	I can explain it like this. In the beginning you feel totally relieved that the cancer has disappeared and actually you take everything that comes along as part of the deal. But in the end you start feeling the appearance of urinary leakage and diminished erection. <sup>5</sup>
<b>Post-traumatic growth</b>	While I don't recommend cancer as a life encounter, my experience is along with the devastation it wreaks there's a positive side to the experience at least for those of us with supportive family and friends. Being thankful for and appreciative of good family, good friends, and good doctors must be balanced alongside the associated hardships. <sup>6</sup>
<b>Transformative learning</b>	...transformative is a very good word ...almost a sea change in ... in the sense that it teaches you about yourself ... who you really are when everything else is stripped away and that is really a valuable experience ...and the additional richness that it's put into my life ...I've become much more mellow with life. <sup>7</sup>
	<b>Theories of RS</b>
	<b>Theory of how response shift occurs</b>
<b>3Rs</b>	"My health is unstable. When I come home from chemotherapy I feel like a '5'. I feel nauseated right after the treatment. But other than that, I do not feel sick. So I can not say my health is a '5'. On the other hand, it isn't a '7' either, because I only feel like a 7 in the last week before the next chemo. So I will opt for a '6', right in between." <sup>8</sup>
	<b>Theory of why response shift occurs</b>
<b>Cognitive Homeostasis/ Set point theory</b>	rather than think I wish I could be doing this, I wish I could be doing that. I don't, I just say right. I'm not going to be able to do this, I'll shrink it ...so cope that way really. <sup>9</sup> "My deficits remain, but I have learned to live relatively peaceful with them, integrating them into my being and managing to enjoy a life which is still rich and active. It is not my previous life. Most people live once; I have had the opportunity to live twice." <sup>10</sup>

1. Curtis R, Groarke A, McSharry J, Kerin M. Experience of breast cancer: Burden, benefit, or both? *Cancer Nursing*. 2014;37:E21-30
2. Parashar D. The trajectory of hope: Pathways to find meaning and reconstructing the self after a spinal cord injury. *Spinal Cord*. 2015;53:565-568
3. Ferrell BR, Grant MM, Funk B, Otis-Green S, Garcia N. Quality of life in breast cancer survivors as identified by focus groups. *Psycho-Oncology*. 1997;6:13-23
4. Baker F, Zabara J, Pollard A, Wingard J. Reintegration after bone marrow transplantation. *Cancer Practice*. 1999;7:190-197
5. Osman M, Jakobsson Ung E, Falk K. Continuity means "preserving a consistent whole"—a grounded theory study. *International journal of qualitative studies on health and well-being*. 2015;10:29872
6. Dowsell G, Lawler J, Dowsell T, Young J, Forster A, Hearn J. Investigating recovery from stroke: A qualitative study. *Journal of clinical nursing*. 2000;9:507-515
7. Leal I, Milbury K, Egebreton J, Matin S, Jonaseh E, Tanni N, et al. Intersection: A qualitative analysis of adjusting to living with renal cell carcinoma. *Palliative & supportive care*. 2018;16:146-154
8. Bickford J, Covency J, Baker J, Hersh D. Living with the altered self: A qualitative study of life after total laryngectomy. *International Journal of Speech-Language Pathology*. 2013;15:324-333
9. Taminias-Bloem EF, van Zanten FJ, Koemanen MA, Rapkin BD, Visser MR, Koning CC, et al. A 'short walk' is longer before radiotherapy than afterwards: A qualitative study questioning the baseline and follow-up design. *Health & Quality of Life Outcomes*. 2010;8:69
10. Appleton L, Flynn M. Searching for the new normal: Exploring the role of language and metaphors in becoming a cancer survivor. *European journal of oncology nursing* : the official journal of European Oncology Nursing Society. 2014;18:378-384
11. Beeken RJ, Eiser C, Dalley C. Health-related quality of life in haematopoietic stem cell transplant survivors: A qualitative study on the role of psychosocial variables and response shifts. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation*. 2011;20:153-160
12. Leveck WM, Kayes NM, Fady JK. Experience of recovery and outcome following traumatic brain injury: A metasynthesis of qualitative research. *Disability & Rehabilitation*. 2010;32:986-999
13. Korfafe IJ, Hak T, de Koning HJ, Essink-Bot M-L. Patients' perceptions of the side-effects of prostate cancer treatment—a qualitative interview study. *Social Science & Medicine*. 2006;63:911-919

### (3229) Capturing the patient experience in systemic lupus erythematosus: are widely used measures fit-for-purpose and adherent to FDA PRO guidance recommendations?

Kayleigh Majercak, MS, University of Maryland Baltimore, Baltimore, Maryland, United States; Ester Villalonga-Olives, PhD, MSc, University of Maryland Baltimore, Baltimore, Maryland, United States; Eleanor Peretto, PhD, MS, National Health Council, Washington, District Of Columbia, United States

**Aims:** The 2009 Food and Drug Administration (FDA) patient-reported outcome (PRO) guidance outlines characteristics of rigorous PRO-measure development. There are a number of widely used PRO measures for Systemic Lupus Erythematosus (SLE), but it is unknown how well the development processes of earlier or updated respective versions align with FDA guidance. The objective of this study was to

assess how well two widely used SLE PRO measures, the LupusQoL and LupusPRO, align with FDA guidance. **Methods:** LupusQoL and LupusPRO were selected as the most widely studied and used in the UK and US. Four versions were reviewed: LupusQoL (2007), LupusQoL-US (2010), LupusPROv1.7 (2012) and LupusPROv1.8 (2018). The methodological review utilized FDA guidance to synthesize evaluation criteria: target population, concepts measured, measurement properties, and documentation across three phases (i.e., item generation, content validity assessment, and other psychometric-property testing). Two reviewers abstracted data independently, compared results, and resolved discrepancies. A third reviewer served as a tie breaker. **Results:** The intended target-SLE populations were British adults (LupusQoL) and ethnically heterogeneous US adults (LupusPRO). LupusQoL assessed health-related quality of life (HRQoL) and LupusPRO assessed HRQoL and non-HRQoL constructs. For all measures, the target population remains unclear as population characteristics (e.g., ethnicity, education, disease severity, etc.) differed or were not consistently reported/not considered across the item generation, content validity, and other psychometric testing phases (e.g., LupusQoL item generation lacked male involvement, LupusPRO content validity population characteristics were not reported, revised measures' target population characteristics differed from original measures). The first phase of development, "item generation," was conducted with concepts elicited via patient engagement interviews until saturation and item derivation from experts. Content validity was assessed via patient feedback with limited item-tracking documentation; measure revisions assumed content validity. Other psychometric testing recommendations (internal consistency, test-retest reliability, construct validity, ability to detect change) were assessed for all measures, except for ability to detect change for revised measures. **Conclusion:** FDA guidance promotes rigorous PRO-measure development. Despite the developers' original efforts on establishing content validity and other measurement properties, there are important limitations in processes and documentation of the target population, thus, calling into question for which target population(s) the measures are fit for purpose.

### (3230) Understanding patient-reported outcomes in the asthma product development process in Japan: A review of labeled products

Bruce Crawford, Syneos Health, Tokyo, Japan; Yoko Sakai, Syneos Health, Tokyo, Japan; Ayumi Shimada, Syneos Health, Tokyo, Japan

**Aims:** With a growing interest in patient-centered approaches, companies look to the patient's voice to understand the benefit of products. In Japan, more than 500 products for asthma were approved or had their label updated in the past decade. This study aims to understand how PROs are used in labeling and physician and patient targeted materials for asthma products in Japan between 2010 and 2019. **Methods:** Asthma treatments approved or labels changed between 2010 and 2019 were identified using the Pharmaceutical and Medical Devices Agency (PMDA) website. From the same resource, labels, patient guidance forms (PFs), and interview forms (IFs) of relevant products were extracted. PFs and IFs were used as supplemental information for patients' and doctors' better understanding, respectively. **Results:** In total, 551 product labels with an indication for asthma were published between 2010 and 2019. After assessment of these labels, as well as their PFs and IFs, only one label and nine IFs were found to contain PROs, whereas no PROs were presented in PFs. Two of the IFs mentioned only "QoL was measured" or "QoL improvement," and did not provide further details such as the type of the PRO measure. In total, there were nine PROs evaluating QoL and/or asthma-related symptoms in the IFs. The Asthma Quality of Life Questionnaire was the most commonly reported PRO measure



( $n = 5$ ), followed by the Asthma Control Questionnaire ( $n = 3$ ). One IF contained a pediatric PRO, while the remaining PROs were collected from adult patients. Most PROs were incorporated as secondary outcomes in clinical trials. More than one PRO was described from several clinical trials in most of the IFs, but few of these PROs were from trials conducted in Japan. Conclusion: Over the last ten years, only 2% of launched asthma products in Japan contained PROs on their label and/or IFs. There was only one product that reported PRO results from pediatric patients identified domestically. There was a lack of consistency in PRO measurement usage for asthma treatments in Japan. In order to characterize the benefit of these drugs, the patient's perspective should be better incorporated in the pharmaceutical product development and regulatory lifecycle.

### **(3231) Identifying risk-adjustment variables to be included in patient-reported outcome-based quality assessments in renal replacement therapy**

Carsten Volland, M.A., Institute for Quality Assurance and Transparency in Healthcare (IQTIG), Berlin, Germany; Tobias Mertzig, Institute for Quality Assurance and Transparency in Healthcare (IQTIG), Berlin, Germany; Gregor Liegl, Dr., Institute for Quality Assurance and Transparency in Healthcare (IQTIG), Berlin, Germany; Julia Böttcher, Institute for Quality Assurance and Transparency in Healthcare (IQTIG), Berlin, Germany; Julia Ginkel, Institute for Quality Assurance and Transparency in Healthcare (IQTIG), Berlin, Germany; Christopher Kienle, Institute for Quality Assurance and Transparency in Healthcare (IQTIG), Berlin, Germany; Mandy Wagner, Dr., Institute for Quality Assurance and Transparency in Healthcare (IQTIG), Berlin, Germany; Konstanze Blatt, Dr., Institute for Quality Assurance and Transparency in Healthcare (IQTIG), Berlin, Germany

**Aims:** The Institute for Quality Assurance and Transparency in Healthcare (IQTIG) has been commissioned to develop patient-reported instruments for measuring quality of care in dialysis centers and renal transplant clinics in Germany. Next to processes of care, different physical and psychosocial patient-reported outcome (PRO) domains have been identified as quality-related aspects of renal-replacement therapy (RRT). Since PRO-based performance measures are potentially biased by patient-related factors, risk-adjustment is crucial for producing fair and valid comparisons between RRT facilities. Thus, this subproject aimed to identify potential risk-adjustment variables, which can be included in patient questionnaires for quality assessment of RRT in Germany. **Methods:** To develop a set of quality-related aspects of RRT, a systematic literature review was conducted, using MEDLINE, Embase, CINAHL and the Cochrane Library. The search and analysis followed a standardized stepwise procedure including the precise definition of criteria for inclusion and full-text extraction of relevant articles by two independent reviewers. Extracted quantitative studies and review articles were additionally analyzed for potential risk-adjustment-variables. **Results:** In sum, 6,095 abstracts were screened, 169 studies met the inclusion criteria of which 66 studies addressed potential risk-adjustment-variables. Ten variables potentially affecting PROs were identified as having the strongest influence: age, education, gender, type of treatment (e.g., home vs. in-center dialysis), duration of dialysis, comorbidities, BMI, family status, housing situation, distance to dialysis center, and employment status. Especially age and gender affected all relevant outcome domains (e.g., symptom burden as well as emotional and social functioning). Other variables, such as education, had mostly influence on emotional and social functioning. While some variables had a rather consistent effect on the outcomes, most variables showed varying effects in terms of effect size and

direction depending on study and/or subpopulation. **Conclusion:** RRT quality-related PRO domains are affected by several factors, which are not related to quality of care. However, final conclusions for RRT quality assessment in Germany are not possible. Thus, in a follow-up project, the ten aforementioned potential risk-adjustment variables will be analyzed in a survey with  $n \sim 1.000$  RRT patients to determine their effects on the specific PRO-based measures used for quality assessment in Germany.

### **(3232) Patients' perspectives of telephysiotherapy in a low resource setting**

Adesola Odole, Ph.D, Department of Physiotherapy, College of Medicine, University of Ibadan, Ibadan, Nigeria; Nse Odunaiya, Ph.D, Department of Physiotherapy, College of Medicine, University of Ibadan, Ibadan, Nigeria; Omobola Akinola, B.Physiotherapy, Department of Physiotherapy, College of Medicine, University of Ibadan, Ibadan, Nigeria; Michael Ogunlana, Ph.D, College of Health Sciences, School of Health Sciences, University of KwaZulu-Natal, Durban, South Africa; Chidozie Mbada, Ph.D, Department of Medical Rehabilitation, College of Health Sciences, Obafemi Awolowo University, Ile-Ife, Nigeria; Ezinne Ekediegwu, M.Sc., Department of Medical Rehabilitation, Faculty of Health Sciences and Technology, Nnamdi Azikwe University, Nnewi Campus, Ile-Ife, Nigeria; Aderonke Akinpelu, Ph.D, Department of Physiotherapy, College of Medicine, University of Ibadan, Ibadan, South Africa

**Aims:** Global trends in healthcare include the infusion of technology and a shift to patient-centered care. The integration of information and communication technology applications into physiotherapy practice globally is linked with its accessibility to patients. For effective integration of tele-physiotherapy in Nigeria, patients' perspectives must be explored. The aim of this study was to investigate the level of awareness, knowledge and perception of tele-physiotherapy among patients receiving physiotherapy in Nigeria. **Methods:** A convergent parallel mixed-method design of a cross-sectional survey (CSS) and a qualitative design were utilized. Conveniently sampled adult patients receiving physiotherapy in public hospitals in Ibadan, Nigeria participated in the CSS. Data for the CSS was collected using a face- and content-validated questionnaire and was analyzed with descriptive statistics. A focus group discussion (FGD) was used to obtain qualitative data using a phenomenology qualitative approach and data were analyzed content thematic analysis. **Results:** Participants for the CSS (78 (51.7%) males and 73 (48.3%) females) were aged  $47 \pm 4.39$  years. Only one patient was aware of the term "tele-physiotherapy," had good knowledge and a positive perception of tele-physiotherapy. The FGD further revealed a very low level of awareness, moderate level of knowledge and a positive perception of tele-physiotherapy. The FGD also explored the reasons behind the low level of awareness and knowledge and themes generated included erratic power supply, underfunding of the health sector, poor policy implementation, lack of proper implementation of tele-physiotherapy, lack of infrastructure and cultural acceptability issues. Solutions such as creating public awareness and education of stakeholders; implementation of policies and provision of facilities were proffered to ensure the feasibility of tele-physiotherapy in Nigeria. **Conclusion:** Patients who received physiotherapy in a low-resource setting (Nigeria) had a very low level of awareness, moderate level of knowledge and a positive perception of tele-physiotherapy. Advocacy for provision of infrastructure; constant power supply, internet facilities and improvement in ICT should be put in place to facilitate, increase the feasibility and effectiveness of tele-physiotherapy in Nigeria.

**(3233) : e-health tools for communication between migrants and health professionals: a systematic literature review**

Frédérique Thonon, PhD, PRO team INSERM 1123, Paris, France; Swati Perrot, PRO team INSERM 1123, Paris, France; Abhijna Vithal Yergolkar, Faculty of Pharmacy, Ramaiah University of Applied Sciences, Karnataka, India; Olivia Rousset-Torrente, PRO team INSERM 1123, Paris, France; James W Griffith, Griffith Lab, Northwestern University, Chicago, Illinois, United States; Olivier Chassany, PRO team INSERM 1123, Chicago, France; Martin Duracinsky, PRO team INSERM 1123, Paris, France

**Aims:** Migrants who have a language barrier face significant hurdles in accessing care. Solutions include formal or informal interpreters but have benefits and drawbacks. Electronic tool can facilitate communication between migrants with low language proficiency and health professionals. We conducted a systematic review of the literature reporting the development and evaluation of applications designed to help communication between allophone migrants and health providers, or to promote health. **Methods:** We searched Pubmed, Embase, Scopus and clinicaltrials.gov. Keywords were defined with the help of a librarian. We included articles in French and English published after 1998 presenting electronic tool for international migrants not fluent with the language of the country they reside, cultural minorities having a language barrier and tourists. We excluded articles examining general translation applications, articles describing only the technical development, articles exploring only the perceptions of users and articles lacking sufficient information about the development or evaluation of an electronic tool. The selection of articles and the data collection were carried out independently by two researchers. Data collection included: health literacy and cultural adaptation; development of the application; evidence about the acceptability and efficacy of the application and use of the application. **Results:** The study is ongoing. The initial search retrieved 158 articles. Of the 40 applications already identified, 13 were designed to facilitate the dialogue between health professionals and allophone migrants and 27 to promote health among migrants. 27 of the applications were developed using scientific methods (mostly qualitative studies). The acceptability or usability has been tested for 29 of the applications and the efficacy has been evaluated for 21 of the apps. Some features associated with greater acceptability of medical translation applications are those that increase interactions and feedback (such as asking to reformulate a question or indicate that the answer is not understood). **Conclusion:** This systematic review will help having an overview of existing applications aiming to improve the communication between allophone migrants and health professionals. E-health tools have the potential to increase migrants' health literacy and health-care access. However more robust evaluation of their efficacy and impact are needed.

**(3234) A review of goal attainment scaling in clinical trials**

Taylor Dunn, MSc, DGI Clinical Inc, Halifax, Nova Scotia, Canada; Justin Stanley, BScE, DGI Clinical Inc, Halifax, Nova Scotia, Canada; Chère Chapman, MSc, MBA, DGI Clinical Inc, Halifax, Nova Scotia, Canada; Kari Knox, RN, DGI Clinical Inc, Halifax, Nova Scotia, Canada; Sanja Stanojevic, PhD, DGI Clinical Inc, Halifax, Nova Scotia, Canada; Susan Howlett, PhD, DGI Clinical Inc, Halifax, Nova Scotia, Canada; Kenneth Rockwood, MD, FRCPC, DGI Clinical Inc, Halifax, Nova Scotia, Canada

**Aims:** Capturing the spectrum of personally meaningful changes with standardized outcomes in clinical trials is challenging. Goal attainment scaling (GAS) is an individualized patient-reported outcome

that quantifies the effects of an intervention based on personal goals. It is well-suited to evaluating interventions for conditions with high heterogeneity in symptoms and progression. Our aim was to review how GAS is used in clinical trials. **Methods:** We searched ClinicalTrials.gov with the terms “goal attainment scale” and “goal attainment scaling,” without any additional restrictions. Each trial was reviewed by the authors. Areas of study (disease or condition) were consolidated with umbrella terms, allowing consistent naming. Studies were summarized based on type (interventional/observational) and trial phase (II, III, etc.). **Results:** Of 162 trials identified, we excluded 3 in which GAS was not an outcome. Start dates ranged from 2004-11-01 to 2021-10-01, with most ( $n = 103$ , 65%) starting in 2015 or later. Nearly half (71, 45%) of the trials were ongoing (pre-recruitment, recruiting or active). The majority were interventional (143, 90%); 16 (10%) were observational studies. Of the 30 interventional trials with phases listed, most were phase III (12, 40%), 8 (27%) were phase II and 9 (30%) were phase IV; one trial was a combination phase II and III trial. GAS was used as a primary outcome in 60 (38%) trials, as a secondary outcome in 89 (56%) trials, and an ‘other’ outcome in 10 (6%) trials. In 6 (4%) trials, GAS was also part of the intervention. The most common applications of GAS were in spasticity (31, 19% of trials), cerebral palsy (25, 16%), and stroke (17, 11%). Several novel uses of GAS were identified: 35 conditions were studied in just one trial, including hemophilia, Down syndrome and epilepsy. **Conclusion:** In recent years, there has been an increase in the number of trials using GAS across multiple disciplines, primarily in interventional studies as a secondary outcome. This shows a growing enthusiasm for GAS as a patient-reported outcome in clinical trials.

**(3235) Developing recommendations for patient-reported outcome measures in clinical quality registries**

Rasa Ruseckaite, PhD, Monash University, Melbourne, Australia; Ashika Maharaj, Monash University, Melbourne, Australia; Joanne Dean, Monash University, Melbourne, Australia; Karolina Kryszynska, University of Melbourne, Melbourne, Australia; Susannah Ahern, Monash University, Melbourne, Australia

**Aims:** Clinical quality registries (CQRs) provide feedback to the clinical community and wider stakeholder groups. Despite CQRs having incorporated patient perspective to support the evaluation of health care provision, no evidence-based guidelines nor recommendations for inclusion of patient-reported outcome measures (PROMs) in CQRs exist. The aim of this project was to develop a consensus set of recommendation statements via a Delphi process for PROMs inclusion in a CQR setting. **Methods:** An online two-round Delphi survey was performed among CQR data custodians, quality of life researchers, biostatisticians and clinicians recruited in Australia and overseas. A list of preliminary statements for the recommendations was based on the findings from the literature and the survey of the Australian registries, conducted in 2019. The statements were grouped into the following domains: rationale, setting, ethics, instrument, administration, data management, statistical methods, and feedback and reporting. **Results:** Of the 18 experts invited to participate in this study, eleven agreed to undertake the first survey (round one) and nine of them participated in the second round. Of the 117 statements presented to the experts in round one, 11 statements (9.4%) were rated unimportant, 55 (47.0%) as very important and there were 51 (43.6%) statements with disagreement on importance. The experts disagreed on PROMs administration, timing of data collection, mode and methods, statistical analyses and data management. Eleven new statements were suggested for the second round. In

total there were 66 items presented to the second round, of which 18 (27%) were deemed very important, 39 (59%) statements had disagreement on importance, and the remaining statements were rated as non-important. 74 statements were proposed for inclusion into the final set. These recommendations will assist new registries planning to implement PROMs data collection in the nearest future. Conclusion: The establishment of recommendations specific to CQRs will provide capacity to maximize the use of the patient's frame of reference and experience as consumers of healthcare to inform quality of care and further improve health outcomes. Incorporation of PROMs in CQRs ensures that the health outcomes that are important to patients are captured for both clinical care and research.

## Theory & Policy I

### (3236) A multilevel approach for the use of routinely collected patient-reported outcome measures (PROMs) data in health systems

Fatima Al Sayah, University of Alberta, Edmonton, Alberta, Canada; Markus Lahtinen, Health Quality Council of Alberta, Calgary, Alberta, Canada; Gouke Bonsel, EuroQol Research Foundation, Rotterdam, Netherlands; Arto Ohinmaa, University of Alberta, Edmonton, Alberta, Canada; Jeff A. Johnson, University of Alberta, Edmonton, Alberta, Canada

Aims: There is a growing recognition throughout the world that the patient's perspective is highly relevant to improving the quality and effectiveness of healthcare. The introduction of patient-reported outcome measures (PROMs) has been a strategy by which patients' perspectives are incorporated into the approaches of delivering healthcare services and evaluating the performance and efficiency of the healthcare system. This has led to an expansion in the routine collection of PROMs data. Little guidance on the use of these data within health systems for those purposes exists, however. We provide a framework for instrumental use of routinely collected PROMs data in health systems, drawing on examples from various PROMs applications in Canada. Methods: We provided an overview of utilization opportunities for PROMs, and proposed a multilevel framework defining the instrumental place of PROMs data in decision-making (including quality cycles, PDCA) at various levels: micro (e.g., patients, clinicians), meso (e.g., healthcare organizations), and macro (e.g., health system, policy-makers). Results: In Canada, and particularly the province of Alberta, tremendous efforts have been directed towards enhancing the use of PROMs within the healthcare system. The EQ-5D is the recommended PROM for use within the provincial healthcare system, and health authorities have invested in the routine collection of EQ-5D (alongside disease-specific PROMs) in various clinical populations. Successful examples of using PROMs within the Alberta healthcare system, include informing clinical practice, enhancing patient-centered care, enable individual decision-making, information for health services programming, enabling performance measurement activities, introducing comparative effectiveness analysis, and stimulating local quality improvement initiatives. Macro-level use of PROMs data is still limited at this stage. Conclusion: We will present several examples of routinely collected PROMs data being used to inform decision-making at the micro (e.g., patient management, usefulness of treatments), meso (e.g., program evaluation, healthcare delivery) and macro (e.g., performance measurement, resource planning and allocation) levels. There are several methodological challenges in using routinely collected PROMs data in health systems (e.g., attrition, missing data, lack of control arm, ill-fitting classifications, large data pitfalls, case-mix adjustment, sample representativeness) that need to be addressed and planned for.

### (3237) A literature review of patient-reported outcomes assessing symptom burden and quality of life in kidney transplantation

Garima Sharma, Novartis Healthcare Private Limited, Hyderabad, India; Christel Naujoks, Novartis Pharma AG, Basel, Switzerland; Matthieu Abbou, Novartis Pharma AG, Basel, Switzerland; Amanda Henry, Novartis Pharma, Dublin, Ireland

Aims: The use of immunosuppressive therapy imposes side effects on kidney transplant recipients, with associated symptom distress causing an adverse impact on physical and emotional quality of life, often leading to lower adherence. This study assessed suitability of available symptom specific and generic PRO measures for inclusion in future clinical trials, assessing potential new therapies. Methods: A targeted literature review was conducted on Ovid from January 2009 until May 2019 to retrieve studies in kidney transplantation recipients focused on measuring immunosuppressive symptom experience, and general health-related quality of life (HRQoL). For the PROs identified, the psychometric properties were assessed against the FDA's 2009 PRO guidance. Results: In total, 764 studies were included for the review. The tools found to evaluate kidney transplant symptom experience included the MTSOSD (Modified Transplant Symptom Occurrence and Symptom Distress Scale) (45 items), MTSOSD-59R (revised 59 items), ESRD-SCL-TM, KTQ, KDQOL-SF, and ESAS. Although MTSOSD-59R demonstrated good content validity with a wide coverage of side-effects, overall it had limited evidence to support its psychometric properties. The quality of life assessment tools identified were SF-36, PROMIS-29 or 57, and EQ-5D. Of these, SF-36 was the most widely used, and despite lack of sufficient psychometric evidence in kidney transplantation, it was found to be sensitive in detecting HRQoL related improvement. PROMIS-29 or 57 was the most recently validated tool for use in kidney transplantation but lacking evidence of responsiveness. Conclusion: A small number of generic, disease and symptom specific instruments were identified for use in kidney transplant trials, and had varying levels of evidence regarding their psychometric properties. The MTSOSD-59R was found acceptable for measuring symptoms associated with immunosuppressive regime. Both SF-36 and PROMIS-29 have demonstrated validity in kidney transplant patients and are considered appropriate for  $\geq \geq \geq \geq$  constructs of HRQoL.

### (3238) Quality of Work Life in informal economy workers of one city of Mexico

Raquel González-Baltazar, PhD, University of Guadalajara, Guadalajara, Jalisco, Mexico; Silvia G. León-Cortés, PhD, University of Guadalajara, Guadalajara, Jalisco, Mexico; Mónica I. Contreras-Estrada, PhD, University of Guadalajara, Guadalajara, Jalisco, Mexico; Brenda J. Hidalgo-González, Specialist, University of Guadalajara, Guadalajara, Jalisco, Mexico; Gustavo Hidalgo-Santacruz, Master, University of Guadalajara, Guadalajara, Jalisco, Mexico

Aims: The Quality of Work Life (QWL) has been studied only in formal work, regardless of economic conditions and employment policies have led to a substantial increase in the informal economy in our country. The informal worker is anyone who is part of a system of independent self-employment. According to the International Labour Organization (ILO), about 60% of workers in Mexico serve on informal employment (ILO, 2014). Currently the QWL is identified as an important indicator of health in workers, so it is important to conduct studies that can describe its behavior in workers. The aim of this research was to evaluate the Quality of Work Life in the informal economy workers in the metropolitan area of one city of Mexico. Methods: Voluntarily participated 507 informal workers who were

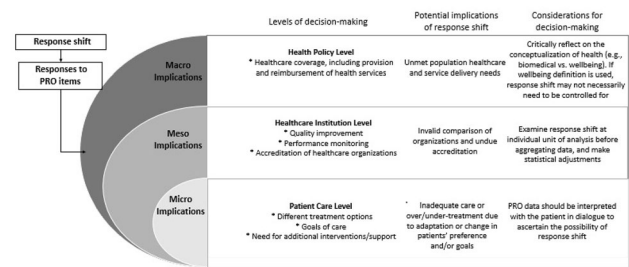
selected randomly, satisfaction with the QWL was measured with the CVT-GOHISALO instrument adapted to workers in the informal economy; the original instrument has validation of content, criterion and construct, with a reliability of 0.9527 of Cronbach's Alpha. The adapted instrument has validation of construct with a reliability of 0.92 of Cronbach's Alpha, with 50 items. Results: As for the study population, 48% of workers were men and 52% women, age range 15–80 years, the highest percentage was between 15 and 29 years. The most common high school education was 43%, followed by secondary with 29%. 67% are between 1 and 15 years of working in the informal economy and 67% work between 6 and 10 h a day. 55% of the population had a low level of satisfaction with their QWL, 39% a medium level and 6% a high level. Conclusion: This study found that in general terms the conditions of this sector workers despite not being poor with economic benefits obtained, evade paying taxes and have more flexibility with schedules; have a low level of satisfaction with their QWL; they do not have access to safety and occupational health, many works in unsanitary conditions and exposed to risks in the workplace.

### (3239) Implications of response shift for micro, meso, and macro healthcare decision-making using patient-reported outcomes

Richard Sawatzky, Trinity Western University, Langley, British Columbia, Canada; Jae-Yung Kwon, University of British Columbia, Vancouver, British Columbia, Canada; Ruth Barclay, University of Manitoba, Winnipeg, Manitoba, Canada; Cynthia Chauhan, Patient Representative, Wichita, Kansas, United States; Lori Frank, RAND Corporation, Arlington, Virginia, United States; Wilbert van den Hout, Leiden University Medical Center, Arlington, Netherlands; Lene Kongsgaard Nielsen, Odense University Hospital, Odense, Denmark; Sandra Nolte, ICON GmbH, Munich, Germany; Mirjam A. G. Sprangers, Amsterdam University Medical Centers, Amsterdam, Netherlands; submitted on behalf of Response Shift—in Sync Working Group

**Aims:** Patient-reported outcomes (PROs) are increasingly used to inform healthcare decision-making. Research has shown that response shift can impact measurement of PROs. As part of an international collaboration, our goal is to provide a framework regarding the implications of response shift at the level of patient care (micro), healthcare institute (meso), and healthcare policy (macro). **Methods:** Empirical evidence of response shift in patients' self-reported health status and preferences provided the foundation for development of the framework. Measurement validity theory, hermeneutic philosophy, and micro-, meso-, macro-level healthcare decision-making informed our theoretical analysis. Building on the perspective of measurement validity as an interpretive process, we drew upon a hermeneutic perspective for insights about response shift implications that recognize the different priorities and perspectives of stakeholders. **Results:** At the micro-level, patients' self-reported health outcomes need to be interpreted via dialogue with the clinician to avoid misinterpretation of PRO results due to response shift. It is also important to consider the potential impact of response shift on study results, when these are used to support decisions. At the meso-level, individual-level data should be examined for response shift before aggregating PROs for decision-making related to quality improvement, performance monitoring, and accreditation. At the macro-level, critical reflection on the conceptualization of health is required to know whether response shift needs to be controlled for when PRO results are used to inform healthcare coverage (see figure of potential implications). **Conclusion:** Given empirical evidence of response shift, there is a critical need for guidelines and knowledge translation to avoid potential misinterpretations of PROs and consequential biases in PRO-informed decision-making. Our framework with guiding questions provides a structure

for developing strategies to address the potential impacts of response shift at micro-, meso-, and macro-levels.



### (3240) Assessment of patient-reported outcomes measures in diabetic retinopathy-A literature review

Garima Sharma, Novartis Healthcare Private Limited, Hyderabad, India; Judit Banhazi, Novartis Pharma AG, Basel, Switzerland; Christine Bouchet, Novartis Pharma AG, Basel, Switzerland; Christel Naujoks, Novartis Pharma AG, Basel, Switzerland; Francesco Patalano, Novartis Pharma AG, Basel, Switzerland

**Aims:** Severe vision impairment from Proliferative diabetic retinopathy has been demonstrated to have a detrimental effect on patient's functioning and overall quality of life. The aim of the targeted literature search to identify suitable patient-reported outcome measures for use in diabetic retinopathy (DR) clinical studies. **Methods:** A targeted literature review was conducted on Ovid from January 1990 until September 2019 to find PROMs across DR studies, and the assessment of their psychometric properties were performed against the FDA's 2009 PRO guidance. **Results:** In total, 1529 abstracts met criteria for our screening process. The only DR-specific tool found to be used across studies was Retinopathy Dependent Quality of life (RetDQoL). National Eye Institute's Visual Function Questionnaire (NEI VFQ-25) was the most commonly used PRO measure across DR studies to assess vision-specific quality of life. The other tools included Impact of Vision Impairment (IVI), VF-14, Activity Inventory (AI), and Activities of Daily Living Scale (ADVS). The Retinopathy Treatment Satisfaction Questionnaire (RetTSQ) was used to measure treatment satisfaction in few studies. Based on the analysis of psychometric evidence, RetDQoL was found to be suboptimal to measure QoL in DR patients because of its complex scoring, and limited information on psychometric properties. While there was mixed evidence of validity for VFQ-25, and modified VFQ-28-R (Rasch-scored version) demonstrated improved performance. The more recently developed DR-specific item banks administered using Computerized adaptive testing (CAT) is the only diabetic retinopathy-specific PRO measure that is validated using rigorous item response theory (IRT) based psychometric techniques. However there was no evidence of its use in any interventional study in DR patients. **Conclusion:** Based on the analysis, the VFQ-28-R and RetTSQ are likely to detect treatment related change whereas additional evidence on responsiveness for DR-specific item banks would determine their suitability for use in DR trials.

### (3241) Predicting EQ-5D-3L health dimensions in people with impaired vision

Antonio Filipe Macedo, Linnaeus University, Faculty of Health and Life Sciences, Department of Medicine and Optometry, Kalmar, Sweden; Kristofer Årestedt, PhD, Linnaeus University, Faculty of Health and Life Sciences, Department of Health and Caring Sciences,

Kalmar, Sweden; Robert Massof, PhD, Johns Hopkins University School of Medicine, Wilmer Eye Institute, Baltimore, Maryland, United States; Ina Marteinsdottir, MD, PhD, Linnaeus University, Faculty of Health and Life Sciences, Department of Medicine and Optometry, Kalmar, Sweden; Evalill Nilsson, MD, PhD, Linnaeus University, Faculty of Health and Life Sciences, Department of Medicine and Optometry, Kalmar, Sweden; Mikael Rask, PhD, Linnaeus University, Faculty of Health and Life Sciences, Department of Health and Caring Sciences, Kalmar, Sweden; Jalal Safipour, PhD, Linnaeus University, Faculty of Health and Life Sciences, Department of Health and Caring Sciences, Växjö, Sweden; Hanna Tuveson, PhD, Linnaeus University, Faculty of Health and Life Sciences, Department of Health and Caring Sciences, Växjö, Sweden; Pedro Lima Ramos, Linnaeus University, Faculty of Health and Life Sciences, Department of Medicine and Optometry, Kalmar, Sweden; Amanda Hellström, PhD, Linnaeus University, Faculty of Health and Life Sciences, Department of Health and Caring Sciences, Kalmar, Sweden; PhD, Linnaeus University, Faculty of Health and Life Sciences, Department of Health and Caring Sciences, Kalmar, Sweden; submitted on behalf of the Portuguese visual impairment study group (PORVIS-group)

**Aims:** To determine predictors of EQ-5D-3L in persons with impaired vision, including demographic and clinical explanatory variables. **Methods:** Data from 494 persons with impaired vision was collected through face-to-face interviews at 4 Portuguese hospitals in the period 2014–2016. Data consisted of Visual Acuity in the Better Eye (VABE) and worse eye, comorbidities, sex, age and the Activity Inventory (AI). The AI is a condition-specific adaptive visual function questionnaire designed to provide visual ABILITY. The EQ-5D-3L covers five dimensions of health; Mobility (MO), Self-Care (SC),

Usual Activities (UA), Pain/Discomfort (PD), Anxiety/Depression (AD). Multiple partial proportional odds regression models (i.e., ordinal regression) were fitted with PROC LOGISTIC in SAS to each dimension. For MO and SC, levels 2 and 3 were merged due to a small number of responses with “extreme” problems and binary logistic regression was used instead. Results: The mean age was 64 years (SD = 14), 50% were females, mean VABE was 0.65logMAR (SD = 0.48) and mean visual ABILITY was 0.62logits (SD = 2.04). The adjusted odds ratio (OR) of significant predictors of reporting problems are given. MO, OR 1.77 (CI 95 1.10–2.83) for females and OR 0.37(CI 95–0.31–0.44) per-unit increase of ABILITY. SC, OR 0.37 (CI 95 0.30–0.46) per-unit increase of ABILITY (single predictor). UA, OR 2.29(CI 95 1.38–3.79) for people with comorbidities, OR 0.33 (CI 95 0.27–0.40) to “moderate” and OR 0.23(CI 95 0.14–0.38) to “extreme” problems per-unit increase of ABILITY. PD, OR 1.59 (CI 95 1.10–2.30) for people with more than 65 years, OR 0.56 (CI 95 0.36–0.88) per-unit increase of VABE and OR 0.62 (CI 95 0.55–0.69) per-unit increase of ABILITY. AD, OR 2.7 (CI 95 1.67–4.38) to “extreme” problems in females and OR 0.50 (CI 95 0.47–0.56) per-unit increase of ABILITY. Conclusion: The health dimensions of the EQ-5D showed a consistent relationship with visual ability. The EQ-5D and the AI agree in their assessment of functional limitations associated with impaired vision. Other variables like acuity in the better eye were less consistent at predicting EQ-5D scores. These results are important when using EQ-5D to investigate the burden of vision impairment.

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.