
DOI: 10.2196/jmir.4568

This is the published version.

©2015, The Authors

Reproduced by Deakin University under the terms of the Creative Commons Attribution Licence

Available from Deakin Research Online:

http://hdl.handle.net/10536/DRO/DU:30081439
Apps for IMproving FITness and Increasing Physical Activity Among Young People: The AIMFIT Pragmatic Randomized Controlled Trial

Artur Direito, MSc; Yannan Jiang, PhD; Robyn Whittaker, PhD; Ralph Maddison, PhD
Faculty of Medical and Health Sciences, National Institute for Health Innovation, University of Auckland, Auckland, New Zealand

Corresponding Author:
Artur Direito, MSc
Faculty of Medical and Health Sciences
National Institute for Health Innovation
University of Auckland
National Institute for Health Innovation, School of Population Health, University of Auckland
Private Bag 92019, Auckland Mail Centre
Auckland, 1142
New Zealand
Phone: 64 9 373 7599 ext 85285
Fax: 64 9 373 1710
Email: a.direito@auckland.ac.nz

Abstract

Background: Given the global prevalence of insufficient physical activity (PA), effective interventions that attenuate age-related decline in PA levels are needed. Mobile phone interventions that positively affect health (mHealth) show promise; however, their impact on PA levels and fitness in young people is unclear and little is known about what makes a good mHealth app.

Objective: The aim was to determine the effects of two commercially available smartphone apps (Zombies, Run and Get Running) on cardiorespiratory fitness and PA levels in insufficiently active healthy young people. A second aim was to identify the features of the app design that may contribute to improved fitness and PA levels.

Methods: Apps for IMproving FITness (AIMFIT) was a 3-arm, parallel, randomized controlled trial conducted in Auckland, New Zealand. Participants were recruited through advertisements in electronic mailing lists, local newspapers, flyers posted in community locations, and presentations at schools. Eligible young people aged 14-17 years were allocated at random to 1 of 3 conditions: (1) use of an immersive app (Zombies, Run), (2) use of a nonimmersive app (Get Running), or (3) usual behavior (control). Both smartphone apps consisted of a fully automated 8-week training program designed to improve fitness and ability to run 5 km; however, the immersive app featured a game-themed design and narrative. Intention-to-treat analysis was performed using data collected face-to-face at baseline and 8 weeks, and all regression models were adjusted for baseline outcome value and gender. The primary outcome was cardiorespiratory fitness, objectively assessed as time to complete the 1-mile run/walk test at 8 weeks. Secondary outcomes were PA levels (accelerometry and self-reported), enjoyment, psychological need satisfaction, self-efficacy, and acceptability and usability of the apps.

Results: A total of 51 participants were randomized to the immersive app intervention (n=17), nonimmersive app intervention (n=16), or the control group (n=18). The mean age of participants was 15.7 (SD 1.2) years; participants were mostly NZ Europeans (61%, 31/51) and 57% (29/51) were female. Overall retention rate was 96% (49/51). There was no significant intervention effect on the primary outcome using either of the apps. Compared to the control, time to complete the fitness test was −28.4 seconds shorter (95% CI −66.5 to 9.82, P=.20) for the immersive app group and −24.7 seconds (95% CI −63.5 to 14.2, P=.32) for the nonimmersive app group. No significant intervention effects were found for secondary outcomes.

Conclusions: Although apps have the ability to increase reach at a low cost, our pragmatic approach using readily available commercial apps as a stand-alone instrument did not have a significant effect on fitness. However, interest in future use of PA apps is promising and highlights a potentially important role of these tools in a multifaceted approach to increase fitness, promote PA, and consequently reduce the adverse health outcomes associated with insufficient activity.
Therefore, novel approaches are crucial to engage those who have limited adherence or exposure to PA interventions [16]. Many existing interventions, including school-based programs because schools integrate almost all young people. A Cochrane systematic review of school-based PA interventions targeting areas such as curriculum, teacher training, educational materials, and accessibility to exercise equipment (44 randomized controlled trials [RCTs], N=36,593, age 6-18 years) provided some evidence of effectiveness for increasing time spent in cardiorespiratory fitness (CRF) of approximately 5% every decade from 1970 to 2000 [3]. Insufficient activity is associated with increased risk of chronic diseases, such as cardiovascular diseases, type 2 diabetes mellitus, some types of cancer, and premature death [4,5]. Likewise, lower levels of CRF, an important marker of overall physiologic health [6,7] sensitive to patterns of PA, also predict future morbidity and mortality [8,9]. Increasing levels of PA and CRF in young people are important public health priorities for enhancing health, well-being, and preventing disease [10].

Physical activity declines with age throughout adolescence [2,11,12] and tracks into adulthood [13], making this an important time to intervene. Common approaches for intervening in this population have included school-based programs because schools integrate almost all young people. A Cochrane systematic review of school-based PA interventions targeting areas such as curriculum, teacher training, educational materials, and accessibility to exercise equipment (44 randomized controlled trials [RCTs], N=36,593, age 6-18 years) provided some evidence of effectiveness for increasing time spent in moderate-to-vigorous intensity PA (MVPA) from 5 to 45 minutes more per day and CRF (VO2peak) from 1.6 to 3.7 mL/kg/min [14]. However, caution interpreting these findings is warranted because the included studies were at moderate risk of bias, with the majority having relied on self-reported activity [15]. Notably, a 2012 meta-analysis (30 RCTs, N=14,326, age ≤16 years), which differs from previous systematic reviews [16] in that it only included studies with accelerometry-measured outcomes (ie, objectively measured), showed small treatment effects for both total PA (standardized mean difference 0.12, 95% CI 0.04-0.20, P<.01) and MVPA (standardized mean difference 0.16, 95% CI 0.08-0.24, P<.001) [17]. These results translate into approximately 4 minutes more walking or running per day, a modest increase with limited clinical significance.

Many existing interventions, including school-based interventions, are limited because they are resource intensive. Young people are considered a hard-to-reach group and typically have limited adherence or exposure to PA interventions [16]. Therefore, novel approaches are crucial to engage those who are insufficiently active to encourage good health [18]. Given young peoples’ ubiquitous use of mobile phones and increased digital literacy [19,20], interventions leveraging this technology may provide a promising intervention for this population.

Young people are early adopters of new consumer technology, such as smartphones and apps. Smartphone ownership among young people is on the rise, having increased from 23% in 2011 to 37% in 2012 in the United States, with similar distribution across ethnicity and family income [21]. Moreover, 58% of all US young people aged 12 to 17 years reported having downloaded apps compared to 34% of adults [22]. Likewise, digital divides have decreased between 2007 and 2013 in New Zealand [23], with 75% of high-income families reporting household access to a smartphone compared with 52% for those with lower income [24]. Alongside the increased ownership of smartphones, there has been a proliferation of apps in the major platform operators’ app stores, which currently offer more than 100,000 mHealth apps [25]. Consumers looking for help via their smartphone face an immense number of mHealth apps [26] and the majority of these (31%) are fitness apps [25]. This high engagement with mobile technology offers an ideal opportunity to leverage the benefits of mobile interventions for health (mHealth), including lower participant burden and flexibility [27].

The effectiveness of mHealth-delivered interventions to promote PA that use texting or short message service (SMS) [28,29], personal digital assistants (PDAs) [30], and apps in addition to other components [31] has been examined. Unlike SMS text messaging, which has a substantial body of literature supporting its use, more complex interventions that capitalize on the computational power of smartphones have only recently been developed. Particularly, few studies have used an entirely mHealth device-based approach to deliver health behavior change interventions [32]. Most [33,34] include other components, such as a website, individual face-to-face or group sessions, phone calls, print materials, or pedometers, making it impossible to tease out the specific effects of each component. Thus, at present, there is a lack of scientific evidence assessing publicly available apps to promote PA. Because these apps are commercially available to the public on app stores and because fitness apps are currently the most popular (78% of users in 2014 compared to 39% in 2013 [35]), it is important to determine whether popular/commonly downloaded apps are effective [36].

Therefore, the primary aim of the Apps for IMproving FITness (AIMFIT) trial was to evaluate the effectiveness of two popular commercially available smartphone apps [37] for improving CRF in young people aged 14 to 17 years compared to usual
behavior alone (the control). Secondary aims were to determine the effect on PA levels, enjoyment, psychological need satisfaction, and self-efficacy. Perceptions of usability and acceptability of the apps were also assessed.

Methods

Trial Design

A 3-arm parallel RCT was conducted in Auckland, New Zealand. Details on the rationale, design, and methods have been previously described [38]. Ethics approval was obtained from the University of Auckland Human Participants Ethics Committee (10054/2013) and the study was conducted according to the principles of the Declaration of Helsinki. There were no deviations to methods after trial commencement. All participants (and guardians for participants younger than 16 years) provided written informed consent. The trial was registered (Australian New Zealand Clinical Trials Registry: ACTRN12613001030763) in accordance with the requirements of the International Committee of Medical Journal Editors and reported in accordance with the CONSORT statement [39] and the CONSORT-EHEALTH extension [40] (see Multimedia Appendix 1). A brief description of the procedures is provided.

Participants

Participants were recruited through advertisements in electronic mailing lists, local newspapers, schools, and flyers posted in community locations. Consenting schools and churches allowed the researcher to present a brief outline of the study. Those interested provided contact details and their eligibility was assessed via telephone. If eligible, participant information and informed consent documentation were either mailed or emailed and participants were scheduled to attend a face-to-face baseline assessment at the university.

Eligible participants were aged 14 to 17 years, lived in Auckland, owned an iPod touch or smartphone running at least Android 2.2 or iOS 6.0, and were able to perform physical activities but were not achieving [41] the PA recommendations (ie, at least 60 minutes of MVPA each day). Exclusion criteria were a medical condition limiting ability to exercise safely, previous use of the apps of interest, and inability to comply with the study protocol. Only 1 child per household was eligible to take part.

Randomization

Participants were enrolled by author AD and were randomly assigned at a 1:1:1 ratio to 1 of 3 conditions. Stratified block randomization in variable blocks was used to maintain balance across gender, an important prognostic factor [24,22]. A biostatistician (YJ) prepared the randomization scheme in advance by using a computer-generated randomization table. Based on the randomization scheme, a research assistant prepared opaque-sealed envelopes containing group referral so that the researcher could not identify group assignment. The envelope was opened by each participant after completion of baseline assessment. Given the nature of the intervention, it was not possible to blind participants. However, allocation concealment was maintained up to the point of randomization.

Interventions

Commercially available apps targeting fitness were identified during previous work evaluating the most popular (ie, top-20 free and top-20 paid) downloaded apps in the Health and Fitness Category of the iTunes New Zealand store [37]. The control group was compared against 2 other groups, each using an app to promote fitness available on both of the major platforms’ app stores (iTunes and Google Play Store). The 2 intervention groups allowed comparison of an identical delivery approach (ie, stand-alone app), but with distinct design features: (1) use of an immersive app or (2) use of a nonimmersive app. Both apps consisted of a fully automated 8-week training program designed to improve fitness and ability to run 5 km; however, the immersive app featured a game-themed design whereby the training program was embedded with a story where the user is trained to collect supplies and protect a town from zombies. Effective self-regulatory behavior change techniques [43,44] underlie the apps, which allow the users to self-monitor and receive feedback on their training. The apps provided information on running and technique, audio instructions on how to perform the training components, and tracked and displayed progress throughout the program. Other features of the apps include the ability to work out with music on the device’s library and links to associated websites to interact with other users.

Participants randomized to the immersive app group received the Zombies, Run! 5K Training app developed by Six to Start with Naomi Alderman for iOS and the Android platform. It was released worldwide for iOS on October 2012. Even though data on number of downloads is not publicly available, the Google Play Store reports 100,000 to 500,000 installs of this app, and the Zombies, Run! community has more than 800,000 players worldwide [45]. Those randomized to the nonimmersive app group received the Get Running-Couch to 5k app, developed by Splendid Things and Benjohn Barnes for iOS and the Android platform. It was released worldwide for Android on July 2009 and the Google play Store reports 10,000 to 50,000 installs of this app [46]. While the trial was conducted, the apps cost NZ $2.45 to $4.19, depending on the app store. Given the commercialized nature of the apps, updates occurred during the trial; however, no major changes affected their content or the intervention. Videos detailing the features of the apps can be accessed via links on the preceding references.

Following randomization, the respective app was paid for and installed by AD on each participant’s mobile device and a short instruction on the features and settings of the app was given. Participants were encouraged to use their app 3 times per week and work their way through each of the workouts, but because this was a pragmatic study [47], access and usage was allowed to vary (ie, participants were able to use at their own pace, ad libitum, as it would occur on an “everyday life” setting). There were no cointerventions, no supplementary modes of delivery, nor usage of prompts (ie, emails, phone calls, SMS text message) to use the app.

Control Group

The control group was asked to continue with their usual physical activities for the duration of the study and was not
offered any information about increasing PA. Both apps were provided (free of charge) to participants after trial completion.

**Procedure**

Assessments were conducted at baseline and 8 weeks at the university by AD. Participants were assessed individually. At both time points, participants completed a field test of CRF (1 mile run/walk test), had their height and weight measured, self-reported their physical activity and related psychological variables, were given an Actigraph accelerometer to wear for the following 7 days (to provide an objective assessment of their free-living PA), and completed a booklet detailing their accelerometer use. AD collected the accelerometers and booklets from the participants’ homes (during the randomization visit that took place after the baseline assessment and at the last visit after the follow-up assessment). Follow-up assessments were not blinded. Participants received a NZ $10 gift card to a local shopping center for each visit to complete study measures (ie, maximum NZ $30 for 3 visits). The vouchers were not conditional on usage of the app; they were offered to compensate for participants’ time and encourage completion of study measures.

**Outcomes**

The primary outcome was CRF, assessed with the 1-mile run/walk test. Following the procedures outlined in the Fitnessgram test administration manual, participants were instructed to run and/or walk at their own pace until completing the distance in the shortest possible time [48]. The output, in seconds, was used to estimate VO$_{2peak}$ using a validated prediction equation [49].

Secondary outcomes included anthropometrics, self-reported PA and associated psychological variables, objectively measured PA, and self-reported acceptability and usability of the apps assessed via an exit survey conducted with intervention participants. A series of closed and open-ended questions were asked to determine features perceived as more and less acceptable as well as which features participants found more useful to support their fitness. Body weight (in kg, without shoes) was measured with a Salter scale to 1 decimal place. Height was measured to the nearest 0.1 cm with a Seca stadiometer. Two measurements were taken for each and the means were used for analysis. Body mass index (BMI) was calculated by using the standard equation (weight in kilograms/height in meters squared). BMI-for-age was calculated using the World Health Organization (WHO) growth standards macro [50].

Using instruments validated in this population, participants self-reported (1) physical activity using the Physical Activity Questionnaire for Adolescents (PAQ-A) [51,52], (2) perceived enjoyment using the Physical Activity Enjoyment Scale (PACES) [53,54], (3) perceived competence, autonomy, and relatedness using the Psychological Need Satisfaction in Exercise Scale (PNSE) [55], (4) self-efficacy using the Physical Activity Self-Efficacy Scale (PASES) [56,57], and (5) frequency of app utilization, acceptability, and usability of the app via an exit survey conducted with the participants in the intervention groups.

Participants were instructed to wear the accelerometer (Actigraph GT1M) on their right hip during waking hours for 7 days after each assessment, removing it when engaging in activities involving water and/or contact sports. A 10-second epoch was used and data were aggregated into minute intervals for subsequent processing. To determine valid wear time, periods of more than 60 minutes of consecutive zeroes and days with less than 600 minutes of valid records were removed before data analysis [58]. Participants were included in the analyses if they had 3 or more valid monitoring days [59]. Mean daily time (min) spent in sedentary and light-to-vigorous activities was calculated for each participant using the cut points proposed by Evenson et al to categorize intensities [60,61].

Adverse events were collected at each study visit or voluntarily reported by contacting the researcher. An adverse event was considered serious if it required hospitalization.

**Sample Size Calculation**

A total of 51 participants (17 per group) was estimated to provide 80% power and α=0.05 overall to detect a difference of 17 seconds in CRF, assuming a 15-second SD in time to complete the 1-mile run/walk test between each of the conditions compared to the control [62].

**Statistical Analyses**

Treatment evaluations were performed on the principle of intent-to-treat, including all randomized participants as allocated. Statistical analyses were performed with SAS version 9.4 software (SAS Institute, Cary, NC, USA). All statistical tests were 2-sided at a 5% significance level, with adjustment for multiple comparisons on the primary outcome. Analysis of covariance (ANCOVA) regression model was used to evaluate the main treatment effects on the primary outcome, adjusting for baseline measure and gender. Model-adjusted means, 95% confidence intervals, and P values were estimated for each group. Mean differences between groups were tested. A similar approach was used for secondary outcomes. Missing data on the primary outcome were imputed with the baseline value. Per-protocol analyses were conducted on those participants with complete baseline and follow-up data and self-reported adherence to the training program (ie, app used 3 times/week) following the same procedures as in the intention-to-treat analyses.

**Results**

**Overview**

Recruitment began October 2013 and finished in June 2014. The final follow-up visit was in September 2014. Figure 1 presents the participant flow diagram. Of 143 individuals screened, 51 eligible participants were randomized to the Zombies, Run immersive app intervention (n=17), the Get Running nonimmersive app intervention (n=16), or the control group (n=18). Baseline demographic characteristics of participants are presented in Table 1.

Participants had a mean age of 15.7 years (SD 1.2, range 14-17 years) and a BMI of 22.9 (SD 4.3) kg/m$^2$. The majority were NZ European (61%, 31/51), whereas 22% (11/51) were Pacific...
Islanders, and 57% (29/51) were female. Follow-up assessments at 8 weeks were completed for 17 (100%, 17/17) immersive app group participants, 15 (94%, 17/16) nonimmersive app group participants, and 17 (94%, 17/18) control group participants, which represents an overall retention rate of 96% (49/51) from baseline.

Table 1. Baseline demographic and clinical characteristics.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Zombies, Run n=17</th>
<th>Get Running n=16</th>
<th>Control n=18</th>
<th>Total N=51</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>15.78 (1.11)</td>
<td>15.69 (1.04)</td>
<td>15.55 (1.32)</td>
<td>15.67 (1.15)</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>8 (47)</td>
<td>6 (38)</td>
<td>8 (44)</td>
<td>22 (43)</td>
</tr>
<tr>
<td>Female</td>
<td>9 (53)</td>
<td>10 (63)</td>
<td>10 (56)</td>
<td>29 (57)</td>
</tr>
<tr>
<td>Ethnicity, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maori</td>
<td>3 (18)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>3 (6)</td>
</tr>
<tr>
<td>NZ European</td>
<td>9 (53)</td>
<td>9 (56)</td>
<td>13 (72)</td>
<td>31 (61)</td>
</tr>
<tr>
<td>Pacific</td>
<td>4 (24)</td>
<td>3 (19)</td>
<td>4 (22)</td>
<td>11 (22)</td>
</tr>
<tr>
<td>Asian</td>
<td>0 (0)</td>
<td>3 (19)</td>
<td>1 (6)</td>
<td>4 (8)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (6)</td>
<td>1 (6)</td>
<td>0 (0)</td>
<td>2 (4)</td>
</tr>
<tr>
<td>BMI (kg/m^2), mean (SD)</td>
<td>23.17 (3.60)</td>
<td>21.85 (3.14)</td>
<td>23.43 (5.56)</td>
<td>22.85 (4.25)</td>
</tr>
<tr>
<td>BMI-for-age^a (z-score), mean (SD)</td>
<td>0.77 (0.86)</td>
<td>0.36 (0.93)</td>
<td>0.64 (1.46)</td>
<td>0.60 (1.12)</td>
</tr>
<tr>
<td>Device, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>iPhone</td>
<td>8 (47)</td>
<td>6 (38)</td>
<td>11 (61)</td>
<td>25 (49)</td>
</tr>
<tr>
<td>Android</td>
<td>5 (29)</td>
<td>7 (44)</td>
<td>5 (28)</td>
<td>17 (33)</td>
</tr>
<tr>
<td>iPod Touch</td>
<td>4 (24)</td>
<td>3 (19)</td>
<td>2 (11)</td>
<td>9 (18)</td>
</tr>
<tr>
<td>Estimated VO2peak^b (mL/kg/min), mean (SD)</td>
<td>43.51 (6.11)</td>
<td>43.58 (5.47)</td>
<td>44.20 (6.95)</td>
<td>43.78 (6.12)</td>
</tr>
</tbody>
</table>

^a WHO growth reference.

^b Prediction equation from 1-mile run/walk test.
**Figure 1.** Flow diagram of the Apps for IMproving FITness (AIMFIT) trial. Those who were unable to complete the postintervention fitness assessment due to injury or sickness still completed self-reported outcomes and were included in all analyses.

**Primary Outcome**

Table 2 lists the effects of the app interventions on the time to complete the 1-mile walk/run fitness test and all secondary outcomes at 8 weeks. On average, time to complete the fitness test decreased in both app groups, but there were no statistically significant differences observed between the intervention groups and the control (immersive app group: adjusted mean difference –28.4 sec, 95% CI –66.5 to 9.8, \( P=0.20 \); nonimmersive app group: adjusted mean difference –24.7 sec, 95% CI –63.5 to 14.2 sec, \( P=0.32 \)). To facilitate interpretation of test scores, estimated \( VO_2_{peak} \) [49] at 8 weeks was 44.09, 44.59, and 43.44 mL/kg/min for the immersive app, nonimmersive app, and control groups, respectively.
### Table 2. Treatment effects at 8 weeks.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Zombies, Run (1), mean (SD)</th>
<th>Get Running (2), mean (SD)</th>
<th>Control (3), mean (SD)</th>
<th>Adjusted differencea (95% CI) at 8 weeks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline (min) 8 week</td>
<td>Baseline (min) 8 week</td>
<td>Baseline (min) 8 week</td>
<td>1 vs 3 ( P ) 2 vs 3 ( P )</td>
</tr>
<tr>
<td>Time to complete 1-mile walk/run (sec)</td>
<td>574.06 (145.68)</td>
<td>560.06 (139.27)</td>
<td>586.56 (129.74)</td>
<td>576.75 (147.91)</td>
</tr>
<tr>
<td>PAQ-A</td>
<td>2.20 (0.66)</td>
<td>2.27 (0.53)</td>
<td>2.09 (0.73)</td>
<td>2.31 (0.74)</td>
</tr>
<tr>
<td>PACES</td>
<td>4.08 (0.47)</td>
<td>4.00 (0.46)</td>
<td>3.99 (0.46)</td>
<td>3.85 (0.56)</td>
</tr>
<tr>
<td>PNSES</td>
<td>4.52 (0.69)</td>
<td>4.49 (0.78)</td>
<td>4.48 (0.89)</td>
<td>4.56 (0.56)</td>
</tr>
<tr>
<td>Competence</td>
<td>4.27 (0.83)</td>
<td>4.24 (0.94)</td>
<td>4.25 (1.09)</td>
<td>4.32 (0.94)</td>
</tr>
<tr>
<td>Autonomy</td>
<td>4.87 (0.98)</td>
<td>4.92 (0.94)</td>
<td>4.94 (0.78)</td>
<td>5.21 (0.56)</td>
</tr>
<tr>
<td>Relatedness</td>
<td>4.40 (0.79)</td>
<td>4.31 (0.87)</td>
<td>4.26 (1.15)</td>
<td>4.14 (0.91)</td>
</tr>
<tr>
<td>PASES</td>
<td>2.50 (0.32)</td>
<td>2.44 (0.40)</td>
<td>2.38 (0.47)</td>
<td>2.43 (0.51)</td>
</tr>
<tr>
<td>Average daily valid use (min)</td>
<td>801.16 (87.85)</td>
<td>784.43 (93.69)</td>
<td>819.39 (85.45)</td>
<td>831.08 (78.74)</td>
</tr>
<tr>
<td>Average daily activity counts (counts/min)</td>
<td>354.27 (123.75)</td>
<td>341.21 (146.22)</td>
<td>269.97 (82.72)</td>
<td>270.20 (84.77)</td>
</tr>
<tr>
<td>Average daily time spent in sedentary activities (min)</td>
<td>526.63 (106.37)</td>
<td>535.22 (113.90)</td>
<td>553.64 (96.80)</td>
<td>570.63 (95.58)</td>
</tr>
<tr>
<td>Average daily time spent in light PA (min)</td>
<td>237.41 (73.32)</td>
<td>216.17 (67.46)</td>
<td>244.46 (61.97)</td>
<td>237.11 (58.45)</td>
</tr>
<tr>
<td>Average daily time spent in moderate PA (min)</td>
<td>30.28 (14.28)</td>
<td>25.07 (12.83)</td>
<td>17.03 (7.81)</td>
<td>18.03 (10.20)</td>
</tr>
<tr>
<td>Average daily time spent in vigorous PA (min)</td>
<td>6.84 (6.03)</td>
<td>7.97 (9.16)</td>
<td>4.26 (4.21)</td>
<td>5.30 (4.34)</td>
</tr>
<tr>
<td>Average daily time spent in MVPA (min)</td>
<td>37.12 (16.84)</td>
<td>33.04 (20.61)</td>
<td>21.29 (11.25)</td>
<td>23.34 (14.04)</td>
</tr>
</tbody>
</table>

a Adjusted for baseline, gender, and multiple comparisons.

**Secondary Outcomes**

No intervention effects were found for self-reported secondary outcomes of physical activity (PAQ-A; immersive app group: adjusted mean difference 0.14, 95% CI –0.26 to 0.54, \( P = .78 \); nonimmersive app group: adjusted mean difference 0.23, 95% CI –0.18 to 0.64, \( P = .42 \) or its predictors of perceived enjoyment (PACES), perceived competence, autonomy, and relatedness (PNSES), or self-efficacy (PASES) (see Table 2 for all outcomes).

For accelerometry, 48 of 51 (94%) participants provided valid data for analysis at baseline, whereas compliance with wearing the device slightly decreased at postintervention (46/51, 90%). Group assignment did not have a significant effect on overall activity (ie, mean counts per minute) or mean daily time spent in MVPA. Compared to the control group, mean baseline daily time spent in MVPA-, gender-, and multiple comparisons-adjusted time in MVPA difference was 1.74 min (95% CI –11.45 to 14.93, \( P = .98 \)) and –1.82 min (95% CI –16.00 to 12.36, \( P = .99 \)) for the immersive and nonimmersive app groups, respectively.

A total of 6 adverse events (1 serious) were reported in 6 participants, 4 of which were in the control group (ankle injury events, lower back pain, and hospitalization because of tonsils removal) and 1 in each of the intervention groups (ankle injury events). None of the adverse events were deemed related to the study intervention.
Approximately two-thirds of participants in the intervention groups reported using the app either 2 (10/32, 31%) or 3 times per week (10/32, 31%), whereas 8 of 32 (25%) only used it 1 time per week (see Multimedia Appendix 2). No differences were evident on timing of use (ie, weekday, weekend, morning, afternoon, evening). Apps were mostly used outdoors (eg, street, park vs gym, home treadmill) and while alone (n=13 alone, n=7 with friend, and n=9 with family).

For the app Zombies, Run!, the features mostly used by participants were the “workout mission tasks” (n=14) and “story and run log” of completed workouts (n=10), whereas social networking features (“share my runs”: n=0; “ZombieLink account”: n=3) were seldom or never used. Results were similar when participants reported the features they liked (“workout mission tasks”: n=14; “story and run log”: n=8) and disliked (“share my runs”: n=5; “ZombieLink account”: n=5).

For the app Get Running, the feature mostly used by participants was the description of the “week-runs” (n=13), whereas only 1 participant reported using the social networking feature “status updates.” The description of the “week-runs” was also the feature participants predominantly liked (n=11), whereas the main feature disliked was the “status updates” (n=3).

Regardless of the app used, similar themes emerged when participants reported their willingness and motives to continue using their app after study participation. Those willing to continue stated personal benefits (eg, “It will help me to build my fitness”, “Because I can improve how far I run”) and app-related motives (eg, “A fun way to get fit”, “Because it is an enjoyable alternative to exercise”). For those unwilling to continue, “not enough time” was the most common barrier, followed by lack of interest (eg, “I didn’t find the app engaging enough”). The nonimmersive app received less positive feedback around motivational aspects (eg, “Using the app became too tedious”).

Overall, participants perceived the layout of the apps and menus as well structured and “straightforward” to use. Being able to receive clear instructions (eg, “Tells me what to do and when”), listen to their own music during the training sessions, task difficulty increasing gradually, and encouragement provided were features highlighted as useful to support participants’ fitness. Some also considered it helpful if the app allowed choosing between different goals and activities (eg, “I prefer to run to my own goals”). For the immersive app, the storyline (“The back story made it interesting”) and the ability to track progress (ie, app used the device’s Global Positioning System [GPS] and/or accelerometer to log distance) (eg, “It records the distance you ran and your time so you are able to view it for next time and compare”) were also reported as important features.

The majority of participants (21/32, 66%) had no prior experience of using their smartphone for PA purposes. Examples of prior experience included listening to music while engaging in PA or previous use of free apps (eg, MyFitnessPal). Overall, 81% (26/32) were interested in trying different PA-promoting apps in the future.

In prespecified per-protocol analyses (ie, the app was used 3 times/week), there were statistically significant differences observed on the primary outcome between the nonimmersive app group and the control (adjusted mean difference −79.39 sec, 95% CI −133.01 to −25.77, P=.003). Time to complete the fitness test also decreased for the immersive app group compared to the control, with the difference not meeting statistical significance (adjusted mean difference −55.29 sec, 95% CI −111.46 to 0.88, P=.06). No intervention effects were found for self-reported secondary outcomes of PA, its predictors, or accelerometry.

Discussion

Principal Findings

This is the first randomized trial comparing the effects of a stand-alone immersive mobile app and a nonimmersive app on CRF, PA levels, and its predictors in young people. Key findings were that fitness improved in both app groups, but these did not significantly differ from the control. Despite the availability of readily available commercial apps to improve health behavior, these findings suggest that, compared to usual care, no major improvements were found for these 2 top downloaded apps.

The small increases in fitness in the present trial (0.6 to 1.0 mL/kg/min) were lower than those observed in a Cochrane review of school-based PA interventions, which found increases of 1.6 to 3.7 mL/kg/min in VO2peak [14]. Similarly, Baquet et al [63], who reviewed the effects of training programs in young people’s aerobic fitness, showed improvements on average of 5% to 6% in VO2peak, whereas fitness only improved 1.3% to 2.3% in this study. The training programs of the apps were consistent with the frequency and duration of programs shown to improve VO2peak [63] (ie, incorporating 3 sessions per week, each lasting 30 minutes to 1 hour). However, use of the apps during the AIMFIT study was ad libitum and not closely monitored, which is how app use would likely occur in real life. Smaller fitness improvements in the current study can most likely be attributed to lower adherence to app use (only 31.3% reported using 3 times/week) or due to participants not having exercised at a sufficient intensity to produce larger fitness effects [63], as observed by no changes in accelerometer-measured MVPA.

Although apps have the potential to increase the reach of health behavior change interventions, our results mirror recent research highlighting that only some participants will consistently use an offered app (approximately 20%) [64] or sustain use over time (approximately 3%) [65]. App design, features, and usability are important factors that impact users’ engagement with an app. Although in AIMFIT both app groups produced comparable fitness effects, the design and features of the immersive app received more positive feedback (and no dropout) and, therefore, these aspects should be considered for future app development. Although typical app use is intermittent [65] and unlikely to induce the more intensive type of immersion observed in other videogame play, incorporating gaming elements that entangle the virtual world (ie, leisure) with actual
experience (ie, exertion) could potentially increase user engagement and sustained use.

An important consideration of app content is whether or not they incorporate behavior change techniques (BCTs). Further, modeling, providing consequences for behavior, providing information on others’ approval, prompting intention formation, self-monitoring, and a behavioral contract were identified as effective BCTs for increasing PA in young people in a recent meta-regression [66]. A notable exception is providing instruction, which was associated with decreased effectiveness of PA interventions in young people, but has been shown to be highly present in top-ranked PA apps [37], including the ones hereby investigated. Providing instructions and not incorporating other effective BCTs or incorrectly applying BCTs (ie, not considering the parameters for them to be effective) may have undermined app efficacy.

Consistent with the primary outcome findings, we found no changes in any of the measured psychological variables. Fulfillment of the 3 basic psychological needs (ie, autonomy, competence, and relatedness)—key elements in the development of intrinsic motivation required to drive behavioral change [67]—did not change. This could be explained by the nature of the training programs, which were based on provision of instructions and may have been perceived as too prescriptive and controlling. Perhaps the alternatives provided by the apps (ie, walking, jogging, sprinting, running) lacked choice and encouragement of the participants’ initiative and, therefore, did not support their autonomy. In addition, skill acquisition and feelings of competence did not change as a function of the intervention. This suggests the BCTs in the apps were insufficient to manipulate these variables or that app use was not intensive enough to evoke change. Further, the apps incorporate features that could capitalize on relatedness and social support, such as online communities, but some of these features load outside of the apps (eg, ZombieLink), which may have contributed to poor use of such features. Interestingly, enjoyment, which is associated with adherence to PA programs, did not change.

Strengths and Limitations

A major strength of AIMFIT was the use of a RCT design to determine the effectiveness of 2 off-the-shelf commercially available interventions. We chose a pragmatic approach in which participants used their own device and apps were used ad libitum. Contact with participants was minimal, which reflects app use in a real-world context and therefore increases the generalizability of the findings. Moreover, the primary outcome was assessed objectively with a valid and reliable measurement, as well as PA via accelerometry, which adds to the study’s internal validity. Unlike the commonly observed high attrition in eHealth and mHealth interventions, follow-up assessments were completed for 96% of participants, which represents a high retention rate. Further, using paid apps instead of free ones has likely reduced the chances of contamination and/or cointerventions. Although the accessibility of these apps on the app stores could have led participants to download them, our study sample was a young population and the apps under investigation required purchase. Although inexpensive, purchasing an app on an app store requires a credit card, which young people do not typically own.

The major limitation of this study was its low statistical power and small sample size. We based our sample size calculation on a smaller standard deviation of the primary outcome than the actual standard deviation observed, which meant the power of the trial was smaller than 80%. We used readily available apps and consequently were limited to the decisions made by the app developers on content, duration of the program, and design features. This also meant that we were unable to access data on app utilization (eg, menus accessed in the app). Further, the relatively short duration of the programs precluded investigation of long-term effects or sustainability. This study also highlights that the peer-reviewed literature will always lag behind consumer technology life cycles because during the lifetime of this study innovative apps were developed at a rate that far outpaced our capacity to test them.

Implications

Among app users, fitness apps are the most popular (78% users in 2014 compared to 39% in 2013) [35]. However, off-the-shelf common commercial mobile apps used as a stand-alone approach to change fitness and promote PA in insufficiently active individuals do not appear effective. Future use as part of a multifaceted approach to increase fitness and to promote PA may be useful. The current one-size-fits-all approach of most common commercial mobile apps is limiting for both users and researchers. More tailored approaches, which are dynamic and responsive to changes in PA behavior over time, are required to improve the ease of use of apps, user engagement, and the apps’ sustained use.

Conclusions

Readily available commercial apps as a stand-alone intervention to improve fitness and increase PA in young people did not increase CRF compared to usual care. Given that smartphone technology appears to resonate with young people and that this type of self-guided intervention has the potential to increase reach at a low cost, this may be best suited as part of a multicomponent intervention, providing additional support and encouragement to the participants (eg, maintenance phases).

Acknowledgments

We would like to acknowledge the trial participants, guardians, and schools who gave time to be in the study. This is an investigator-initiated study supported by internal funding from the University of Auckland Postgraduate Research Student Support. AD is supported by a Foundation for Science and Technology scholarship (FCT-Portugal SFRH/BD/95762/2013). FCT had no role in experimental design, data collection, or manuscript preparation.
Authors' Contributions
AD and RM conceived the study, participated in its design and coordination, and helped draft the manuscript. YJ and RW participated in the design of the study and helped draft the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
CONSORT-EHEALTH checklist V1.6.2 [40].

[PDF File (Adobe PDF File), 146KB - jmir_v17i8e210_app1.pdf]

Multimedia Appendix 2
Supplementary word file containing graphics with exit survey questions for intervention participants.

[PDF File (Adobe PDF File), 417KB - jmir_v17i8e210_app2.pdf]

References


46. Splendid Things. Get Running (Couch to 5k) URL: http://splendid-things.co.uk/getrunning/ [accessed 2015-04-21] [WebCite Cache ID 6XxqSPOZf]


Abbreviations

BCT: behavior change technique
BMI: body mass index
CRF: cardiorespiratory fitness
MVPA: moderate-to-vigorous physical activity
PA: physical activity
PACES: Physical Activity Enjoyment Scale
PAQ-A: Physical Activity Questionnaire for Adolescents
PASES: Physical Activity Self-Efficacy Scale
PNSES: Psychological Need Satisfaction in Exercise Scale
RCT: randomized controlled trial
SMS: short message service

©Artur Direito, Yannan Jiang, Robyn Whittaker, Ralph Maddison. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 27.08.2015. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.