

# Methods of the ITC Four Country Smoking and Vaping Survey, wave 1 (2016)

Mary E. Thompson<sup>1</sup>, Geoffrey T. Fong<sup>2,3,4</sup>, Christian Boudreau<sup>1</sup>, Pete Driezen<sup>2</sup>, Grace Li<sup>2</sup>, Shannon Gravely<sup>2</sup>, K. Michael Cummings<sup>5,6</sup>, Bryan W. Heckman<sup>5,6</sup>, Richard O'Connor<sup>7</sup>, James F. Thrasher<sup>8,9</sup>, Georges Nahhas<sup>5,6</sup>, Ron Borland<sup>10,11</sup>, Hua-Hie Yong<sup>10,12</sup>, Ann McNeill<sup>13,14</sup>, Sara C. Hitchman<sup>13,14</sup> & Anne C. K. Quah<sup>2</sup>

## ABSTRACT

**Aim** To describe the methods of the 2016 International Tobacco Control (ITC) Four Country Smoking and Vaping (4CV) Survey, conducted in 2016 in Australia (AU), Canada (CA), England (EN) and the United States (US). **Methods** The respondents were cigarette smokers, former smokers (quit within the previous 2 years), and at-least-weekly vapers, aged 18 years and older. Eligible cohort members from the ITC Four Country Survey (4C) were retained. New respondents were sampled by commercial firms from their panels. Where possible, ages 18–24 and vapers were oversampled. Data were collected online, and respondents were remunerated. Survey weights were calibrated to benchmarks from nationally representative surveys. **Results** Response rates by country for new recruits once invited ranged from 15.2 to 49.6%. Sample sizes for smokers/former smokers were 1 504 in AU, 3006 in CA, 3773 in EN and 2239 in the US. Sample sizes for additional vapers were 727 in CA, 551 in EN and 494 in the US. **Conclusion** The International Tobacco Control Four Country Smoking and Vaping Survey design and data collection methods allow analyses to examine prospectively the use of cigarettes and nicotine vaping products in jurisdictions with different regulatory policies. The effects on the sampling designs and response quality of recruiting the respondents from commercial panels are mitigated by the use of demographic and geographic quotas in sampling; by quality control measures; and by the construction of survey weights taking into account smoking/vaping status, sex, age, education and geography.

**Keywords** Cohort survey, nicotine vaping products, online data collection, policy evaluation, sampling design.

Correspondence to: Mary E. Thompson, Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada, N2L 3G1. E-mail: methomps@uwaterloo.ca

[Correction added on 16 September 2019, after first online publication: Changes have been made to the article to improve clarity.]

## INTRODUCTION

This paper describes wave 1 of the ITC Four Country Smoking and Vaping (4CV) Survey conducted from July to November 2016 in Australia (AU), Canada (CA), England (EN) and the United States (US). The 4CV Survey is an expansion of the original ITC Four Country (4C) Survey [1–4], which had focused on evaluating the impact of tobacco control policies on smoking-related beliefs, attitudes and behaviors. During the 13-year period 2002–15, between nine and 11 waves were conducted in each of the four countries.

This paper briefly outlines the objectives of the 4CV Survey, the sampling design for the survey, data collection, survey outcome rates such as response and completion rates and the approach used for weighting data in an effort to generate a sample broadly representative of each country's population of adult smokers, recent former smokers and vapers at the time of the survey. The paper

also summarizes the strengths and weaknesses of the 4CV Survey methodology. A full description of the methods of the 4CV Survey is available in the ITC 4CV wave 1 (4CV1) Technical Report [5].

### Objectives of the ITC 4CV survey

Responding to the rapid evolution of the nicotine delivery market, the ITC Project team designed the 4CV Survey to address important issues regarding the use of this broad array of products, with special attention to the use of nicotine vaping products (NVPs). An important specific objective was to examine how policies on smoked tobacco, NVPs and other alternative nicotine products may influence the use of these classes of products and transitions among them by present and former smokers. The behavioral effects of these policies on known psychosocial precursors (mediators) of behaviors such as quit attempts and quit

success can serve as inputs to simulation modeling efforts to estimate the effects on population health [3,4,6].

The 4CV Survey provides an opportunity to study these questions in jurisdictions with different regulatory policies [7–10]. The policy environments for smoking and NVPs in the four countries and the conceptual model that describes how policies on cigarettes (and other smoked tobacco products), NVPs and other alternative nicotine products are hypothesized to affect use of these products and the transitions among them are described in greater detail in other papers in this supplement.

#### 4CV1 survey content

The 4CV wave 1 (4CV1) questionnaire addressed vaping patterns, both current and past; brand choices and purchasing; reasons for use; knowledge, beliefs and perceptions of harmfulness of NVPs; and attitudes toward regulation. At the same time, all sections of the 4C questionnaire on cigarette policies were retained, although the number of specific measures was reduced to allow for the addition of questions about NVPs. Table 1 provides a list of measures in the 4CV1 questionnaire on vaping and on cigarette smoking. More detailed information on the topics is available in the 4CV1 Technical Report [5] and the questionnaire [11].

#### Focus on population subgroups

There are some important differences between the designs of the 4CV Survey and the 4C Survey because of the new

research directions and the need for focus on certain subgroups. Because of our interest in NVP use and how it relates to smoking cessation, we included people who were former smokers of cigarettes (within 2 years) at the time of recruitment; we were interested in the stability of cessation versus relapse over this period and the fact that NVPs had only recently gained popularity as a cessation method. We expanded the recruitment criteria to include current (at least) weekly vapers, and oversampled these in CA, EN and the US to ensure that sufficient numbers were captured in our sample for analysis purposes; it was judged that vaping at least weekly would be necessary for NVP use as an aid to smoking cessation [12,13]. In AU, a supplementary dedicated vapor sample was recruited via referral sampling from online vapor forums and vape stores. Data from this group are not included in the core data set used for most analyses; given the sources and sampling method, these respondents are not expected to be representative of the population of all Australian vapers. As other studies had noted that younger adults were most likely to take up vaping (e.g. Kasza *et al.* [14]), smokers aged 18–24 years were oversampled in CA, EN and the US. Unfortunately, limited resources precluded oversampling younger adults in AU.

#### Inclusion criteria for 4C respondents

Cohort members in the final 4C wave in each of the four countries were invited to the 4CV Survey. In AU, members of the existing cohort were excluded if they had not smoked

**Table 1** Measures in the International Tobacco Control (ITC) Four Country Smoking and Vaping (4CV) survey questionnaire.

*Demographic variables:* gender, age, ethnicity, education, income, state of health

*Other personal moderators:* quitting history, nicotine dependence, levels of stress including financial stress and depressed mood, use of intoxicants (e.g. alcohol, cannabis) and experiences of use

*Environmental moderators:* number of smokers/nicotine vaping product (NVP) users in the household and in the social network

*Policy-specific (proximal) variables* (same measures for cigarettes/smoking and NVP use unless indicated):

- (1) Price paid per unit of product, total weekly cost, product type/variant, purchasing unit, price perceptions
- (2) Use of cessation services, recall of advice, NVP and/or other medicines use in conjunction with professional assistance, advice on appropriateness of NVP use
- (3) Advertising/ marketing: noticing advertising and frequency in key channels (TV, print, internet), susceptibility to advertising, reports about whether NVP advertising makes respondent think about cigarettes
- (4) Health warnings and packaging: salience and noticing of health warnings (if any), brand usage, perceived risks, perceived impact on product use; forgoing cigarettes/NVPs because of the warnings
- (5) Vapor-free laws: exposure to vaping across a range of settings, perceived impact on product use, reports on restrictions
- (6) Restrictions on access: perceived availability
- (7) Nicotine content, flavor and other product characteristics: nicotine content and flavors of vaping brands used, perceived addictiveness of NVPs and cigarettes and NVP appeal
- (8) Media campaigns: awareness and recall of media campaigns on NVPs

*Non-policy-specific (distal) psychosocial mediator variables:* social norms for both vaping and smoking, outcome expectancies, intentions for NVP use, reasons for NVP use, self-efficacy and intentions to quit smoking; relative harmfulness, health concerns, functions of smoking, substitutability of functions to NVP

*NVP and tobacco use behaviors:* key outcomes along with some of the distal variables for intermediary analyses. Use of NVPs and other nicotine products: frequency of use, duration and intensity of use (e.g. cigarettes per day); usual brand/type of product; quit attempts (smoking), duration of abstinence (smoking), product-switching

**Table 2** Original target sample sizes (realized sample sizes).

Component sample	AU	CA	EN	US
Smokers/former smokers aged 25+	1500 (1504)	2350 (2271)	2960 (2886)	1590 (1725)
Smokers/former smokers aged 18–24		750 (735)	1100 (887)	500 (514)
Total smokers/former smokers	1500 (1504)	3100 (3006)	4060 (3773)	2090 (2239)
Additional (at least) weekly vapers	NA	715 (727)	500 (551)	500 (494)

Realized sample sizes in 'Smokers/former smokers' rows include also some long-term former smokers in the International Tobacco Control (ITC) cohort retained by virtue of vaping at least weekly; the vapers in the final row are all newly recruited in 4CV1. AU = Australia; CA = Canada; EN = England; US = United States. NA = not applicable.

cigarettes for more than 2 years and did not vape at least weekly at the time of the 4CV1 Survey. In the other three countries, former smokers who had not smoked cigarettes for more than 2 years were retained, to provide supplementary samples of longer-term quitters.

#### Inclusion criteria for new respondent subgroups

New respondents in each country were recruited from the panels of commercial survey firms. A new respondent (aged 18 or older) could be recruited as someone who:

- had smoked at least 100 cigarettes in lifetime and was currently smoking at least monthly or 'less than monthly but occasionally'; or
- had smoked at least 100 cigarettes in their life-time and had quit within the previous 2 years; or
- was currently vaping at least weekly.

Note that a respondent eligible because of smoking or cessation status might vape only monthly, occasionally, or not at all; similarly, a respondent eligible because of vaping status might be a long-term quitter or a never smoker. The recruitment of less-than-monthly smokers was motivated by the wish to study transitions of occasional smokers into smoking, vaping or cessation over time.

#### The use of commercial firms for recruitment

The gold standard of probability sampling, such as address-based sampling with face-to-face recruitment or random-digit dialing (RDD) with telephone recruitment, is becoming increasingly difficult to attain, due to the high costs of maintaining strong contact and participation rates. The ITC Project has witnessed this dramatic increase in the difficulty in conducting national surveys employing rigorous probability sampling methods. The 4C wave 1 Survey in 2002 was a telephone survey with recruitment by RDD. This method was maintained for the first six waves, with ever-increasing costs. From wave 7 (2008–09) onwards, we began to move our survey administration to the web,

while continuing to recruit respondents through RDD. The shift was gradual, to allow modeling to take account of measurement differences between the two survey modes, and to prevent the loss of participants who were without web access or who otherwise preferred telephone administration. The RDD method was supplemented in waves 9 and 10 (2013–14) by recruitment from commercial panels, as an increasing proportion of the sample members (including telephone recruits) were responding online.

At the final 4C wave (ending in 2015), the data collection costs had tripled and become unattainable. From considerations of practicality and cost, it was decided that new recruits to the 4CV Survey would be entirely sourced from web panels in each of the four countries. We chose panels that (1) had high standards regarding the recruitment of their panel members;<sup>1</sup> (2) offered strong collaborative relationships with our survey management team in survey development, programming, testing and validation; and (3) quoted higher expected rates of retention over time. All survey response data were to be collected through an online questionnaire, hosted by the Survey Research Centre (SRC) at the University of Waterloo.<sup>2</sup>

Below we describe the sampling designs and further details on data collection for the ITC 4CV1 Survey; survey outcome rates; and the construction of survey weights.

#### Sample sizes, sampling designs and data collection

In CA, EN and the US, the new recruit sample for the 4CV1 Survey was to consist mainly of cigarette smokers and former smokers (within 2 years), aged 18 and older. Because of the importance of longitudinal analysis over waves, the 4CV1 sample sizes for the smoker/former smoker respondents aged 25+ were chosen so as to provide at least 1400 who would be present in both waves 1 and 2 (to be conducted 18 months later), using retention rate estimates that the survey firms provided at the design stage. Table 2

<sup>1</sup>High standards include use of probability-based methods where feasible, e-mail confirmation of recruits signing up through the panel website, proper remuneration of panel members, monitoring and removal of non-engaged responders.

<sup>2</sup>This requirement was later relaxed for 46 US respondents who were members of the 4C cohort and who agreed to participate by telephone but not online.

shows the original 4CV1 target sample sizes and realized sample sizes, for this group and the oversampled ones.

In AU, the sampling design was targeted at 1500 respondents satisfying the eligibility criteria, without any oversampling.

To mitigate potential non-representativeness of the samples from the survey firms, geographic, age and sex quotas were applied to sample sizes, as indicated in Table 3. Further details on the determination of quotas are given in the 4CV1 Technical Report [5].

### Sample sources

Table 4 shows the sources and their sample sizes in each of the countries.

### Remuneration of respondents

The 4CV1 questionnaire was lengthy; median survey lengths by country ranged from 39 to 48 minutes

(Table 5). The survey tended to be longer for respondents who both smoke and vape and respondents who had made failed quit attempts. It was shorter for users of cigarettes only or vaping products only and for respondents who had quit completely. Because of the significant time commitment, respondents received remuneration for their time upon completion of the questionnaire. The remuneration took various forms (cheque, gift card, reward points, entry into a prize draw), depending on the source and country of respondents. Remuneration by cheque ranged between \$20 and \$30 (US). The remuneration not only aided recruitment, but also is expected to impact retention rates at future waves favorably. Remuneration details are available in the 4CV1 Technical Report [5].

### Quality control

The commercial survey firms applied their own quality control checks to completed questionnaires before

**Table 3** Types of quotas applied to sample composition.

<i>Sample component</i>	<i>Quota cells</i>
AU smokers/former smokers age 18+	Geographic regions crossed with sex (realized approximately)
CA, US smokers/former smokers age 18–24	Geographic regions; language in Canada
CA smokers/former smokers age 25+	Geographic regions; language
US smokers/former smokers age 25+	None; most of the sample was from the 4C cohort, expanded in 4C wave 9
CA, US, EN additional vapers age 18+	Geographic regions
EN smokers/former smokers age 18–24	Geographic regions crossed with sex
EN smokers/former smokers age 25+	Geographic regions crossed with sex and age group

AU = Australia; CA = Canada; EN = England; US = United States.

**Table 4** Sources and sample sizes for International Tobacco Control (ITC) Four Country Smoking and Vaping (4CV) survey respondents.

<i>Country, subsample</i>	<i>Sources</i>	<i>Sample size</i>
AU, smokers/former smokers 18+	ITC 4C cohort (retention rate 43.1%)	515
	Roy Morgan single source (probability-based)	504
	Survey sampling international	485
CA, smokers/former smokers 18+	ITC 4C cohort (retention rate 41.6%)	567
	Léger Marketing, where possible from their probability-based panel	2439
CA, additional vapers 18+	Léger Marketing	727
EN, smokers/former smokers 18+	ITC 4C cohort (retention rate 35.7%)	254
	Ipsos	3519
EN, additional vapers 18+	Ipsos	551
US, smokers/former smokers aged 25+	ITC 4C cohort	1372
	GfK knowledge panel (probability based)	127
	Ipsos	226
US, smokers/former smokers aged 18–24	ITC 4C cohort (retention rate 44.2%)	6
	Lucid (GfK partner panel, opt-in)	496
	Ipsos	12
US, additional vapers 18+	Ipsos	494
Total		12 294

AU = Australia; CA = Canada; EN = England; US = United States.

**Table 5** Outcome rates by country, for new recruits.

Country	AU	CA	EN	US-GfK	US-all <sup>c</sup>
1. Total interviewed	1575	3182	4129	131	1377
2. Refusals/breakoffs	116	309	328	12	335
3. Not known if eligible <sup>a</sup>	6338	67 194	171 438	434	1422
4. Estimated rate of eligibility and quota-not-full	49.3%	19.7%	13.3%	28.0%	36.6%
5. Estimated number of eligible and quota-not-full non-respondents in 3	3123	13 204	22 784	122	580
6. Response rate <sup>d</sup>	32.7%	19.1%	15.2%	49.6%	60.1%
7. Cooperation rate <sup>d</sup>	93.1%	91.1%	92.6%	91.6%	80.4%
8. Median survey length <sup>b</sup>	48 min	44 min	39 min	NA	45 min

<sup>a</sup>Sent to Survey Research Centre (SRC), did not respond, unknown if eligible. Roy Morgan (and their partner SSI) in AU and IPSOS in EN pre-screened respondents before sending them to SRC's website. Léger in CA and GfK in the US pre-identified individuals who would probably be eligible before sending them to SRC's website. <sup>b</sup>The questionnaire differed somewhat from country to country. 'Rates in the 'US-all' column should be viewed with caution and are likely to be overestimates: Lucid (for respondents aged 18–24) and IPSOS (for respondents recruited as VNP users) were not able to provide the total number of respondents they invited to the survey. <sup>c</sup>The denominator of the response rate is the sum of 1, 2 and 5. The denominator of the cooperation rate is the sum of 1 and 2. The numerator for each of those rates is 1. AU = Australia; CA = Canada; EN = England; US = United States.

assigning them a 'complete' status. Once the data were received by the ITC Data Management Core, they were further examined for identification and removal of poor-quality responses [15]. Three indicators of possible poor quality were used: seconds per question, percentage of responses that were either refused or don't know (RDK) and the percentage of questions which had the first response checked (top-box). Some extreme values occurred for both of these variables, e.g. times of less than 1.7 sec per question, which by published estimates do not allow time for reading the questions (reading speeds quoted by Zhang & Conrad (2013) [15] are 200–300 milliseconds per word), and RDK responses for more than 70% of the questions completed.

The basic strategy for removal was to create a group of normal respondents by dropping all 'suspicious' respondents, namely those with very low seconds per question and/or high percentage RDK or top-box, and to use this main group to calculate normal behavior ranges. Points were then assigned for the degree of departure of each of the three indicators from normal behavior, and the questionnaires of those for whom the points totals were too high were discarded. In this manner, 87 records (0.7% of the sample) with low seconds per question and/or high percentage RDK or top-box (26 from the US, 13 from CA, 48 from EN and none from AU) were removed from the initial data set.

### Survey outcome rates

Response rates and cooperation rates are provided in Table 5.

Detailed information about the disposition codes and the computation of rates can be found in the 4CV1

Technical Report [5]. The cooperation rate is defined as the number of completed interviews as a percentage of the number of those who entered the survey and proceeded as far as confirming their eligibility. The response rate is defined as the number of completed interviews as a percentage of an estimated number who were invited to the survey at a time when their quota was 'open' and who were eligible to participate. The estimated number in the denominator of the response rate was obtained from observed eligibility rates for those whose eligibility status is known, and the observed quota open rate for the survey as a whole.

### Survey weights

As with most survey weights, the 4CV1 weights are constructed to correct and adjust for sample misrepresentation caused by unequal sampling probabilities, frame error (i.e. undercoverage and multiplicity) and non-response, as well as to improve precision of estimates through the use of auxiliary information (e.g. smoking prevalences).<sup>3</sup>

Except in the case of the cross-sectional inflation weights, all weights were rescaled to sum to overall sample size within each country. The cross-sectional weights are intended to be used for analyses using 4CV1 data alone; the longitudinal weights are constructed for analyses using both 4CV1 data and 4C data of respondents who were present in later waves of the 4C Survey.

### Cross-sectional sampling weights

We constructed 4CV1 cross-sectional inflation weights for the data set. The inflation weight of a respondent can be interpreted as the number of people in the population represented by that respondent. Inflation weights were

<sup>3</sup>See Levy & Lemeshow [16] for a more detailed discussion of the rationale for weights and their construction.

<sup>4</sup>By design, there are no 4C AU respondents in the data set who were quit for longer than 2 years and were not vapers at the time of data collection.



computed by dividing all respondents into four broad user groups:

- A: users of cigarettes but not vaping products
- B: users of both cigarettes and vaping products (individuals who smoke at least occasionally and vape at least monthly)
- C: exclusive vapers (individuals who vape at least monthly, but are not current or former cigarette users)
- D: former cigarette users, who may or may not be vapers

Former cigarette users were further categorized into four subgroups determined by their quit duration (< 1 year and 1–5 years in the US, CA and EN; < 1 year and 1–2 years in AU<sup>4</sup>) crossed with whether or not they vape at least monthly.

Calibration or benchmark figures were then obtained from national population surveys for each of groups A–D and the four subgroups of D. The benchmark surveys used for our purposes are described in the Supporting information, Table S1. See Supporting information, Tables S2 and S3 for more details of the sources, and see the references for the online documentation of these surveys [17–21].

In addition to an overall estimated number of individuals in each of user groups A–D and the four subgroups of D, estimates were obtained from the same sources for the following cross-tabulations: user group by sex, user group by age group, user group by geographic region, user group by ethnicity (US only), user group by education (except in CA where education was not collected in the benchmark survey) and user group by language (CA only). A raking procedure (Battaglia *et al.* 2009 [22]) was then applied to calibrate the weights using the above-mentioned cross-tabulations; this was conducted separately for each country. Some cells from the above-mentioned cross-tabulations were collapsed in order to reduce the instability of the sampling weights that would result due to small cell sizes.

The cross-sectional inflation weights are designed to make respondents in each of groups A–D and subgroups of D representative of the corresponding population at the time of 4CV1 data collection. For example, the cross-sectional inflation weight of someone in the CA group B sample (individuals who smoke cigarettes and also vape) is the estimated number of people in the CA group B population (in the same age–sex group) represented by that individual. Rescaled cross-sectional weights are also provided, which are simply the 4CV1 cross-sectional inflation weights rescaled to sum to the sample size of each country. As a consequence, they have an average value of 1 in each country.

The inflation weights are needed for estimating population totals and numbers; the rescaled cross-sectional

weights can be used for most other purposes, such as estimation of means and proportions and regression analyses.

The cross-sectional inflation weights should not normally be used in analyses involving two or more countries. This is due to the fact that the numbers of smokers (and vapers) differ greatly by country. From the data used to calibrate the weights, there were about 39.8 million cigarette smokers in the US at the time of 4CV1 data collection, compared to only 3.6 million such individuals in CA, 7.4 million in EN and 2.7 million in AU. Hence, any joint analysis using data from all four countries will be dominated by the US if the inflation weights are used. If the rescaled weights are used, EN and CA will have a slightly greater impact on the results (since the sample sizes in EN and CA are larger than those of AU and the US; see Supporting information, Table S1), but no country will dominate the analysis. Various other rescaled weights have been created for specific user groups for multi-country analyses. More information about those weights is available in the 4CV1 Technical Report [5].

### Longitudinal sampling weights

The rescaled waves w-to-4CV1 longitudinal weights (where w = 8, 8.5, 9 or 10) are computed for the respondents from the relevant countries who completed the 4C wave w survey, and were successfully retained and interviewed at 4CV1. These weights are designed to make these smokers (and former smokers) representative of their country's population of smokers at the time of 4C wave w data collection. They would be used in analyses of associations between outcomes such as cessation at the time 4CV1 and predictors for the same individuals in 4C data.

## DISCUSSION

This paper summarizes the objectives and methods of the ITC 4CV Survey. The strengths of the 4CV design are its deployment in four countries that are culturally and economically similar, but have different regulatory environments that affect the marketing and sale of NVPs; its longitudinal structure, allowing for the observations of transitions in product use and non-use; its relatively large sample sizes; and its focus on subpopulations of particular interest for examining the relationship between smoking and NVP use and transitions between them.

There are challenges and limitations connected with the implementation of the sampling design that are important to note. The 4CV1 samples were not selected as probability samples from the combined population of smokers,

<sup>4</sup>By design, there are no 4C AU respondents in the data set who were quit for longer than 2 years and were not vapers at the time of data collection.

recent ex-smokers and vapers of each country. With a limited budget, we recruited specified numbers of adult cigarette users, former smokers (within 2 years) and vapers from commercial databases, only some of which were put together from members of probability samples from one or several sources; respondents had to have accepted the invitation from the firm to be in the database, and to accept the invitation to participate in the 4CV Survey. Accordingly, in the 4CV1 Survey, it is generally not feasible to compute inclusion probabilities for sample members, or to adjust these for differential non-response. The effects of not using direct probability sampling, and of using respondents from commercial panels, have been mitigated to some extent by the use of sex, age and geographic quotas in sampling; by efforts to eliminate speeders from the data set; and by the construction of survey weights taking into account cigarette/vaping status, sex, age, education (where possible) and geography. Well-established national general population surveys employing probability sampling methods were used to weight the 4CV1 data so that estimated prevalence rates for smoking and vaping status are approximately in line with what would be expected in each country.

For many analyses, such as those examining the relationships among attitude and behavior variables, the 4CV1 data are expected to be trustworthy. Given the approximate nature of the benchmarks, the usual advice for analyses with weighted data is especially important: in linear and logistic regression analyses, it is important to include among the explanatory variables or controls the main weighting variables, namely the smoker/vaper/former smoker user groups, sex, age group and education.

However, 4CV1 'descriptive' estimates of population means and proportions, such as prevalences of certain behaviors, must be interpreted with caution. One reason is that the benchmark surveys are contemporary with the 4CV1 Survey for AU and the US, whereas national surveys conducted in 2015 were used to calibrate the sampling weights for CA and EN. It follows that if the 4CV1 CA and EN data are used for prevalence estimates, these prevalence estimates would be valid for an earlier time. More importantly, for all the countries, without a direct probability sampling design as a basis, prevalence estimates from 4CV1 for categories not used in the calibration could well be biased. For example:

- i The subsample of former smokers cannot be expected to be representative of former smokers in their populations. A sizable number of former smokers in the 4CV1 sample were initially recruited in the 4C cohort as smokers and cannot be said to have been sampled at random from the population of former smokers at the time of the 4CV1 data collection. Moreover, dropout from the 4C sample has probably been associated with cessation, and the cessation experience of the longitudinal respondents

could well have been affected by being in the sample. For these reasons, among others, former smokers in the 4CV1 sample should not be regarded as representative of former smokers in their populations. Calibration has been used to make the recent former smokers in the sample more representative with respect to demographics (sex, age group and education), quit duration and vaping, but there is no assertion that the recent former smokers are representative with respect to these last two variables crossed with demographic variables.

- ii Among vapers in the sample who are smokers or former smokers, monthly or occasional vaping is quite common. At the same time, the inclusion criteria for the additional vaper samples excluded those who ONLY vape, but do so less than weekly. For this and other reasons, it is useful to include vaping frequency in models of the associations of outcomes of interest with vaping.

## SUMMARY

The ITC 4CV Survey uses a sampling design, data collection methods, together with an analytical strategy intended to account for departures from the ideal situation of data from respondents recruited by probability sampling in all populations of interest in all countries. The result is a data set allowing for sound analyses to address the objectives: to measure and understand the impact of policy and non-policy factors on the use of cigarettes and NVPs and on transitions among these products or away from their use (i.e. cessation).

## ETHICS APPROVAL

The survey protocols and all materials, including the survey questionnaires, were cleared for ethics by Institutional Review Board, Medical University of South Carolina; Research Ethics Office, King's College London, UK; Office of Research Ethics, University of Waterloo, Canada; and Human Research Ethics, Cancer Council Victoria, Australia. All participants provided consent to participate.

## Declaration of Interests

K.M.C. has received payment as a consultant to Pfizer, Inc., for service on an external advisory panel to assess ways to improve smoking cessation delivery in health care settings. K.M.C. also has served as paid expert witness in litigation filed against the tobacco industry. J.E.T. and G.T.F. have served as expert witnesses on behalf of governments in litigation involving the cigarette industry. All other authors have no conflicts of interest to declare.

## Acknowledgements

This study was supported by grants from the US National Cancer Institute (P01 CA200512), the Canadian Institutes of Health Research (FDN-148477), and by the National Health and Medical Research Council of Australia (APP 1106451). S.G. was funded by a 3-year Canadian Cancer Society (CCS) Career Development Award in Prevention (703858). G.T.E. was supported by a Senior Investigator Award from the Ontario Institute for Cancer Research. B.W.H. was supported by NIDA (K23 DA041616). The sponsors had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; preparation, review or approval of the manuscript; and decision to submit for publication.

## Authors' affiliations

Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON, Canada,<sup>1</sup> Department of Psychology, University of Waterloo, Waterloo, ON, Canada,<sup>2</sup> School of Public Health and Health Systems, University of Waterloo, Waterloo, ON, Canada,<sup>3</sup> Ontario Institute for Cancer Research, Toronto, ON, Canada,<sup>4</sup> Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, SC, USA,<sup>5</sup> Hollings Cancer Center, Medical University of South Carolina, Charleston, SC, USA,<sup>6</sup> Department of Health Behavior, Roswell Park Comprehensive Cancer Center, Buffalo, NY, USA,<sup>7</sup> Department of Health Promotion, Education and Behavior, Arnold School of Public Health, University of South Carolina, Columbia, SC, USA,<sup>8</sup> Tobacco Research Department, Center for Population Health Research, Cuernavaca, Mexico,<sup>9</sup> Cancer Council Victoria, Melbourne, VIC, Australia,<sup>10</sup> School of Psychological Sciences, University of Melbourne, Melbourne, VIC, Australia,<sup>11</sup> School of Psychology, Deakin University, Geelong, VIC, Australia,<sup>12</sup> Addictions Department, King's College London, London, UK<sup>13</sup> and UK Centre for Tobacco & Alcohol Studies, Nottingham, UK<sup>14</sup>

## References

1. ITC Project (2004). ITC Four Country Wave 1 (2002) Technical Report. University of Waterloo, Waterloo, Ontario, Canada; Medical University of South Carolina, Charleston, South Carolina, United States; VicHealth Centre for Tobacco Control, Carlton, Australia; Cancer Control Victoria, Melbourne, Australia; King's College London, London, United Kingdom; University of Stirling, Stirling, United Kingdom; and the Open University, Buckinghamshire, United Kingdom. 2004 (Archived at <http://www.webcitation.org/73jHXktnd> on November 6, 2018).
2. ITC Project (2011). ITC Four Country Waves 2 to 8 (2003–2011) Technical Report. University of Waterloo, Waterloo, Ontario, Canada; Medical University of South Carolina, Charleston, South Carolina, United States; VicHealth Centre for Tobacco Control, Carlton, Australia; Cancer Control Victoria, Melbourne, Australia; King's College London, London, United Kingdom; University of Stirling, Stirling, United Kingdom; and the Open University, Buckinghamshire, United Kingdom. 2011 (Archived at <http://www.webcitation.org/73jHg8tRf> on November 6, 2018).
3. Fong G. T., Cummings K. M., Borland R., Hastings G., Hyland A., Giovino G. A. *et al.* The conceptual framework of the International Tobacco Control (ITC) policy evaluation Project. *Tob Control* 2006; **15**: iii3–iii11.
4. Thompson M. E., Fong G. T., Hammond D., Boudreau C., Driezen P., Hyland A. *et al.* Methods of the International Tobacco Control (ITC) Four Country Survey. *Tob Control* 2006; **15**: iii12–iii18.
5. International Tobacco Control (ITC) Project. ITC Four Country Smoking and Vaping Survey Wave 1 (2016) Technical Report. University of Waterloo, Waterloo, Ontario, Canada; Medical University of South Carolina, Charleston, South Carolina, United States; Cancer Council Victoria, Melbourne, Australia; King's College London, London, United Kingdom. 2017 (Archived at <http://www.webcitation.org/73jHzqz0z> on November 6, 2018).
6. IARC Handbooks of Cancer Prevention. *Tobacco Control*, vol. 12: *Methods for Evaluating Tobacco Control Policies*. Lyon, France; 2008.
7. Canadian Cancer Society. E-cigarettes: Where we are in 2016. 2016 (Archived at <http://www.webcitation.org/73jHstCxM> on November 6, 2018).
8. Yong H. H., Borland R., Balmford J., Hitchman S. C., Cummings K. M., Driezen P. *et al.* Prevalence and correlates of the belief that electronic cigarettes are a lot less harmful than conventional cigarettes under the different regulatory environments of Australia and the United Kingdom. *Nicotine Tob Res* 2017; **19**: 258–63.
9. Adkison S. E., O'Connor R. J., Bansal-Travers M., Hyland A., Borland R., Yong H. H. *et al.* Electronic nicotine delivery systems. International Tobacco Control Four-Country Survey. *Am J Prev Med* 2013; **44**: 207–15.
10. Gravely S., Driezen P., Ouimet J., Quah A. C. K., Cummings K. M., Thompson M. E. *et al.* Awareness, trial, and use of electronic cigarettes among adult current smokers and ex-smokers in 14 countries with differing EC regulatory policies: Results from the International Tobacco Control Policy Evaluation Project. In press.
11. 4CV1 Questionnaire (Archived at <http://www.webcitation.org/73jI5oUbj> on November 6, 2018).
12. Biener L., Hargraves J. L. A longitudinal study of electronic cigarette use among a population-based sample of adult smokers: association with smoking cessation and motivation to quit. *Nicotine Tob Res* 2015; **17**: 127–33.
13. Hitchman S. C., Brose L. S., Brown J., Robson D., McNeill A. Associations between e-cigarette type, frequency of use, and quitting smoking: findings from a longitudinal online panel survey in Great Britain. *Nicotine Tob Res* 2015; **17**: 1187–94.
14. Kasza K., Ambrose B. K., Conway K. P., Borek N., Taylor K., Goniewicz M. *et al.* Tobacco-product use by adults and youths in the United States in 2013 and 2014. *N Engl J Med* 2017; **376**: 342–53.
15. Zhang C., Conrad F. G. Speeding in web surveys: the tendency to answer very fast and its association with straightlining. *Surv Res Methods* 2013; **8**: 127–35.
16. Levy P. S., Lemeshow S. *Sampling of Populations: Methods and Applications*, 4th edn. Hoboken, NJ: John Wiley & Sons; 2008.
17. National Drug Strategy and Household Survey 2016 (Archived at <http://www.webcitation.org/73jIFsDvM> on November 6, 2018).
18. Canadian Tobacco Alcohol and Drugs Survey 2015 (Archived at <http://www.webcitation.org/73jIO05Xv> on November 6, 2018).
19. Opinions and Lifestyle Survey 2015 (Archived at <http://www.webcitation.org/73jISTQpO> on November 6, 2018).
20. Smoking Toolkit Study (Archived at <http://www.webcitation.org/73jIWUMyf> on November 6, 2018).



21. National Health Interview Survey 2016 (Archived at <http://www.webcitation.org/73jJVYDUL> on November 6, 2018).
22. Battaglia M. P., Hoaglin D. C., Frankel M. R. Practical considerations in raking survey data. *Surv Pract* 2009; **2**: 1–13.

**Table S1** National Benchmark Surveys.

**Table S2** Questions in the Benchmark Surveys.

**Table S3** Benchmarks for the User Groups and Subgroups for Weights Calibration.

### Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.