Original Papers

A Novel Clinician-Orchestrated Virtual Reality Platform for Distraction During Pediatric Intravenous Procedures in Children With Hemophilia: Randomized Controlled Trial (e10902)
Amy Dunn, Jeremy Patterson, Charmaine Biega, Alice Grishchenko, John Luna, Joseph Stanek, Robert Strouse................................................................. 2

Young People’s Knowledge of Antibiotics and Vaccinations and Increasing This Knowledge Through Gaming: Mixed-Methods Study Using e-Bug (e10915)
Charlotte Eley, Vicki Young, Catherine Hayes, Neville Verlander, Cliodna McNulty................................................................. 13

Serious Games in Surgical Medical Education: A Virtual Emergency Department as a Tool for Teaching Clinical Reasoning to Medical Students (e13028)
Seung-Hun Chon, Ferdinand Timmermann, Thomas Dratsch, Nikolai Schuelper, Patrick Plum, Felix Berlth, Rabi Datta, Christoph Schramm, Stefan Haneder, Martin Späth, Martin Dübbers, Julia Kleinert, Tobias Raupach, Christiane Bruns, Robert Kleinert................................................................. 28

Energy Expenditure and Enjoyment During Active Video Gaming Using an Adapted Wii Fit Balance Board in Adults with Physical Disabilities: Observational Study (e11326)
Laurie Malone, Mohanraj Thirumalai, Sangeetha Padalabalanarayanan, Whitney Neal, Sean Bowman, Tapan Mehta. ................................................................. 39

Effectiveness of a Behavior Change Technique-Based Smartphone Game to Improve Intrinsic Motivation and Physical Activity Adherence in Patients With Type 2 Diabetes: Randomized Controlled Trial (e11444)
Christoph Höchsmann, Denis Infanger, Christopher Klenk, Karsten Königstein, Steffen Walz, Arno Schmidt-Trucksäss................................................................. 54

Active Video Games for Rehabilitation in Respiratory Conditions: Systematic Review and Meta-Analysis (e10116)
Joshua Simmich, Anthony Deacon, Trevor Russell................................................................. 66

Creating a Theoretically Grounded Gaming App to Increase Adherence to Pre-Exposure Prophylaxis: Lessons From the Development of the Viral Combat Mobile Phone Game (e11861)
Laura Whiteley, Leandro Mena, Lacey Craker, Meredith Healy, Larry Brown................................................................. 81
A Novel Clinician-Orchestrated Virtual Reality Platform for Distraction During Pediatric Intravenous Procedures in Children With Hemophilia: Randomized Controlled Trial

Amy Dunn¹, MD; Jeremy Patterson²,³, BFA; Charmaine F Biega⁴, RN; Alice Grishchenko³, MFA; John Luna³, MFA; Joseph R Stanek⁴, MS; Robert Strouse², MFA

¹Nationwide Children's Hospital, Division of Hematology, Oncology and BMT, The Ohio State University, Columbus, OH, United States
²The Research Institute at Nationwide Children's Hospital, Columbus, OH, United States
³The Ohio State University Advanced Computing Center for the Arts and Design, Columbus, OH, United States
⁴Nationwide Children's Hospital Division of Hematology/Oncology/BMT, Columbus, OH, United States

Corresponding Author:
Amy Dunn, MD
Nationwide Children's Hospital
Division of Hematology, Oncology and BMT
The Ohio State University
700 Children's Drive
Columbus, OH, 43205
United States
Phone: 1 614 722 3564
Fax: 1 614 722 3669
Email: amy.dunn@nationwidechildrens.org

Abstract

Background: Needles are frequently required for routine medical procedures. Children with severe hemophilia require intensive intravenous (IV) therapy to treat and prevent life-threatening bleeding and undergo hundreds of IV procedures. Fear of needle-related procedures may lead to avoidance of future health care and poor clinical outcomes. Virtual reality (VR) is a promising distraction technique during procedures, but barriers to commercially available VR platforms for pediatric health care purposes have prevented widespread use.

Objective: We hypothesized that we could create a VR platform that would be used for pediatric hemophilia care, allow clinician orchestration, and be safe and feasible to use for distraction during IV procedures performed as part of complex health care.

Methods: We created a VR platform comprising wireless, adjustable, disposable headsets and a suite of remotely orchestrated VR games. The platform was customized for a pediatric hemophilia population that required hands-free navigation to allow access to a child’s hands or arms for procedures. A hemophilia nurse observing the procedure performed orchestration. The primary endpoint of the trial was safety. Preliminary feasibility and usability of the platform were assessed in a single-center, randomized clinical trial from June to December 2016. Participants were children with hemophilia aged 6-18 years. After obtaining informed consent, 25 patients were enrolled and randomized. Each subject, 1 caregiver, and 1 hemophilia nurse orchestrator assessed the degree of preprocedural nervousness or anxiety with an anchored, combined modified visual analog (VAS)/FACES scale. Each participant then underwent a timed IV procedure with either VR or standard of care (SOC) distraction. Each rater assessed the distraction methods using the VAS/FACES scale at the completion of the IV procedure, with questions targeting usability, engagement, impact on procedural anxiety, impact on procedural pain, and likability of the distraction technique. Participants, caregivers, and nurses also rated how much they would like to use VR for future procedures. To compare the length of procedure time between the groups, Mann-Whitney test was used.

Results: Of the 25 enrolled children, 24 were included in the primary analysis. No safety concerns or VR sickness occurred. The median procedure time was 10 (range 1-31) minutes in the VR group and was comparable to 9 (range 3-20) minutes in the SOC group (P=.76). Participants in both the groups reported a positive influence of distraction on procedural anxiety and pain. Overall, in 80% (34/45) of the VR evaluations, children, caregivers, and nurses reported that they would like to use VR for future procedures.
Conclusions: We demonstrated that an orchestrated, VR environment could be developed and safely used during pediatric hemophilia care for distraction during IV interventions. This platform has the potential to improve patient experience during medical procedures.

Trial Registration: ClinicalTrials.gov NCT03507582; https://clinicaltrials.gov/ct2/show/NCT03507582 (Archived by WebCite at http://www.webcitation.org/73G75upA3)

(JMIR Serious Games 2019;7(1):e10902) doi:10.2196/10902

KEYWORDS
anxiety; distraction; hemophilia; intravenous; mobile phone; needle; pediatric; virtual reality

Introduction
Medical procedures involving needles induce fear in most people, especially children [1-6]. However, treatment for some medical conditions requires frequent needle sticks. Balancing the treatment needs for life-threatening conditions with the fear and anxiety imposed by the mechanism of these treatments is an important challenge for the medical community. In addition to the load needle sticks impose on patients and families, the medical system and medical providers also bear burden and expenses related to needles. One of the conditions that requires the most needle-intensive care is hemophilia. Both hemophilia A (HA) and B (HB) are severe congenital bleeding disorders. Without intravenous (IV) infusions of clotting factor concentrates, children with hemophilia experience life- and limb-threatening bleeding. Most children with hemophilia begin routine IV infusions of factor concentrate between the age of 1 and 3 years and continue infusions 2-3 times per week for life. This translates to thousands of necessary needle sticks over the course of their lives and a high treatment burden related to needles [7]. Needle fear related to IV procedures, particularly in children with hemophilia, can progress to blood-injection-injury phobia (needle phobia), treatment avoidance, and poor adherence, all of which can contribute to poor medical outcomes [8]. Psychological interventions, such as distraction and hypnosis, can reduce needle-related pain and distress in pediatrics [9]. A recent systematic review and meta-analysis demonstrated that distraction could lead to reductions in child- and observer-reported pain and distress during needle procedures. However, among the major elements of distraction, the relative importance of (1) no or low-tech versus high-tech distracters; (2) child engaging versus passive distracters; (3) degree of adult involvement; and (4) the availability of child choice in the success of these distracters is unclear [10]. Specific virtual reality (VR)-based distractions have shown promise in clinical settings, such as dentistry, IV placement, occupational or physical therapy, and burn care [9,11-24]. In particular, VR approaches have demonstrated decreased pain associated with dental procedures, physical therapy, and burn dressing changes. Although various consumer-grade VR technologies are available, most carry barriers that prevent clinical implementation [11-18,20-24]. Cost is another barrier, with the top-selling commercial unit having an approximate cost of US $500. This includes the head-mounted display (HMD), wireless controllers, and proximity sensors, but not a personal computer capable of running VR content, with a cost of approximately US $2000. On the low end, VR content is available for smartphones and low-cost HMDs, at around US $800 for the smart device and approximately US $35 for HMD. In either case, these estimates do not include the cost of content or technical support to configure and maintain the systems. Additionally, the unit is tethered to the personal computer and, as a result, cannot easily move between locations. Infection concerns exist with nondisposable headsets particularly in a pediatric setting, such as a hematology or oncology clinic. Additionally, most commercial VR platforms do not allow a medical provider to direct or modify the VR environment in response to patient distress and also focus on the experience of patients and not of caregivers.

Despite the barriers to VR implementation described above, we hypothesized that a children’s hospital-based team with expertise in hemophilia patient care and user experience technology could create a safe and clinically feasible VR ecosystem that met the needs of patients, caregivers, and medical providers during IV procedures. Assessing safety and testing the feasibility of the designed ecosystem in a complex clinic environment were the goals of the project. A comprehensive hemophilia clinic visit was selected to test the safety, feasibility, and likeability of the ecosystem. During these visits, patients see multiple providers in addition to having IV procedures, so the efficiency of clinic flow is vitally important. We specifically designed the ecosystem for children with hemophilia because patients with hemophilia need a frequent distraction from needles safe and successful integration into a complex clinic environment would demonstrate the likelihood of feasibility in less complex situations.

Methods
Virtual Reality Platform Design
The VR platform design team consisted of the Nationwide Children’s Hospital Hemophilia Treatment Center (HTC) staff and the user experience technology team. The HTC staff included a hematologist, nurse clinicians, a social worker, and a psychologist. The user experience team included a project lead, industrial designer, and 2 game designers, 1 of which also served as the illustrator and visual designer. The team designed the platform for boys and girls aged 6-18 years with HA or HB. Ideal features were identified through recurring meetings, which included observation of HTC visits and IV procedures. Specific platform design needs were identified. First, the cost of the overall system needed to be low. The system needed to be technically easy to implement and maintain for the nursing team. Games needed to be pediatric friendly but of sufficient
quality to maintain the interest of patients and families. Additionally, the video content needed to limit the possibility of VR sickness. Importantly, the majority of IV infusions require that the medical team has easy access to the hands and arms of a patient. This meant that game control mechanics needed to be hands free and help ensure the patient remained still. Because medical care is fluid and each patient experience is unique, the design team needed to allow the medical team to orchestrate the patient experience in the real-time to meet individual needs. We also sought ways to include caregivers in the game experience because parental anxiety around IV procedures can contribute to poor outcomes, and family-centered care is an important tenet of pediatric care. The team was also mindful of the need of the system to minimize the risk of transmitting infections agents from any game component. Each element of the ecosystem—the HMD, games, and the orchestration—was refined with input from volunteer pediatric hemophilia nurse clinicians, pediatric hematologists, and children without hemophilia. The ecosystem components were also demonstrated at hemophilia community events, and patients with hemophilia and their families gave feedback to the design team. The final VR intervention used for the clinical trial consisted of custom, cordless, multisized, disposable headsets, which enabled the use of VR through iPod Touch, and immersive custom games with hands-free navigation. Navigation techniques included head glances and breath. The mechanism for nurse orchestration (Figure 1; Multimedia Appendix 1) was software running on an iPad dashboard that wirelessly communicated to iPod Touch. The orchestration dashboard offered a suite of tools allowing a nurse to respond to patient needs by deploying mini-games or providing relocation to a new setting in the VR ecosystem. The dashboard also allowed parents to monitor their child’s progress. Moreover, an expert practitioner could observe patient state and, upon noticing stress, could trigger an intervention to act as a distraction method. The system’s communication platform utilized an internal server as a relay between the dashboard and the iOS device, which ran VR activities. All signal messages passed through the intermediary system and were logged and tracked against the identifier for the orchestrated session and the type of orchestration command. Message types recorded included commands to enter or exit a mini-game, transport the patient to a virtual location, connect to a paired iOS device, adjust the volume, open the video feed from the iOS device’s camera, and calibration of the headset units for fit and viewing preference. After finalization of the platform, safety and feasibility of platform integration into a routine comprehensive HTC visit were explored. Comprehensive visits included care from multiple providers during a specified timeframe, and anything that impeded clinic flow was judged not feasible. The complexity of the care for patients with hemophilia is reflected in the number and diversity of providers available in each HTC clinic. Patients are seen for comprehensive care visits once or twice per week and during those visits are seen by numerous providers based on individual needs. This can include hematologists, nurse clinicians, research nurses, psychologists, social workers, physical therapists, nutritionists, orthopedists, dentists, genetic counselors, radiologists, advocacy coordinators, and phlebotomy staff. For the pilot study, the length of the IV procedure time was chosen as a surrogate marker of feasibility.

**Study Design, Patients, and Randomization**

A prospective, monocentric, unblinded, randomized clinical trial was performed (Clinical Trials.gov NCT03507582). Multimedia Appendix 1 provides details of the trial protocol. After Institutional Review Board approval, children with HA or HB, aged 6-18 years, who were cared for in our HTC were screened for eligibility. Eligible patients had to be seen for a comprehensive HTC visit defined as seeing ≥3 specialty hemophilia providers during one visit and have a clinically indicated IV laboratory draw or factor infusion. Patients seen between July 1 and December 31, 2016, were eligible for participation if they and their parents or guardians could speak and write in English. Patients who were unable to use the VR equipment (visual, cognitive, or hearing impairment that precluded engagement with the VR environment), who had a known history of severe motion sickness, or uncontrolled seizures were not eligible (n=1). Those who met screening criteria were approached for participation (n=25). A hemophilia nurse obtained paper informed consent and age-appropriate assent. Enrolled patients were block randomized by computer-generated random allocation to study groups using a prespecified seed. The nurse performing randomization was blinded to group allocation. Recruited participants were randomly assigned using a 2:1 ratio to VR or standard of care (SOC) distraction with a stratified block design to maintain age strata of 12 patients aged 6-12 years and 12 patients aged 13-18 years. SOC distraction was any technique routinely available in the hematology clinic that did not include VR (ie, smart devices, bubbles, and videos).

**Study Intervention**

All enrolled patients had a clinically indicated IV procedure. A procedure timer and survey instruments were incorporated into the orchestration iPad. A modified visual analog (VAS)/FACES (Figure 2) rating scale was used to assess the distraction techniques. The scale did not undergo specific psychometric testing; however, the FACES scale is routinely used in our hemophilia clinic. A unipolar, horizontal scale was employed to increase understandability, uniformity, and sensitivity [25-27]. All participants, caregivers, and nurses (raters) were educated in how to use the VAS/FACES scale prior to beginning the study procedure by sliding a bar below the scale picture in response to an anchored question. Raters were not blinded to participant’s distraction group, and each rater answered the study questions independently. Immediately before the IV procedure, children in the VR group were introduced to the headset, game options, and navigation techniques. Then, before being positioned for the procedure, each patient, a caregiver with them for the procedure, and a hemophilia nurse assessed the participant’s level of preprocedural anxiety using a (VAS)/FACES scale answering the question: How worried or nervous is or are you or your child or your patient about the IV procedure? Sliding the bar to the left (toward 0) represented a low degree of preprocedural worry or nervousness and to the right (toward 100) represented the highest degree of nervousness or worry. The terms nervous and worry were substituted for anxiety owing to the young age of many participants. After being positioned for the procedure, patients in the SOC group used whatever distraction they preferred, whereas patients in

https://games.jmir.org/2019/1/e10902/
the VR group put on the headset, launched the game of their choice, and explored the VR environments. The nurse clinician was located next to the patient and deployed VR orchestration tools as she deemed necessary. The guardian was present throughout in both the groups. The study methodology was not altered during the trial. No revisions to the VR platform were performed or required during the trial.

**Primary Outcome**

The primary aim of the study was to assess safety concerns and feasibility of integration of the VR platform into a comprehensive HTC visit. Safety concerns included discomfort from HMD, infection, and VR sickness. Barriers to feasibility that were assessed included technical issues with the set-up or orchestration. The primary surrogate marker of feasibility was the duration of the IV procedure in both study groups. Procedure time was the length of time from the moment a patient was positioned for the procedure to the completion of the procedure, and the primary and secondary outcomes did not change during the course of the trial. No VR platform performance issues or unexpected events related to the platform were encountered during the trial.

*Figure 1.* Manual orchestration of child engaging events in a virtual reality environment from a connected virtual digital interface from an embedded viewpoint.

*Figure 2.* An example of the modified visual analog/FACES scale. VR: virtual reality.

**Question:** How was it to use the VR equipment?

Easy

![Easy](image)

0

Hard

![Hard](image)

100

50
Secondary Outcomes

Effectiveness of the Distraction Technique

Both patient groups assessed the effectiveness of their distraction techniques following the procedure by answering 3 questions as follows. (1) “Did the distraction technique keep you/your child/your patient engaged?” The anchors were 0=It really kept them engaged and 100=It really did not keep them engaged; (2) “Do you think the distraction technique(s) changed you/your child/your patient’s nervousness/anxiety level during the IV procedure?” The anchors were 0=it decreased nervousness/anxiety a lot and 100=it increased nervousness/anxiety a lot; (3) “How did the distraction technique(s) affect pain during your/your child/your patient’s IV procedure?” The anchors were 0=it made pain a lot better and 100=it made pain a lot worse.

Usability and Likeability

For patients randomized to the VR arm, data on the use of the VR equipment were recorded by the nurse orchestrator at the end of each procedure. Data were categorized if a participant wore the VR equipment: (1) during the entire procedure; (2) part of the procedure; or (3) only prior to the procedure. In addition, participants were asked to rate the usability of the VR equipment using the VAS/FACES scale to answer “How easy was it for you/your child/your patient to use the VR equipment?” A score of 0 represented “really easy to use,” and a score of 100 correlated with “really hard to use.” Lastly, participants were asked to use the VAS/FACES scale to assess the VR likeability by answering “How much would you/your child/your patient like to use VR for future IV procedures?” A score of 0 equated to “they would really like to use VR again” and 100 meant “they would really not like to use it again.”

Statistical Analysis

This was a pilot feasibility study in a rare population. With insufficient background data on the overall feasibility of using VR technology in a clinical setting, a sample of 24 patients was selected to collect pilot data on the safety and usability, logistical issues of implementation, and to assess the durability and adaptability in the design of the equipment. The study was designed to be randomized and include a control SOC group to perform a preliminary test on the hypothesis that the length of time for IV procedures would be similar between the groups, but the sample size did not allow testing for equivalence. The justifications for this sample size were based on the rationale about feasibility, obtaining adequate precision on numerical estimates, and regulatory considerations [19]. This sample size allowed for evaluating the potential utility of the VR distraction technique on a wide range of ages and would result in parameter estimates that would aid in adequately powering future studies that would directly assess the benefits of VR in clinical settings. Clinical data, including survey instrument data and adverse events, were entered into a hospital-compliant internet data entry system (REDcap) that included password protection and internal quality checks. Demographic data were captured from the patients’ electronic medical records. All demographic and safety data were described using summary statistics. To evaluate the hypothesis that the procedure time would be similar between the intervention and SOC groups, Mann-Whitney test was used. The length of IV procedure time was summarized by presenting mean and corresponding 95% CI as well as median and range. Kruskal-Wallis tests were applied to compare VAS/FACES scores given by the 3 rater groups as the secondary aim. Scores for the rater groups were summarized with medians and ranges. Statistical analyses were performed using the base R statistical package (R Foundation for Statistical Computing, Vienna, Austria).

Results

Demographics

In this study, 26 patients were screened for enrollment. There was 1 screen failure in a patient with moderately controlled seizures; therefore, 25 eligible patients aged 6-18 (median age, 13) years participated in study (Figure 3). In total, 16 patients were randomized to VR and 9 to SOC group; 1 patient in the VR arm was excluded from the analysis because of inability to wear the HMD over glasses. Of the remaining 24 patients, 83% (20/24) participants were males and 17% (4/24) were females; 54% (13/24) had HA and 46% (11/24) had HB, 42% (10/24) had severe hemophilia, 58% (14/24) had nonsevere hemophilia, and 54% (13/24) were on routine prophylaxis (Table 1). No patient harms or unintended effects were seen in either the SOC or VR group.

Outcomes

Safety

No adverse events occurred during the trial. No patient experienced VR sickness, seizures, discomfort from the HMD, or infection-related events related to the VR experience.

Procedure Time

The median procedure time was 10 (range 1-31) minutes in the VR group, and was similar to 9 (range 3-20) minutes in the SOC group (P=.76; Table 1). The mean procedure time was 12.3 minutes (95% CI 7.2-17.4) in the VR group and 10.1 minutes (95% CI 5.7-14.6) in the SOC group.

Virtual Reality Equipment Usability

In this study, 60% (9/24) patients wore VR equipment during their entire procedure, 27% (4/24) utilized VR for part of their procedure, and 13% (2/24) only used VR prior to the procedure. Of the 4 participants who used VR for a part of their procedure, 1 removed the headset after the IV stick, while 3 used VR at the beginning of their procedures, removed the headset to watch the IV stick, and then resumed VR. The 2 participants, who only used VR prior to their procedures, chose to watch the entirety of the IV procedure. No technical issues were noted with the orchestration dashboard, headsets, or game hardware.
Figure 3. Trial flow diagram. VR: virtual reality.

Table 1. Participants’ characteristics and median procedure times.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Virtual reality (n=15)</th>
<th>Control (n=9)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males, n (%)</td>
<td>12 (80)</td>
<td>8 (89)</td>
<td>.99</td>
</tr>
<tr>
<td>Median age, years</td>
<td>12.2</td>
<td>12.8</td>
<td>.56</td>
</tr>
<tr>
<td><strong>Hemophilia type, n (%)</strong></td>
<td></td>
<td></td>
<td>.99</td>
</tr>
<tr>
<td>Hemophilia A</td>
<td>8 (53)</td>
<td>5 (56)</td>
<td></td>
</tr>
<tr>
<td>Hemophilia B</td>
<td>7 (47)</td>
<td>4 (44)</td>
<td></td>
</tr>
<tr>
<td><strong>Hemophilia severity, n (%)</strong></td>
<td></td>
<td></td>
<td>.36</td>
</tr>
<tr>
<td>Mild</td>
<td>3 (20)</td>
<td>3 (33)</td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>4 (27)</td>
<td>4 (44)</td>
<td></td>
</tr>
<tr>
<td>Severe</td>
<td>8 (53)</td>
<td>2 (22)</td>
<td></td>
</tr>
<tr>
<td>Routine prophylaxis, n (%)</td>
<td>9 (60)</td>
<td>4 (44)</td>
<td>.68</td>
</tr>
<tr>
<td>Median procedure time, minutes</td>
<td>10</td>
<td>9</td>
<td>.76</td>
</tr>
<tr>
<td>Median orchestration events, n</td>
<td>17</td>
<td>N/A*</td>
<td>N/A</td>
</tr>
</tbody>
</table>

aN/A: not applicable.
Table 2. Median pre- and postprocedural visual analog/FACES scale score.

<table>
<thead>
<tr>
<th>Question</th>
<th>Virtual reality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nervousness prior to intravenous procedure (0=None, 100=Very)</strong></td>
<td></td>
</tr>
<tr>
<td>Nurse response</td>
<td>14</td>
</tr>
<tr>
<td>Caregiver response</td>
<td>25</td>
</tr>
<tr>
<td>Patient response</td>
<td>3</td>
</tr>
<tr>
<td><strong>Assessment of the distraction technique</strong></td>
<td></td>
</tr>
<tr>
<td>Engaged or attentive (0=Yes it did, 100=Did not)</td>
<td></td>
</tr>
<tr>
<td>Nurse response</td>
<td>13</td>
</tr>
<tr>
<td>Caregiver response</td>
<td>13</td>
</tr>
<tr>
<td>Patient response</td>
<td>18</td>
</tr>
<tr>
<td><strong>Impact on anxiety (0=decreased, 100=increased)</strong></td>
<td></td>
</tr>
<tr>
<td>Nurse response</td>
<td>3</td>
</tr>
<tr>
<td>Caregiver response</td>
<td>15</td>
</tr>
<tr>
<td>Patient response</td>
<td>8</td>
</tr>
<tr>
<td><strong>Impact on pain (0=lessened, 100=worsened)</strong></td>
<td></td>
</tr>
<tr>
<td>Nurse response</td>
<td>4</td>
</tr>
<tr>
<td>Caregiver response</td>
<td>4</td>
</tr>
<tr>
<td>Patient response</td>
<td>3</td>
</tr>
<tr>
<td><strong>Likability (0=really like, 100=did not like)</strong></td>
<td></td>
</tr>
<tr>
<td>Nurse response</td>
<td>1</td>
</tr>
<tr>
<td>Caregiver response</td>
<td>4</td>
</tr>
<tr>
<td>Patient response</td>
<td>3</td>
</tr>
<tr>
<td><strong>Questions for the virtual reality group only</strong></td>
<td></td>
</tr>
<tr>
<td>Ease of using the virtual reality equipment (0=Easy, 100=Hard)</td>
<td></td>
</tr>
<tr>
<td>Nurse response</td>
<td>2</td>
</tr>
<tr>
<td>Caregiver response</td>
<td>9</td>
</tr>
<tr>
<td>Patient response</td>
<td>7</td>
</tr>
<tr>
<td>Virtual reality for future intravenous procedures (0=likely, 100=not likely)</td>
<td></td>
</tr>
<tr>
<td>Nurse response</td>
<td>3</td>
</tr>
<tr>
<td>Caregiver response</td>
<td>8</td>
</tr>
<tr>
<td>Patient response</td>
<td>12</td>
</tr>
</tbody>
</table>

The VR equipment usability was favorably scored (0=easy to use, 100=hard to use) with median scores of 7, 9, and 2 by participants, guardians, and the nurse, respectively (Table 2). Additionally, patients, caregivers, and nurses positively rated the desire to use VR for future procedures on a scale of 0 (would really like to use VR again) to 100 (would really not like to use VR again); in majority (36/45, 80%) of evaluations, children, caregivers, and nurses reported that they would like to use VR for future procedures (score <50).

**Orchestration Events**

The number of deliberate orchestration events per patient, including commands to enter or exit a mini-game, transport the player patient to a virtual location, connect to a paired iOS device, adjust the volume, open the video feed from the iOS device’s camera, or calibrate the headset units for fit and viewing preference, was available for 13 of the 15 patients. A server storage issue led to data loss, resulting in missing orchestration data for 2 patients. In the remaining 13 patients, there were a median of 17 orchestration events (range 7-28) per procedure.

**Nervousness or Worry and Pain**

The groups did not differ statistically in preprocedural nervousness or worry, as rated by the participant ($P=0.67$), caregiver ($P=0.37$), or nurse ($P=0.27$; Table 2). Median nervousness rating for the VR group was 3 (range 0-94) and 12 for the SOC group (range 0-100). Preprocedural patient nervousness did differ between age groups ($P=0.002$) with those aged 6-12 years (median 50, range 0-100) being significantly more nervous compared with those aged 13-18 years (median...
0, range 0-13). Preprocedural patient nervousness did not statistically differ between those on and not on prophylaxis ($P=0.64$). Median nervousness rating was 6 (range 0–94) for those on prophylaxis and 27 for those not on prophylaxis (range 0–100). Both groups favorably viewed distraction techniques in terms of the effect on procedural anxiety, with median responses being 8 for the VR group and 10 for the SOC group (0=the distraction technique decreased anxiety and 100=the technique increased anxiety). Both VR and SOC distraction techniques had a positive influence on procedural pain (0=made pain better, 100=made pain worse); no statistically significant differences were observed between raters in the VR or SOC groups (Table 2).

**Level of Engagement**

Scores of participants, nurses, and caregivers did not significantly differ between VR and SOC distraction techniques in terms of the ability to engage and hold attention (0=held attention and was engaging and 100=did not hold attention and was not engaging; Table 2).

**Discussion**

**Principal Findings**

This study represented the first known effort to translate VR from a research environment to a functioning pediatric clinic environment as a comprehensive platform including custom content (games), tools, and hardware (HMDs). The final platform included a high-tech, VR-based, child engaging distraction with child choice incorporated into a platform that allowed for need-based adult involvement. The final HMDs were multisized, lightweight, and disposable and had an estimated unit cost of less than US $55. If manufactured at scale, the estimated cost of the headphones, headsets, lenses, cardboard, elastic, and hooks and loops is less than US $20. HMDs accommodated 2 distinct size faces—one for average face size of patients over 15 years of age and another for that of patients under 15 years of age. The game navigation was hands-free and wireless. Games produced were designed to appeal to a wide age range of male and female patients. Games were also engineered to achieve a consistent 30 frames per second playback rate on limited resource devices. Additionally, games afforded the player control over a child engaging gameplay environment without requiring the use of the patients’ hands—a unique affordance ingrained in this platform. The system achieved hands-free gameplay via the introduction of proprietary hardware in the custom HMD, which is capable of monitoring the breathing of the play and providing this feedback to the games. The system did not require a high-end computer or technical expertise for installation, and the orchestration features allowed for customization to patient needs and family engagement. Successful incorporation into the clinic was demonstrated by the lack of safety and technical issues. Although our sample size did not allow equivalence comparison, we had similar lengths of procedure times in both groups.

The games themselves mitigated the risk of the player developing simulator sickness with the standard technique of introducing a static frame of reference during periods of motion. While static frames of reference are a conventional technique used to prevent VR sickness, the technique’s manifestation in our system was novel in approach in that the technique utilized world elements as the static frame of reference. This execution of the technique, therefore, did not sacrifice the immersive quality of the environment. In addition, the static frames introduced a form of kinesthetic reinforcement for the players’ physical sensation to the virtual world, thus lending to their immersive quality. For example, in the snorkeling game in the game world, players observed a projection of a virtual diving mask in their game view. The projection not only acted as a static frame of reference that thematically fits within the game environment but also created a relationship between the physical sensation of the physical HMD and the virtual mask as seen in the game world.

Our environments incorporated a core component of distraction theory most significantly, with affordance of control and choice as a means to distract [28]. This method is inherent in child engaging games, with our platform bringing this method to the patient population through hands-free game controls and thus representing the first set of pediatric-focused games, which utilize this technique while not requiring hand movement to interact and thus bringing this technique via games to patients receiving IV procedures. In total, the techniques employed throughout our ecosystem resulted in games that apply distraction theory through child engaging VR environments while not triggering simulator sickness in participants, as demonstrated in the collected data.

Pain and anxiety related to procedures is a concern for patients, families, and providers, particularly in pediatric settings. Because children with hemophilia have frequent needle procedures, they represent a population that could really benefit from VR-based distraction. While there is strong evidence for the success of distraction during pediatric procedures, it remains unclear which of the 4 main elements of effectiveness (high vs no or low tech, child engaging vs passive, degree of adult involvement, and availability of child choice) contribute to the success of the technique [9,10]. Encouragingly, 87% of participants wore the VR equipment during some or all of their procedure, and overall scores regarding the impact of VR on pain, anxiety, usability, likeability, and level of engagement were favorable.

**Limitations**

The trial was limited by small sample size and single-institution design. Additionally, we studied only a single intervention, so we were unable to test if the VR platform would continue to perform well with future use in the same patient. The study was underpowered to evaluate the equivalence of procedure time. The study was also underpowered to compare VR versus SOC attributes, but the majority of caregivers and providers had favorable ratings of VR, and 80% of participants, all of whom had previous experience with SOC distraction in our clinic, reported a desire to use VR for future procedures. The outcomes related to pain and anxiety were self-reported. Addition of objectives measures of distress, such as blood pressure and heart rate, would strengthen future studies.
Conclusions

We demonstrated that a custom-designed VR platform could be used safely during IV procedures in a pediatric hemophilia population with specific design needs. These results warrant future exploration to assess the impact of our platform on its ability to reduce the burden of IV procedures on patients, parents, and clinical care providers. Future studies will be needed to validate our findings in other disease populations, clinical settings, and institutions with larger participant numbers. While patients with hemophilia may have more IV experiences than most children, it is likely that the positive effects of high-quality distraction would be generalizable to IV procedures in any pediatric population. We plan larger studies comparing (1) our child engaging platform versus a passive distracter; (2) the degree to which adult involvement impacts successfulness; (3) the importance of the availability of child choice in distracters; (4) ability to decrease procedural chemical sedation; and (5) cost. Additionally, nonrandomized trials allowing children to choose their distraction technique of choice during procedures would generate useful information. The inclusion of objective measures of pain and anxiety would strengthen future studies. This study suggests that a custom VR ecosystem with clinician orchestration is a promising modality to provide distraction during IV procedures in pediatrics with the potential to mitigate the perception of procedural pain and anxiety.

Acknowledgments

The authors would like to thank the patients and staff of the Nationwide Children’s Hospital HTC in Columbus, Ohio, for their support and participation. Additionally, we thank Ms Beth Burkhart, MBA, for assistance with the development and maintenance of the REDcap database, Ms Jessie Haines for administrative support, Mr Roy Goudy for video assistance, and Dr Michael Dunn for critical review of the manuscript. This study was funded by Ohio State University Center for Clinical and Translational Science/Nationwide Children’s Hospital Research Information Solutions & Innovations Voucher #3796: UL1TR001070 and The National Hemophilia Foundation Nursing Excellence Award.

Conflicts of Interest

AD reports grants, personal fees, and nonfinancial support from Bayer, CSL Behring, Pfizer, Shire, and Hema Biologics; grants and personal fees from Bioverativ; personal fees and nonfinancial support from Medscape and NovoNordisk; and grants from Alnylam and Octapharma during the conduct of the study, and other from Little Seed Inc outside the submitted work. JP reports grants from The Ohio State University, outside the submitted work. JL reports grants from The Ohio State University, during the conduct of the study. CFB reports personal fees from Shire. JRS, AG, and RS have no conflict of interest to declare.

Authors’ Contributions

AD conceptualized and designed the study, designed the data collection instruments, supervised data collection, drafted the initial manuscript, and reviewed and revised the manuscript. JP conceptualized and designed the study, and reviewed and revised the manuscript. CFB conceptualized and designed the study, designed the data collection instruments, collected data, and reviewed and revised the manuscript. JL, AG, and RS designed the study, collected data, reviewed, and revised the manuscript. JRS designed the study, reviewed, and revised the manuscript.

Multimedia Appendix 1

Voxel Bay video experience.

[MP4 File (MP4 Video), 23MB - games_v7i1e10902_app1.mp4 ]

Multimedia Appendix 2

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 82KB - games_v7i1e10902_app2.pdf ]

References


Abbreviations

HA: hemophilia A
HB: hemophilia B
HMD: head-mounted display
HTC: Hemophilia Treatment Center
IV: intravenous
SOC: standard of care
VAS: visual analog scale
VR: virtual reality

©Amy Dunn, Jeremy Patterson, Charmaine F Biega, Alice Grishchenko, John Luna, Joseph R Stanek, Robert Strouse. Originally published in JMIR Serious Games (http://games.jmir.org), 09.01.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on http://games.jmir.org, as well as this copyright and license information must be included.
Abstract

Background: e-Bug, led by Public Health England, educates young people about important topics: microbes, infection prevention, and antibiotics. Body Busters and Stop the Spread are 2 new e-Bug educational games.

Objective: This study aimed to determine students’ baseline knowledge, views on the games, and knowledge improvement.

Methods: Students in 5 UK educational provisions were observed playing 2 e-Bug games. Before and after knowledge and evaluation questionnaires were completed, and student focus groups were conducted.

Results: A total of 123 junior and 350 senior students completed the questionnaires. Vaccination baseline knowledge was high. Knowledge increased significantly about antibiotic use, appropriate sneezing behaviors, and vaccinations. In total, 26 student focus groups were conducted. Body Busters was engaging and enjoyable, whereas Stop the Spread was fast-paced and challenging but increased vaccination and health behavior intentions.

Conclusions: e-Bug games are an effective learning tool for students to enhance knowledge about microbes, infection prevention, and antibiotics. Game-suggested improvements should help increase enjoyment.

Introduction

Educating children and young people is important in the fight against antibiotic resistance. Through education, we can raise awareness, enhance knowledge, and modify behavioral intentions about hygiene and antibiotic use in our future generation of antibiotic users. E-Bug, led by Public Health England (PHE), is an international health education resource that teaches children and young people about hygiene, the spread of infections, antibiotic use, and resistance. e-Bug includes Web-based lesson plans and activities for educators and educational games for students hosted on the e-Bug website. Evaluation of e-Bug activities to be undertaken in schools and science shows has been well-documented [1-6], and the National Institute of Clinical Excellence [7] has suggested that schools may use the evidence-based e-Bug resources to educate children and young people in an age-appropriate way about hygiene, prevention of infections, and antibiotic use.

Google Analytics has been used to monitor Web traffic to the e-Bug website since 2010 [2]. The e-Bug website had 94,675 visitors from September 1, 2016, to August 31, 2017, and 100,955 visitors in the previous academic year. The e-Bug...
games’ home page was the second highest visited page with 28,610 page views from September 1, 2016, to August 31, 2017; the senior students games’ home page had 10,154 views during the same period.

The internet is a suitable tool for health promotion, and internet-based health interventions have been shown to change health behaviors. Gamification, where features of gaming are used in other disciplines, has become increasingly popular in recent years, aiming to make science and health education more available and exciting to the general public. Serious games are those games where the primary focus is not entertainment but education and learning [8]. A meta-analysis of serious games in regard to their effect on cognitive processes and motivation found the games to positively affect cognitive processes, including learning and retention compared with traditional educational methods, with no difference to motivation [9].

Evidence suggests that gamification and serious games for health and well-being are most effective when targeting health-related behaviors [10]. For instance, positive associations between gamification, serious games, and school-aged knowledge and behavior have been reported in public health topics such as asthma [11], fruit and vegetable consumption [12], and oral hygiene [13,14].

The e-Bug Web-based educational games [15] have been previously evaluated including a study that evaluated 3 e-Bug games using a mixed-methods approach; the 3 e-Bug games showed an improvement in knowledge and focused on the use of antibiotics for bacterial versus viral infections and ensured that the course of antibiotics is completed [16].

The aim of this study was to evaluate the 2 e-Bug educational games: Body Busters, previously evaluated by Hale et al and then modified with new content [16], and Stop the Spread, a new educational game launched in 2016. The questions of the research study are as follows:

- What is students’ baseline knowledge about the game-learning outcomes?
- What is students’ change in knowledge following the games?
- What are students’ views on the games to suggest improvements?

Figure 1 details the style of play and learning outcomes of the 2 games. Both e-Bug games are responsive on all devices including computers and tablets. Pilot game testing was conducted at 3 schools before the game launch to ensure the games worked correctly and the instructions were clear.
**Figure 1.** The style of play and learning outcomes of Body Busters and Stop the Spread.

<table>
<thead>
<tr>
<th>Game</th>
<th>Style of game</th>
<th>Learning outcomes</th>
</tr>
</thead>
</table>
| **Body Busters**| Pac-man style game in which users need to collect all the antibiotics to kill the bacteria. Useful bacteria give the users extra lives, but these bacteria are also killed by the antibiotics. Harder levels now also include viruses, which are not killed by the antibiotics. | • There are different types of microbes (bacteria and viruses)  
• Bacteria can be harmful or useful  
• Antibiotics kill useful and harmful bacteria but do not kill viruses |
| **Stop the Spread** | Users prevent the spread of infection in a school by catching sneezes (red circles) with a tissue and throwing the used tissue into a bin. As sneezes go uncaught, more children in the playground become infected (green circles). The game is over once every child in the playground is infected. The longer the game lasts, the more points the user scores. An additional level includes the concept of vaccination and herd immunity. In this level users vaccinate children in the playground as quickly as they can to slow the spread of infection. | • Infections spread by sneezing  
• Catch sneezes in a tissue  
• Put used tissues in the bin  
• Using tissues stops the spread of infections  
• Vaccinations can help prevent the spread of infection |
Methods

Research Design
The study was a mixed-method evaluation using quantitative and qualitative methods. Quantitative methods included before and after students’ knowledge questionnaires; qualitative methods included students’ focus groups and open-ended questions and responses from the students’ postgaming evaluation questionnaire.

Sampling and Recruitment
Educational providers, including schools and summer schools, were invited to take part in the study through convenience recruitment of educators at educational and scientific conferences and then through snowball sampling. Sampling aimed to ensure a representation of school-aged children across 3 local authorities in the United Kingdom, including rural and urban schools, different socioeconomics, and selective grammar and nonselective state schools (see Table 1). Local authorities were Gloucestershire, Buckinghamshire, and South Wales.

Ethics
All researchers who observed the sessions had a Disclosure Barring Check, through PHE, to work with children. This study did not require National Research Ethics Service approval as it was outside the National Health Service and was classed as a service evaluation. PHE provided written confirmation approving the service evaluation in July 2016. Educational providers gave informed written consent before the study took place; students were involved and their parents were given the option for students to opt out at any point during data collection. Teachers reported that no students opted out of the research. Consent was deemed accepted if the participants completed the before and after knowledge questionnaires. Questionnaires were collected in line with the Data Protection Act 1998 and Caldicott 1999 regulations on handling and distributing sensitive participant information. Focus group participants provided verbal informed consent for participation in the study, audio recording, and the publishing of anonymized quotes.

Data Collection and Analysis
Data collection took place between August 2016 and July 2017.

Quantitative Data
Before and after knowledge questionnaires were used to evaluate whether playing the e-Bug games had made any change to students’ knowledge. Questionnaires were based on previously validated questionnaires used to evaluate the e-Bug games and activities [2,3,16]; 4 additional questions (1, 2, 3, and 5) were included in the questionnaire to cover additional learning outcomes that could be indirectly improved through game play and are hereby referred to as general questions (see Multimedia Appendix 1 for questionnaires).

Data collection consisted of (1) students completed pregame-play questionnaire 1 alone without consultation, (2) students played on Body Busters for 5 min, (3) students played on Stop the Spread for 5 min, and (4) students completed postgame-play questionnaire 2 alone without consultation, which also included some additional open-ended evaluation questions. The game play was for 5 min to follow the methodology of previous evaluations [16], and as 5 min is the estimated amount of time it takes to play 1 round of the games from testing, it allowed the study to measure knowledge change after single game play rather than repeated game play.

A researcher was present in each session to monitor and observe game play and hand out and collect questionnaires from the students. Data collection occurred in a convenient room where students had their own computer. The rationale for that was to model how the games might be played in a real-life teaching situation. Figure 2 provides further details on the data collection process.

McNemar test was used for each response from the multiple-choice question to determine the significance of the difference in the proportion of correct answers before and after game playing. Moreover, 95% CIs of the odds ratio were estimated to determine the odds of students answering correctly. Analysis was performed separately for junior and senior school-aged pupils, as knowledge change could differ between age groups. All statistical analysis was completed in STATA, version 14.2.

The postgame-play questionnaire 2 included an additional 7 questions on game enjoyment, including 2 Likert scale questions (students circled a number scale of 1-10) and 5 open-ended questions. Likert scale responses were inputted into MS Excel, and mean enjoyment scores for each game were calculated for junior, senior, and all students.

Table 1. Demographics of educational providers.

<table>
<thead>
<tr>
<th>Educational provider</th>
<th>Local authority</th>
<th>Type of education</th>
<th>Questionnaires (N=473), n (%)</th>
<th>Focus groups (N=26), n (%)</th>
<th>Students (N=126), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Gloucestershire</td>
<td>Summer school</td>
<td>61 (13)</td>
<td>14 (54)</td>
<td>61 (48)</td>
</tr>
<tr>
<td>B</td>
<td>Gloucestershire</td>
<td>Grammar</td>
<td>29 (6)</td>
<td>1 (4)</td>
<td>6 (5)</td>
</tr>
<tr>
<td>C</td>
<td>South Wales</td>
<td>State</td>
<td>100 (21)</td>
<td>4 (15)</td>
<td>24 (19)</td>
</tr>
<tr>
<td>D</td>
<td>Bedfordshire</td>
<td>State</td>
<td>183 (39)</td>
<td>4 (15)</td>
<td>20 (16)</td>
</tr>
<tr>
<td>E</td>
<td>Bedfordshire</td>
<td>Grammar</td>
<td>100 (21)</td>
<td>3 (12)</td>
<td>15 (12)</td>
</tr>
</tbody>
</table>
**Qualitative Data**

In total, 5 open-ended questions on enjoyment were included in the postgame-play questionnaire 2 to provide deeper qualitative data from all students.

Semistructured focus groups were facilitated immediately after the game intervention by VLY, CVH, and CVE who are all trained qualitative researchers for the e-Bug project, PHE. All 26 focus groups took place in person in a convenient room at the educational establishment. Focus groups of size 4 to 6 students, chosen by the class teacher and represented a mix of student abilities, lasted for 6 to 20 min depending on participant age. Focus groups were audio recorded, transcribed verbatim, and checked for accuracy by CVH or CVE. Furthermore, 26 focus groups were planned with all 5 schools participating; no new themes emerged from the later focus groups, and researchers agreed that data saturation had been reached.

The topic guide for the focus groups was based on previous e-Bug evaluation topic guides for e-Bug and included additional questions on Stop the Spread learning outcomes. The schedule was piloted during the e-Bug game development in 3 testing sessions in schools in March 2016.

All focus group data and open-ended responses on enjoyment were inputted into NVivo 10 (QSR International) qualitative analysis software. NVivo 10 was used to organize, code, and analyze the focus group transcripts by CVH and open-ended evaluation responses by CVE. A subset of focus group data (2 junior and 2 senior transcripts) was analyzed by a second researcher (CVE) to ensure reliability. Both researchers discussed the data and coding to agree on the emerging themes before developing a thematic framework. Any discrepancies between researchers were resolved through discussion until an agreement was reached. The thematic framework was discussed by the research team.
Results

Main Findings

The study recruited 473 students (123 junior and 350 senior students) aged 7 to 16 years from 5 educational providers across 3 local authorities in the United Kingdom (illustrated in Table 1).

Before and after knowledge questionnaires were completed by 473 students. Baseline knowledge about vaccinations was high in junior students (>60% correct responses) and was higher in senior students (>80% correct responses), except for 1 question for which the baseline knowledge 55%. Baseline knowledge about antibiotics was low in junior and senior students (<40% and <67% correct responses, respectively). Senior students had greater pregaming knowledge than junior students in 11 out of the 12 questions and had greater postgaming knowledge in 10 out of the 12 questions. However, knowledge change was greater in 9 out of the 12 questions for junior students.

Quantitative results showed significant improvements in knowledge (P<.05) about antibiotic use, appropriate sneezing behaviors, and vaccinations for both age groups: junior (7-11 years) and senior (11-16 years).

In total, 26 student focus groups with 126 students were conducted, 10 junior (7-11 years) and 16 senior (11-16 years) with approximately 4 to 6 students per focus group. Researchers observed that students enjoyed playing Body Busters more, and they were keen to answer questions to gain more lives, whereas Stop the Spread was more difficult and on occasion required some researcher explanation. The mean enjoyment score for Body Busters was 8.4/10 for juniors and 7.2/10 for seniors; the mean enjoyment score for Stop the Spread was 6.2/10 for juniors and 5.1/10 for seniors (illustrated in Table 2).

Qualitative Data

Reported Views of Body Busters Game

Body Busters Positive Perceptions: User Experience

Qualitative results from focus groups and open questionnaire responses for Body Busters were overall very positive with a few suggestions for improvement. Many students of both age groups reported positive perceptions of user experience; at least 1 participant in each focus group reported positive levels of enjoyment and nearly all other participants agreed. Many students wanted to play for longer as the game was very engaging, similar to Pac-Man, at the correct level of difficulty, and students reported that they had learned through the gaming experience:

I could be on there [Body Busters] for like the whole day or an hour, or actually 2 hours. [Junior student, Focus group 9]

I feel like it was teaching us that antibiotics are to be used to kill a bacterial infection, but also that not all bacteria is harmful to the body. [Senior student, Focus group 5]

I enjoyed the game play - collecting antibiotics and dodging the bad bacteria. The actual game was fun. [Senior student questionnaire response]

I realised not all microbes are harmful. [Senior student questionnaire response]

Body Busters Positive Perceptions: Game Functions

Many students of both age groups reported positive perceptions of the game functions, including game recovery aids, the pace of the game, and the game aesthetics. Some junior students also reported that they liked the microbe characters and the concept of being able to gain more lives. Some senior students also reported that they liked the different levels of the game and the useful pictorial instructions:

I liked getting more lives from the good bacteria and I liked collecting the antibiotics. [Junior student, Focus group 17]

It was quite like, there was an equation, so the robot plus the circle equals health up, and that was quite a good way of formatting it without writing it out as paragraphs. [Senior student, Focus group 15]

The game had good bacteria so you could regain your lives. [Senior student questionnaire response]

Body Busters Negative Perceptions

A common negative theme for Body Busters was the slower pace of the game when users lost a life, reported by many junior and senior students:

I was cornered by two enemies and lost lots of lives-then I was really slow for the rest of the game. Becoming slow made it much less enjoyable. [Senior student questionnaire response]

Some students suggested ways to modify this aspect of the game, including when the user loses a life, do not slow the avatar down but instead increase the speed or size of the harmful bacteria and viruses. Other students reported that the instructions could be improved by keeping the pictorial instructions but adding written instructions for clarity, and a few students suggested having a visible key for the different microbes in the game play:

Maybe...if you lose your health maybe there’s like a circle around the bacteria so that their like range gets bigger, but you’re still the same speed, or they get bigger themselves and you just increase the size of the map so two of them can go. [Senior student, Focus group 15]

I think you should make it not go so slow and I think when you press start it should like count 1, 2, 3 go, so then you actually know where you are and you know where everything else is, you can just go off. [Junior students, Focus group 17]
Table 2. Enjoyment scores for the e-Bug games (score out of 10).

<table>
<thead>
<tr>
<th>Students</th>
<th>Body Busters</th>
<th>Stop the Spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junior</td>
<td>8.4</td>
<td>6.2</td>
</tr>
<tr>
<td>Senior</td>
<td>7.2</td>
<td>5.1</td>
</tr>
<tr>
<td>All students</td>
<td>7.6</td>
<td>5.5</td>
</tr>
</tbody>
</table>

**Reported Views of Stop the Spread Game**

**Stop the Spread Positive Perceptions: User Experience**

Some students in both age groups reported positive perceptions of user experience, including reporting an increase in knowledge about the spread of infection and the importance of vaccinations:

- Vaccinations protect other people, so if you have a vaccination then you won’t get any colds and stuff, and then you’re protecting other people because you can’t pass anything on to them. [Junior student, Focus group 16]

- I think it teaches you that if you know you have a cold before you go to school...you have a tissue and you make sure you sneeze into the tissue and put it in the bin, so it’s safe, or if it’s too bad...you might want to stay away as you’ll infect loads of people and make it even worse. [Senior student, Focus group 21]

- The learning points - very obvious how severely something can spread. [Senior student questionnaire response]

In most junior focus groups, at least 1 student voiced an increase in understanding appropriate health behaviors, especially about sneezing, and most participants concurred with them. In some focus groups, students reported intent to change health behaviors:

- It was teaching us that you should always use a tissue and put that in the bin. [Junior student, Focus group 16]

- It shows you what to do not, like if you feel like you’re going to sneeze and cough, do it into a tissue. [Junior student, Focus group 17]

- That we need to sneeze into a tissue not just sneeze out...Put it in the bin, don’t keep it in your coat pocket like I used to do. [Junior student, Focus group 19]

**Stop the Spread Positive Perceptions: Game Functions**

Some students, junior and senior, enjoyed the fast pace of the game and different levels of difficulty, especially the vaccination levels. A few senior students reported on the game aesthetics:

- I enjoyed the colours, so when it showed you they were red and then purple, that was helpful. [Senior student, Focus group 22]

- You could protect people by injecting them. [Senior student questionnaire response]

**Stop the Spread Negative Perceptions**

A common negative theme of Stop the Spread was that the game was too difficult. In most focus groups, the majority of students felt the game was “too hard” because it was “too fast” and “too many children were sneezing at the same time.” Many students reported a lack of engagement to continue to play the game as they had negative emotions, such as feeling “stressed” and “annoyed”:

- Oh gosh level 1 was fine, but as the levels went on I was like oh gosh how are we meant to do this now. [Junior student, Focus group 12]

- Maybe it was a little bit too hard. [Junior student, Focus group 18]

- It’s really fast, and it’s really hard. [Senior student, Focus group 2]

In most focus groups, suggestions for improvements were provided, including slow the pace of the game down to make it easier, have fewer children sneezing at the same time, or slow the time down between students sneezing, include more levels of different difficulty such as an easy level or have a tutorial level, and make the instructions clearer and simpler:

- The thing is that they all sneezed at the same time and I didn’t have enough time to put the tissue in the bin, so that is really a struggle. [Junior student, Focus group 16]

- Yeah too fast paced, maybe slow it down to start off with one person infected, and then do levels as you did with the Pac man one. [Senior student, Focus group 22]

- The instructions could have been clearer. [Senior student questionnaire response]

Suggestions for game improvements and modifications from the qualitative focus groups and evaluation questions are summarized in Textboxes 1 and 2.

**Quantitative Data**

Tables 3 and 4 show the percentage correct before the game intervention, percentage correct after the game intervention, and the P value for junior (7-11 years) and senior (11-15 years) school-aged students, respectively. Table 5 shows a comparison of baseline and postgaming knowledge between age groups.
Textbox 1. Suggestions for Body Busters game improvements from qualitative focus groups and evaluation questions.

Body Busters

1. When the user loses a life, do not slow the avatar down but either:
   • increase the speed of harmful bacteria and viruses or
   • increase the size of harmful bacteria and viruses
2. Make instructions clearer
   • keep the pictorial instructions
   • add written instructions
3. Have a visible key for the different microbes in the game play
4. Add more levels in different areas of the body

Textbox 2. Suggestions for Stop the Spread game improvements from qualitative focus groups and evaluation questions.

Stop the Spread

1. Slow the pace of the game down to make it easier
   • have fewer children sneezing at the same time
   • slow the time down between students sneezing
2. Include more levels of different difficulty
   • have an option for an easy level
   • have a tutorial or practice level
3. Make the instructions simpler and clearer

Table 3. Improvement scores by question for junior schools.

<table>
<thead>
<tr>
<th>Question or statement</th>
<th>Correct before, %</th>
<th>Correct after, %</th>
<th>Odds ratio (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Which of these microbes causes coughs and colds?</td>
<td>35</td>
<td>29</td>
<td>0.57 (0.21-1.46)</td>
<td>.29</td>
</tr>
<tr>
<td>What is the best way to treat an infection with a virus?</td>
<td>44</td>
<td>47</td>
<td>1.25 (0.55-2.92)</td>
<td>.70</td>
</tr>
<tr>
<td>Antibiotics help cure colds</td>
<td>33</td>
<td>28</td>
<td>0.60 (0.23-1.46)</td>
<td>.31</td>
</tr>
<tr>
<td>Which of these infections could antibiotics be used to treat?</td>
<td>31</td>
<td>46</td>
<td>2.88 (1.24-7.43)</td>
<td>.01a</td>
</tr>
<tr>
<td>Most coughs and colds get better without antibiotics</td>
<td>59</td>
<td>68</td>
<td>2.22 (0.97-5.54)</td>
<td>.06b</td>
</tr>
<tr>
<td>All microbes are bad or harmful</td>
<td>78</td>
<td>84</td>
<td>1.50 (0.68-3.41)</td>
<td>.36</td>
</tr>
<tr>
<td>You cannot infect other people around you through coughs and sneezes</td>
<td>79</td>
<td>78</td>
<td>0.93 (0.42-2.07)</td>
<td>&gt;.99</td>
</tr>
<tr>
<td>The more people are vaccinated, the more people are protected from that infection</td>
<td>60</td>
<td>72</td>
<td>3.00 (1.23-8.35)</td>
<td>.01a</td>
</tr>
<tr>
<td>By getting vaccinated, you can also protect others around you from infection</td>
<td>59</td>
<td>68</td>
<td>2.22 (0.97-5.54)</td>
<td>.06b</td>
</tr>
<tr>
<td>Antibiotics (list); kill good and bad bacteria</td>
<td>22</td>
<td>33</td>
<td>2.30 (1.05-5.41)</td>
<td>.04a</td>
</tr>
<tr>
<td>Vaccinations (list); protect us from catching and spreading diseases</td>
<td>79</td>
<td>78</td>
<td>0.75 (0.21-2.46)</td>
<td>.79</td>
</tr>
<tr>
<td>The best way to stop microbes in coughs and sneezes spreading is to (list)</td>
<td>61</td>
<td>71</td>
<td>4.00 (1.29-16.4)</td>
<td>.01a</td>
</tr>
</tbody>
</table>

aSignificant at .05.
bApproaching significance at .06.
Table 4. Improvement scores by question for senior schools.

<table>
<thead>
<tr>
<th>Question or statement</th>
<th>Correct before, %</th>
<th>Correct after, %</th>
<th>Odds ratio (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Which of these microbes causes coughs and colds?</td>
<td>47</td>
<td>46</td>
<td>0.88 (0.52-1.49)</td>
<td>.71</td>
</tr>
<tr>
<td>What is the best way to treat an infection with a virus?</td>
<td>46</td>
<td>46</td>
<td>1.06 (0.65-1.71)</td>
<td>.91</td>
</tr>
<tr>
<td>Antibiotics help cure colds</td>
<td>61</td>
<td>49</td>
<td>0.24 (0.11-0.45)</td>
<td>&lt;.001a</td>
</tr>
<tr>
<td>Which of these infections could antibiotics be used to treat?</td>
<td>56</td>
<td>63</td>
<td>1.80 (1.08-3.06)</td>
<td>.02a</td>
</tr>
<tr>
<td>Most coughs and colds get better without antibiotics</td>
<td>67</td>
<td>73</td>
<td>1.79 (1.10-2.95)</td>
<td>.02a</td>
</tr>
<tr>
<td>All microbes are bad or harmful</td>
<td>95</td>
<td>92</td>
<td>0.40 (0.15-0.95)</td>
<td>.04a</td>
</tr>
<tr>
<td>You cannot infect other people around you through coughs and sneezes</td>
<td>92</td>
<td>90</td>
<td>0.72 (0.37-1.37)</td>
<td>.36</td>
</tr>
<tr>
<td>The more people are vaccinated, the more people are protected from that infection</td>
<td>84</td>
<td>89</td>
<td>1.89 (1.04-3.55)</td>
<td>.04a</td>
</tr>
<tr>
<td>By getting vaccinated, you can also protect others around you from infection</td>
<td>55</td>
<td>65</td>
<td>2.54 (1.56-4.26)</td>
<td>&lt;.001a</td>
</tr>
<tr>
<td>Antibiotics (list); kill good and bad bacteria</td>
<td>36</td>
<td>46</td>
<td>3.12 (1.78-5.74)</td>
<td>&lt;.001a</td>
</tr>
<tr>
<td>Vaccinations (list); protect us from catching and spreading diseases</td>
<td>93</td>
<td>94</td>
<td>1.50 (0.63-3.73)</td>
<td>.42</td>
</tr>
<tr>
<td>The best way to stop microbes in coughs and sneezes spreading is to (list)</td>
<td>79</td>
<td>83</td>
<td>1.94 (1.03-3.79)</td>
<td>.04a</td>
</tr>
</tbody>
</table>

Table 5. Comparison between age groups of baseline and postgaming knowledge.

<table>
<thead>
<tr>
<th>Question or statement</th>
<th>Baseline, %</th>
<th>Postgaming, %</th>
<th>Difference in knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junior correct before</td>
<td>Senior correct before</td>
<td>Difference in knowledge</td>
<td>Junior correct after</td>
</tr>
<tr>
<td>1. Which of these microbes causes coughs and colds?</td>
<td>34</td>
<td>47</td>
<td>13a</td>
</tr>
<tr>
<td>2. What is the best way to treat an infection with a virus?</td>
<td>44</td>
<td>46</td>
<td>2b</td>
</tr>
<tr>
<td>3. Antibiotics help cure colds</td>
<td>33</td>
<td>61</td>
<td>28a</td>
</tr>
<tr>
<td>4. Which of these infections could antibiotics be used to treat?</td>
<td>31</td>
<td>56</td>
<td>25a</td>
</tr>
<tr>
<td>5. Most coughs and colds get better without antibiotics</td>
<td>59</td>
<td>67</td>
<td>8a</td>
</tr>
<tr>
<td>6. All microbes are bad or harmful</td>
<td>78</td>
<td>95</td>
<td>17a</td>
</tr>
<tr>
<td>7. You cannot infect other people around you through coughs and sneezes</td>
<td>79</td>
<td>92</td>
<td>13a</td>
</tr>
<tr>
<td>8. The more people are vaccinated, the more people are protected from that infection</td>
<td>60</td>
<td>84</td>
<td>24a</td>
</tr>
<tr>
<td>9. By getting vaccinated, you can also protect others around you from infection</td>
<td>59</td>
<td>55</td>
<td>4b</td>
</tr>
<tr>
<td>10. Antibiotics (list); kill good and bad bacteria</td>
<td>22</td>
<td>36</td>
<td>14a</td>
</tr>
<tr>
<td>11. Vaccinations (list); protect us from catching and spreading diseases</td>
<td>79</td>
<td>93</td>
<td>13a</td>
</tr>
<tr>
<td>12. The best way to stop microbes in coughs and sneezes spreading is to (list)</td>
<td>61</td>
<td>79</td>
<td>18a</td>
</tr>
</tbody>
</table>

aA higher knowledge in senior students.
bA higher knowledge in junior students.

Junior Student Knowledge About Antibiotics and Vaccinations

Over 70% of junior students had high baseline knowledge of learning outcomes covered in questions 6, 7, and 11, and there was only a small nonsignificant increase in correct answers:

(6) All microbes are bad or harmful (78%-84%; P=.36)
(7) You cannot infect other people around you through coughs and sneezes. True or false (79%-78%; P>.99)
(11) Vaccinations protect us from catching and spreading diseases (79%-78%; P=.79).
Low baseline knowledge (<40% correct answers) was seen in questions 1, 3, 4, and 10:

(1) Which of these microbes causes coughs and colds? Bacteria or fungus or virus (34%-29%; \(P = .29\))

(3) Antibiotics help cure colds? True or false (33%-28%; \(P = .31\))

(4) Which of these infections could antibiotics be used to treat? Bacterial or viral or fungal (31%-46%; \(P = .01\))

(10) Antibiotics: *kill good and bad bacteria* (22%-33%; \(P = .04\)).

Knowledge of the 7- to 11-year-old students significantly improved for 4 of the questions (4, 8, 10, and 12):

(4) Which of these infections could antibiotics be used to treat? Bacterial or viral or fungal (31%-46%; \(P = .01\))

(8) The more people are vaccinated; the more people are protected from that infection. True or false (60%-72%; \(P = .01\))

(10) Antibiotics: *kill good and bad bacteria* (22%-33%; \(P = .04\))

(12) The best way to stop microbes in coughs and sneezes spreading is to: “catch coughs and sneezes in a tissue and throw the tissue away” (61%-71%; \(P = .01\)).

Knowledge improvement for 2 other questions (5 and 9) was approaching significance (\(P = .06\)):

(5) Most coughs and colds get better without antibiotics. True or false (59%-68%; \(P = .06\))

(9) By getting vaccinated, you can also protect others around you from infection. True or false (59%-68%; \(P = .06\)).

Questions that saw the greatest improvement in knowledge for junior students were question 4 (31%-46%; \(P = .01\)), 8 (60%-72%; \(P = .01\)), and 10 (22%-33%; \(P = .04\)). Figure 3 provides before and after knowledge percentages and levels of significance.

There was no evidence of a significant knowledge change for questions 6, 7, and 11 or the 3 general questions (1, 2, and 3), which covered knowledge that could be indirectly gained from playing the 2 games.

**Senior Student Knowledge About Antibiotics and Vaccinations**

Senior school students had greater baseline knowledge than junior students in 10 out of the 12 questions. Senior students had high baseline knowledge (>92% correct scores) to the same 3 questions as junior students (6, 7, and 11). In addition, the 2 questions (8 and 12) on vaccinations and sneezing behaviors had scores less than 70%. Senior students had significantly higher baseline knowledge scores for all bar 2 questions, so it was difficult to improve as much as the junior students. Low baseline knowledge (<40% correct scores) was seen in question 10 (Antibiotics [list] correct answer *Kill good and bad bacteria*).

---

**Figure 3.** Percentage of junior students answering questions correctly before and after playing Body Busters and Stop the Spread Games.
There were significant improvements in knowledge among senior students, with 6 out of the 12 questions showing a significant increase in the odds of students answering correctly (4, 5, 8, 9, 10, and 12), which were the same 6 questions that showed significant improvements among junior students. Questions that saw the greatest improvement in knowledge for senior students were questions 9 (By getting vaccinated, you can also protect others around you from infections; 10.7%) and 10 (Antibiotics kill good and bad bacteria; 10.8%). Figure 4 provides before and after knowledge percentages and levels of significance.

Question 6 (All microbes are bad or harmful; false), evaluating a learning outcome of Body Busters, and a general question 3 (Antibiotics help cure colds; false) showed a significant decrease in knowledge (95%-92% and 61%-49%, respectively). There was no evidence of a significant knowledge change for questions 7 (You cannot infect other people around you through coughs and sneezes; false) and 11 (Vaccinations protect us from catching and spreading diseases), for which over 90% answered correctly before playing the games. Two other general questions, 1 (Which of these microbes causes coughs and colds?) and 2 (What is the best way to treat an infection with a virus?), did not see a significant change in knowledge.

**Discussion**

**Principal Findings**

This study indicates that playing the 2 e-Bug games had a significant ($P<.05$) positive effect on students’ knowledge on 6 out of the 12 questions:

1. Antibiotics are used to treat bacterial infections (question 4)
2. Antibiotics kill good and bad bacteria (question 10)
3. Most coughs and colds get better without antibiotics (junior $P=.06$; question 5)
4. The best way to stop microbes in coughs and sneezes spreading is to catch coughs and sneezes in a tissue and throw the tissue away (question 12)
5. The more people are vaccinated, the more people are protected from that infection (question 8)
6. By getting vaccinated, you can also protect others around you from infection (junior $P=.06$; question 9).

However, the games were indicated to have a detrimental effect on 2 true or false questions in the older students aged 11 to 16 years:

3. Antibiotics help cure colds (false)
6. All microbes are bad or harmful (false).
Suggestions for this detrimental effect include the following: question 3 was perhaps not obvious that viruses cause colds in either of the games, and question 6 had a very high baseline knowledge of 95%; therefore, it would have been difficult to see an improvement in knowledge. Modifications to the games will be required to address this detrimental effect, and developers should consider the age group of their target audience.

Another main finding of the study was the comparison between age groups of baseline knowledge and postintervention knowledge: junior (7-11 years) and senior (12-15 years). Baseline knowledge for senior students was higher than junior students on 11 out of the 12 questions; juniors scored 4.3% higher in question 9. Postintervention knowledge for senior students was higher than that for junior students in 10 out of the 12 questions; the 2 questions that were lower than juniors (2 and 9) were only lower by 0.3% and 4.3%, respectively. Positive knowledge change for juniors was greater in 9 out of 12 questions compared with senior students, suggesting that the e-Bug games had a greater impact on junior student knowledge; the researchers therefore recommend that the e-Bug games should be targeted at junior school–aged children and should be further promoted to this age group.

The high baseline knowledge for senior students could be a reflection that the questions were too easy for older students and perhaps senior students obtained other learning from the games, which researchers could have picked up with different or more difficult questions.

Overall, both junior and senior students reported Body Busters to be more enjoyable than Stop the Spread on the Likert scale responses and thematic analysis of the focus group transcripts; this is supported by Google Analytics. During the academic year, September 1, 2016, to August 31, 2017, junior and senior students viewed Body Busters (8905 and 3814 views, respectively) more than Stop the Spread (6803 and 3027 views, respectively). During the same academic year, on average, junior and senior students played Body Busters (01:40 and 01:51 min, respectively) more than Stop the Spread (01:01 and 01:36 min, respectively). This may be because Body Busters was easier and more enjoyable; however, Stop the Spread lead to a greater improvement in knowledge, particularly about appropriate sneezing behaviors and vaccinations in this study. Suggestions for improvements on both games were provided by students, and the e-Bug team will consider the suggestions when making modifications for improvement.

Qualitative focus groups enabled the exploration of a range of students’ views; it also brought synergism, snowballing of ideas, and stimulation of participants, which will assist in making improvements to the games.

The study evaluation allowed students to play each game in a classroom setting for only 5 min. This mimicked a real-life classroom setting where they would usually play Web-based games and discuss the games together, and 5 min is the usual length of time it takes to play 1 game. However, the 5 min of game play might not replicate the normal duration for game play; 5 min per game might not have been long enough for some students to gain the desired knowledge. For example, in Body Busters, viruses only appeared in level 3, and some students might have struggled to reach this level, which could partly explain any variation in percentage correct answers. The intervention in this study was used in isolation and perhaps the learning outcomes can be achieved better when the games are used as a tool to reinforce teaching about each topic in the classroom or in the home environment.

The questionnaires used in this study were based on questionnaires that have been used in previous e-Bug evaluations [2,3,16]. However, to eliminate any question style bias, future work with young students should use a questionnaire design that has the same format for each question, that is, all multiple-choice questions with 1 correct answer or all true or false questions to make it easier to understand for younger students.

Baseline knowledge was very high in senior students, especially about vaccinations and sneezing behaviors, so there is little need for improvements; however, modifications, including adding more levels to the games or adding extra learning outcomes, are required.

In the focus groups and evaluation questionnaires, junior students found it difficult to vocalize their thoughts beyond close-ended questions. Furthermore, many junior students found writing answers to the open responses on the evaluation form difficult to express their views, which was observed by researchers during data collection; however, data saturation was reached during the focus groups, suggesting no new themes would emerge.

Comparison With Existing Literature

Improvements in students’ knowledge after the delivery of an e-Bug lesson [2,3,4,6] and the e-Bug Web-based games [16] have been well documented. Our research adds to the body of literature to support the value of the e-Bug resources and Web-based games in educating children on hygiene and antibiotic topics. Limited research has been conducted about school-aged children’s knowledge of antibiotics and vaccination topics in England. One e-Bug evaluation across 3 European countries found that junior- and senior-aged students in 1 county in England had high baseline knowledge about the spread of infection (68%-78%) and low levels of baseline knowledge about the treatment and prevention of infection (29%-34%) [6]; this is reflected in this study as students had greater baseline knowledge about vaccinations than antibiotics generally.
A previous evaluation [16] of an earlier version of Body Busters showed that it increased knowledge of antibiotics in children, created a flow-like state in players, and was enjoyed the most out of the 3 games evaluated. Suggestions for modifications to the game included the following: more information in the introductory text, make the difference between viruses and bacteria more obvious, and create a steady increase in difficulty level as the game progresses [16]. The changes suggested by Hale et al were made in 2015 and aided the enjoyment reported by students in this study. However, this study is a much larger evaluation including more student questionnaires and focus group responses from a wider student sample across the United Kingdom; therefore, this evaluation provides new evidence to support the importance of the e-Bug project, for future modifications to the games and for future e-Bug game developments.

Systematic reviews and meta-analyses of gamification, serious games, and apps about health topics found that they can have a positive impact on health and well-being in the general population [17-20]. Other gamification studies reported positive associations between gamification and school-aged knowledge in several health topics; a randomized controlled trial (RCT) of a serious game promoting oral health found a significant improvement in knowledge of children after playing the game compared with before playing the game [13]. Likewise, an RCT for a game educating children about asthma found an increase in knowledge and improved attitudes at postintervention and follow-up as compared with control [11]. This study adds to the literature to support the positive effect of serious games on knowledge and attitudes of school children on important health issues. A meta-analysis of serious games for healthy lifestyle promotion was found to be appealing to individuals regardless of age or gender, showing this could be an intervention suitable for a more general audience than just children [20].

Other antimicrobial games have been developed but very few have been evaluated and on such a large scale. Recently, there has been an increase in serious games about the topic of microbiology and antimicrobial resistance, such as the Longitude prize’s Superbugs, showing this is a rapidly growing area of serious games. An infection prevention gamification tool called Germ Defence, encouraging individuals of any age to pledge to wash their hands more often, has been evaluated as part of a large RCT [21]. The RCT with over 20,000 participants found that those who used the Germ Defence website had fewer colds, flu, and stomach upsets than those who had not seen the website [21], providing further evidence for the positive effect of gamification on knowledge and health behaviors.

**Implications for Future Research**

Students had significant increases in knowledge about antibiotics, showing that e-Bug helps to reinforce the 2017 national antibiotics campaign keep antibiotics working. Students had significant increase in knowledge and behavioral intentions about appropriate sneezing behaviors, which directly support the 2013 national campaign catch it, bin it, kill it. Future national infection-related public health campaigns could link to the e-Bug games to encourage schools to use the games in their teaching and reinforce the campaigns.

Further research is required to investigate whether the knowledge gained from the e-Bug games is maintained or has changed future behavior. Additional qualitative research with teachers is needed to explore and understand how the e-Bug educational games can be used in a lesson to support learning.

e-Bug will continue to follow NICE guidance and work with educators and students to develop and promote resources for teaching children and young people about microbes, infection, and antibiotics in a fun and interactive way.

**Conclusions**

Science pedagogy Web-based games, including the e-Bug games, have the potential to engage and excite children and young people about important public health topics and aid in the learning of knowledge. To increase gaming, the e-Bug games should be both fun and challenging.

This study shows that 2 e-Bug educational games, Body Busters and Stop the Spread, covering learning topics about microbes, infection prevention, and antibiotics, are valuable to school-aged children’s knowledge. Body Busters is greatly enjoyed by and engaged school-aged children; a few modifications about antibiotics that can kill good and bad bacteria are required to reinforce learning outcomes. Stop the Spread is enjoyed by school-aged children to a lesser degree and more modifications, including slowing the game down, are required to retain user engagement. However, learning outcomes are very well covered in Stop the Spread. Health commissioning schools should target and promote the Body Busters and Stop the Spread e-Bug games, especially toward junior students (aged 7-11 years), as they showed the greatest improvement in knowledge. Further levels with more learning outcomes will facilitate increased learning in older students (aged 12-15 years).

**Acknowledgments**

The authors would like to thank the staff in the PHE Primary Care Unit, particularly Stephen Cooper and Rowshonara Syeda for their help in data collection and Dr Carla Brown for her comments on the manuscript. The authors would also like to thank Grove Street Media for their work in developing the 2 e-Bug games. The authors would like to thank all educational providers, teachers, and students who took part in the research. This work was supported by PHE.

**Conflicts of Interest**

At the time this study was conducted, CVE, VLY, CVH, and CAMM all worked for PHE to produce and disseminate e-Bug teaching resources for schools covering microbes, the spread, treatment and prevention of infection, and antibiotic resistance.
References


Abbreviations

NICE: National Institute of Health and Care
PHE: Public Health England
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 30.04.18; peer-reviewed by R Alkoudmani, M Roccetti; comments to author 19.07.18; revised version received 28.08.18; accepted 28.08.18; published 01.02.19

Please cite as:
Eley CV, Young VL, Hayes CV, Verlander NQ, McNulty CAM
Young People's Knowledge of Antibiotics and Vaccinations and Increasing This Knowledge Through Gaming: Mixed-Methods Study Using e-Bug
JMIR Serious Games 2019;7(1):e10915
URL: https://games.jmir.org/2019/1/e10915/
doi: 10.2196/10915
PMID: 30707096
Serious Games in Surgical Medical Education: A Virtual Emergency Department as a Tool for Teaching Clinical Reasoning to Medical Students

Seung-Hun Chon¹, MD; Ferdinand Timmermann²; Thomas Dratsch², DPhil; Nikolai Schuelpert³, MD, MME; Patrick Plum¹, MD; Felix Berlth¹, MD; Rabi Raj Datta¹, MD; Christoph Schramm⁴, MD; Stefan Haneder⁵, MD; Martin Richard Späth⁶,⁷, MD; Martin Dübbers¹, MD; Julia Kleinert¹, MD; Tobias Raupach⁸,⁹, MD, MME; Christiane Bruns¹, MD; Robert Kleinert¹, MD

¹Department of General, Visceral and Cancer Surgery, University Hospital of Cologne, Cologne, Germany
²University of Cologne, Cologne, Germany
³Department of Haematology and Medical Oncology, University Medical Centre Göttingen, Göttingen, Germany
⁴Department of Gastroenterology and Hepatology, University Hospital of Cologne, Cologne, Germany
⁵Institute of Diagnostic and Interventional Radiology, University Hospital of Cologne, Cologne, Germany
⁶Department II of Internal Medicine and Center for Molecular Medicine, University of Cologne, Cologne, Germany
⁷Cologne Excellence Cluster on Cellular Stress Responses in Aging Associated Diseases, University of Cologne, Cologne, Germany
⁸Department of Cardiology and Pneumology, University Medical Centre Göttingen, Göttingen, Germany
⁹Division of Medical Education Research and Curriculum Development, University Medical Centre Göttingen, Göttingen, Germany

Corresponding Author:
Seung-Hun Chon, MD
Department of General, Visceral and Cancer Surgery
University Hospital of Cologne
Kerpener Strasse 62
Cologne, 50937
Germany
Phone: 49 221 478 4864
Fax: 49 221 478 86227
Email: seung-hun.chon@uk-koeln.de

Abstract

Background: Serious games enable the simulation of daily working practices and constitute a potential tool for teaching both declarative and procedural knowledge. The availability of educational serious games offering a high-fidelity, three-dimensional environment in combination with profound medical background is limited, and most published studies have assessed student satisfaction rather than learning outcome as a function of game use.

Objective: This study aimed to test the effect of a serious game simulating an emergency department (“EMERGE”) on students’ declarative and procedural knowledge, as well as their satisfaction with the serious game.

Methods: This nonrandomized trial was performed at the Department of General, Visceral and Cancer Surgery at University Hospital Cologne, Germany. A total of 140 medical students in the clinical part of their training (5th to 12th semester) self-selected to participate in this experimental study. Declarative knowledge (measured with 20 multiple choice questions) and procedural knowledge (measured with written questions derived from an Objective Structured Clinical Examination station) were assessed before and after working with EMERGE. Students’ impression of the effectiveness and applicability of EMERGE were measured on a 6-point Likert scale.

Results: A pretest-posttest comparison yielded a significant increase in declarative knowledge. The percentage of correct answers to multiple choice questions increased from before (mean 60.4, SD 16.6) to after (mean 76.0, SD 11.6) playing EMERGE ($P<.001$). The effect on declarative knowledge was larger in students in lower semesters than in students in higher semesters ($P<.001$). Additionally, students’ overall impression of EMERGE was positive.

Conclusions: Students self-selecting to use a serious game in addition to formal teaching gain declarative and procedural knowledge.
Introduction

E-learning programs are now widely used for teaching declarative or theoretical knowledge [1,2]. One important part of clinical education is the transition of declarative knowledge to procedural knowledge [3], which is the basis for mastering clinical workflows in diagnosis and therapy or clinical reasoning. Virtual patient simulators [4] and serious games offer the possibility for teaching declarative and procedural [5] knowledge [6,7]. They are valuable in training specific medical scenarios that are hard to reproduce in the daily curriculum, such as rare clinical conditions, or are resource intensive (ie, major incident triage training) [8]. Even in common medical scenarios, serious games and virtual patient simulators enable virtual experience [9], as students can face the consequences of different decisions without putting real patients at risk.

Repetitive training allows the internalization of clinical patterns that are relevant for the necessary procedural performance [10]. Home-based distance learning with individual pace and number of repetitions allows equalization of students’ knowledge enabling effective face-to-face-teaching. Serious games extend the possibilities of virtual patient simulators because they are known to motivate students [11] as identification with an avatar (immersion) influences important incentives for intrinsic motivation, such as a sense of autonomy and a sense of achievement [12]. However, the amount of immersion depends on technical and, in particular, graphics quality, which should be “state of the art” [13,14]. The availability of serious games for clinical education is limited [15]. Commercial development of such a high-fidelity project is resource intensive [16], which may require subsidies because development outlay would necessitate high levels of sales to achieve profitability. In addition to technical quality, the success of serious games strongly depends on the amount of curricular content available [15]. Commercial serious games often lack medical content and are thus limited from a curricular perspective [17].

Therefore, it is desirable that the development of such educational methods is in the hands of universities because they are the main promoters of innovative educational methods and they are responsible for curricular content. However, the benefit of an educational tool of this kind is questionable if the impact on knowledge gain is still under discussion [11,15]. Moreover, most serious games are stand-alone solutions that cover only small aspects of clinical education [17] and are often not available outside published studies.

EMERGE is a serious game that combines both factors: it fosters training of clinical reasoning with state-of-the-art graphics that are likely to facilitate high immersion. We recently demonstrated the noninferiority of EMERGE compared to small-group teaching with respect to student learning outcome on clinical reasoning [18]. The aim of this study is to test EMERGE as an educational tool and determine the effect on student motivation and knowledge gain.

Methods

Description of the Serious Game

EMERGE allows free navigation through a virtual emergency department. A digital mentor supports the student while dealing with the interface and treating patients. When starting the simulation, students get information about the incoming patient from an emergency physician, they dispatch the patient to an examination room, take the patient’s medical history, order diagnostic tests, and establish a diagnosis and treatment. When taking the patient’s medical history, students are free to choose from an alphabetical list of 70 questions. Students can ask the questions in any order. If one of the questions is repeated, the virtual patient always responds with the same answer. Students are free to choose any diagnostic test or treatment that a modern emergency department offers. Students are not restricted to medically indicated tests nor to the sequence for taking these tests. All medical decisions have real consequences. For instance, patients respond to treatments and medications have effects and side effects. EMERGE was developed at Göttingen Medical School in collaboration with the University Medical Centre Hamburg-Eppendorf ensuring the educational quality of the program. EMERGE was programmed by PatientZero Games GmbH, at a cost of approximately €200,000.

EMERGE displays the classic features of a modern computer game: free interaction between the player and the game, a certain challenge (ie, a critically ill patient), a game story as the student plays a virtual doctor, and high immersion owing to high-fidelity graphics and an intuitive graphical user interface. A demo of EMERGE is freely accessible [17,19].

Participants

Students of all clinical years (5th to 12th semester) of medical education at Cologne University (Cologne, Germany) were invited to sign up for the study via mailing lists, flyers, and social networks. Each semester consists of approximately 200 students; thus, 1600 students were invited. A total of 140 medical students (46 males, 94 females; mean age 24.1 years, age range 20-33 years; response rate 140/1600, 8.8%) volunteered to participate (see Table 1 for demographic information). Overall 67.1% (94/140) of our participants were female medical students. This distribution of male and female medical students is in line with the national average in Germany with approximately two-thirds of medical students being female [20]. Students received €15 for their participation. The Ethics Committee of the Medical Faculty at the University of Cologne approved the evaluation. The Institutional Review Board was informed, and there were no objections.
Table 1. Demographic information about the students who played EMERGE (N=140).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of students, n</td>
<td>46</td>
<td>94</td>
</tr>
<tr>
<td>Age, mean (SD)</td>
<td>24.5 (2.6)</td>
<td>23.8 (2.7)</td>
</tr>
<tr>
<td>Semester, n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5th semester</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>6th semester</td>
<td>10</td>
<td>28</td>
</tr>
<tr>
<td>7th semester</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>8th semester</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td>9th semester</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>10th semester</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Practical year (11th and 12th semester)</td>
<td>12</td>
<td>6</td>
</tr>
</tbody>
</table>

Study Design
The effect of EMERGE on declarative and procedural knowledge was analyzed in a pretest-posttest design. Students first completed the questions measuring procedural and declarative knowledge. After students had played EMERGE, they again completed the questions measuring procedural and declarative knowledge. Additionally, after students had finished playing EMERGE, they also answered several questions measuring their impressions of EMERGE.

Procedure
In December 2017 and January 2018, 35 gaming sessions lasting 90 minutes each were conducted with four students allocated to each session (only four computers were available for this study). At the beginning of the experiment, students were greeted by the experimenter (FT). Each computer corresponded with one of four different case scenarios. Students were unaware of the associations between computers and case scenarios. Students were free to decide which patient to interact with first. The time to play EMERGE was not limited. Students played for a mean 31.7 (SD 7.4) minutes. The game ended after students had diagnosed and treated the patients or transferred them to another department for treatment. At the end of the game, students were debriefed by a virtual doctor, who informed them about the correct diagnosis and treatment for each patient. After the students had finished playing EMERGE, they again completed the 20 multiple choice questions and the written Objective Structured Clinical Examination (OSCE) questions. Students were given 40 minutes to complete the questions. After that, students were instructed to imagine being a first-year intern in a real emergency department. Students were further instructed to behave as if their decisions had real consequences. After the instructions, the experimenter launched EMERGE and introduced them to the general controls of EMERGE. A virtual map showed students the way to one of the various rooms within the department (Figure 1), where two virtual patients were already waiting for them. Students were free to decide which patient to interact with first. The time to play EMERGE was not limited. Students played for a mean 31.7 (SD 7.4) minutes. The game ended after students had diagnosed and treated the patients or transferred them to another department for treatment. At the end of the game, students were debriefed by a virtual doctor, who informed them about the correct diagnosis and treatment for each patient. After the students had finished playing EMERGE, they again completed the 20 multiple choice questions and the written OSCE questions.
Medical Content

To increase the generalizability of the study and to ensure that the results would not be limited to just one condition, we created four distinct clinical cases: pneumothorax, sigmoid diverticulitis, mesenteric ischemia (easy), and mesenteric ischemia (difficult). These cases were selected because they are part of the standard surgical curriculum for German medical students [21]. At the University of Cologne, the surgical curriculum is spread out over several semesters. In the second clinical semester, students attend the main surgical lecture, covering all the declarative knowledge for the diseases used in this study. Additionally, students complete the first part of surgical bedside teaching. In the fifth clinical semester, students complete the second part of surgical bedside teaching. Moreover, as a part of the curriculum at the University of Cologne, clinical subjects are taught throughout the whole 6 years (12 semesters) of medical school, which results in students being presented with similar aspects several times during their training. For all diseases, correct diagnostic and treatment patterns were defined in process charts, as described in a previous article [22]. Briefly summarized, these process charts served as a blueprint for optimal diagnostic and therapeutic workflows. Congruence between student pathways and blueprints formed the basis for the assessment of student performance. Blueprints were based on published guidelines for each disease [23-25]. The four different clinical cases were combined into four different case scenarios (see Textbox 1). In each case scenario, students were presented with two medical conditions. Because the diagnosis of mesenteric ischemia requires profound clinical reasoning and experience, this condition was divided into an easy and a difficult case. In the easy case, serum lactate was high, and the patient suffered from an arrhythmia. These signs are strong indicators of this clinical condition. However, there is no laboratory value that is associated with mesenteric ischemia. A functional liver is able to eliminate lactate, even when there is a severe ischemic condition. Due to restraints in the recruiting of participants for the study, only four different case scenarios could be tested. The cases of pneumothorax, sigmoid diverticulitis, and mesenteric ischemia (easy) were each used twice in the study. The case of mesenteric ischemia (difficult) was only used once. Because only 140 students participated in the study, it was not possible to create a fifth case scenario with an additional 35 participants to repeat the mesenteric ischemia (difficult) case.

Figure 1. Sample screenshots of EMERGE. Students could freely interact with the environment.
### Textbox 1. Combinations of the different clinical cases in the four groups (N=140).

- **Case scenario 1 (n=35)**
  - Pneumothorax
  - Sigmoid diverticulitis

- **Case scenario 2 (n=35)**
  - Mesenteric ischemia (easy)
  - Sigmoid diverticulitis

- **Case scenario 3 (n=35)**
  - Pneumothorax
  - Mesenteric ischemia (easy)

- **Case scenario 4 (n=35)**
  - Mesenteric ischemia (easy)
  - Mesenteric ischemia (difficult)

### Impact on Declarative Knowledge

Students gain in declarative knowledge was determined by asking 20 multiple choice questions before and immediately after working with EMERGE. Influence of preexisting knowledge (ie, concordance validity) was measured by subgroup analysis with regard to study year.

### Impact on Procedural Knowledge

Students’ gain in procedural knowledge was measured by presenting students with a modified clinical case from OSCE describing a patient with sigmoid diverticulitis. This disease was intentionally selected for this analysis. Based on our own personal experience as OSCE examiners at our university, this case was the most complex of all the vignettes. Because the study lasted several weeks, it was expected that participants would exchange knowledge and experience. In order not to bias participants toward sigmoid diverticulitis, only half of the students playing EMERGE were presented with a patient with sigmoid diverticulitis in the game. In this subgroup, we were able to measure the effect of treating a patient with sigmoid diverticulitis on procedural knowledge. Time to read the case was not limited. After students read the case, they were asked the following six questions:

1. Please list the five most likely potential diagnoses for the patient.
2. What additional anamnestic questions would you ask the patient?
3. What procedures would you order to confirm your diagnosis and in what order?
4. How would you treat perforated sigmoid diverticulitis?
5. How would you treat nonperforated sigmoid diverticulitis?
6. How would you treat appendicitis?

Students had to write down their answers. Answers were compared to a blueprint created by two expert physicians (RK and SC) based on current medical guidelines and scored accordingly. Answers were scored by FT. Critical cases were discussed among all three authors (RK, SC, and FT) and resolved. Interrater reliability was not calculated.

An overview of the experimental set-up can be seen in Figure 2. Pretest and posttest multiple choice questions and OSCE case were identical before and after EMERGE.

Figure 2. Testing declarative and procedural knowledge before and after working with EMERGE.

![Figure 2](http://games.jmir.org/2019/1/e13028/)

http://games.jmir.org/2019/1/e13028/
Students’ Impressions of EMERGE

Students also rated 10 statements to measure three different aspects of their experience using EMERGE: (1) overall impression of EMERGE, (2) usability of EMERGE, and (3) student attitudes toward e-learning and computers. These 10 items had been used in a previous study on a Web-based immersive patient simulator [22]. Students rated these aspects on a 6-point Likert scale ranging from 1 (fully agree) to 6 (fully disagree). The overall impression of EMERGE was measured using the following statements: “Using EMERGE is fun,” “EMERGE teaches new knowledge,” “EMERGE prepares me for clinical practice,” “I would use EMERGE,” and “My overall impression of EMERGE.”

The usability of EMERGE was measured using the following statements: “EMERGE is easy to learn” and “EMERGE is easy to use.”

Student attitudes toward e-learning and computers were measured using the following statements: “I use computers on a daily basis,” “Computers, consoles, and cell phones are my hobby,” and “I mostly learn with books” (reversed item).

Data Analysis

To assess students’ gain in declarative knowledge, students were presented with 10 multiple choice questions per clinical case. Because each student was presented with two clinical cases, students answered a total of 20 multiple choice questions. Students completed the same 20 questions before and after playing EMERGE. For each student, we calculated the percentage of questions they answered correctly. The percentage of correctly answered questions was analyzed in a mixed ANOVA. Effect sizes for ANOVAs are given as partial eta squared (η²). To determine the sample size for the study, a statistical power analysis was performed. This statistical power analysis was not based on data from prior studies but on general considerations about the trade-offs between the ability to detect certain effects and the feasibility to acquire a sufficiently large sample. Given a specified effect size (Cohen d), power, and alpha, the sample size was calculated using GPower (GPower 3.1). Because the implementation of new teaching methods can be expensive and time-consuming, potential new teaching methods should have a sufficiently large benefit. However, promising new teaching methods should not be overlooked. As a compromise, our study needed to be sufficiently powered to detect medium-sized effects with a Cohen d of 0.3 for the within-group comparisons. With an alpha=0.05 and power=0.80, the projected sample size needed to detect an effect of Cohen d of 0.3 for the within-group comparisons was N=90 (GPower 3.1). With an alpha=0.05 and power=0.80, the projected sample size needed to detect a large effect (Cohen d of 0.5) for the between-group comparisons was N=64 per group (GPower 3.1).

To assess students’ gain in procedural knowledge, we presented students with a clinical vignette (sigmoid diverticulitis) and asked them six questions about the vignette. Students provided their answers in written form. Because only half the students playing EMERGE were presented with a patient with sigmoid diverticulitis in the game, we were able to measure the effect of treating a patient with sigmoid diverticulitis on procedural knowledge.

Students’ answers to the first question were graded as correct if they had listed the correct diagnosis among the top three diagnoses. The gain in procedural knowledge was assessed by comparing the number of students who listed the correct diagnosis among the top three diagnoses before playing EMERGE to the number of students who listed the correct diagnosis among the top three diagnoses after playing EMERGE. The number of students before and after playing EMERGE were compared using generalized estimating equations (GEEs).

In the second question the number of diagnostic questions asked before and after playing EMERGE was analyzed using a mixed ANOVA.

Students’ answers to the third question were graded as correct if they had listed all diagnostic procedures in the correct order. Correct answers before and after playing EMERGE were compared using GEEs.

Students’ answers to the fourth, fifth, and sixth questions were graded as correct if they provided the correct treatment for each condition. Correct answers before and after playing EMERGE were compared using GEEs.

For all correlational analyses, Kendall tau (τ) was used as a robust measure of correlation. Data were analyzed using SPSS version 25 (IBM Corp, Armonk, NY, USA).

Results

Knowledge Gain: Impact on Declarative Knowledge

Working with EMERGE showed a positive impact on declarative knowledge. The percentage of correct answers to multiple choice questions increased from before (mean 60.4, SD 16.6) to after (mean 76.0, SD 11.6) playing EMERGE (t_{139}=-13.92, P<.001, d=1.25).

To test whether declarative knowledge differed between students in lower or higher semesters, we conducted a 2×7 mixed ANOVA (time×semester) as an exploratory analysis. There was a significant main effect of time on the percentage of correct answers (F_{6,133}=130.67, P<.001, partial η²=.496). This suggests that participants answered significantly more questions correctly after playing EMERGE compared to before playing. There was also a significant main effect of semester (F_{6,133}=4.76, P<.001, partial η²=.177), which suggests that students in higher semesters answered significantly more questions correctly than students in lower semesters. There was also a significant interaction effect between time and semester (F_{6,133}=18.36, P<.001, partial η²=.453), which suggests that the knowledge gain after playing EMERGE was larger for students in lower semesters (see Figure 3). This was further supported by a negative significant correlation between knowledge gain (percentage of correct answers after playing EMERGE minus percentage of correct answers before playing EMERGE) and semester (Kendall τ_{140}=-0.298, P<.001).
Knowledge Gain: Impact on Procedural Knowledge

Students’ gain in diagnostic accuracy was analyzed using GEEs in a 2×2 design (time×sigmoid case). There were no significant effects (all Wald $\chi^2_1 <1.4$). For descriptive results, see Table 2.

To test whether there was a gain in procedural knowledge, we compared the number of diagnostic questions asked in the sigmoid diverticulitis case before and after playing EMERGE in a 2×2 mixed ANOVA (time×sigmoid case). There was a significant main effect for time ($F_{1,138}=20.20$, $P<.001$, partial $\eta^2=.128$). All other effects were not significant (all $F<.775$).

This suggests that students asked significantly more diagnostic questions after playing EMERGE (mean 3.11, SD 1.65) compared to before (mean 2.42, SD 1.38), regardless of whether they had been presented with a sigmoid diverticulitis case in the game or not.

To test whether the order of correct diagnostic procedures improved after playing EMERGE, we conducted an analysis using GEEs in a 2×2 design (time×sigmoid case). There was a significant main effect of time (beta=–2.54, Wald $\chi^2_1=13.5$, $P<.001$). All other effects were not significant (all Wald $\chi^2_1<2.8$). This suggests that more students chose the correct diagnostic pathway after playing EMERGE (33/140, 23.6%) compared to before playing EMERGE (7/140, 5.0%), regardless of whether they had been presented with a sigmoid diverticulitis case in the game or not.

To test whether the treatment suggestions for the perforated sigmoid diverticulitis case improved after playing EMERGE, we conducted an analysis using GEEs in a 2×2 design (time×sigmoid case). There were no significant main or interaction effects (all Wald $\chi^2_1<0.9$). For descriptive results, see Table 2.

To test whether the treatment suggestions for the nonperforated sigmoid diverticulitis case improved after playing EMERGE, we conducted an analysis using GEEs in a 2×2 design (time×sigmoid case). There was a significant effect of time (beta=–2.88, Wald $\chi^2_1=21.8$, $P<.001$). There was also a significant effect of sigmoid case (beta=–2.46, Wald $\chi^2_1=14.7$, $P<.001$). There was also a significant interaction effect (beta=2.51, Wald $\chi^2_1=14.5$, $P<.001$). To break down this interaction effect, chi-square tests were performed. A significantly larger proportion of students who had been presented with a sigmoid case in EMERGE, 67 of 70 participants (96%), provided correct treatment suggestions after playing EMERGE than students who had not been presented with a sigmoid case, 46 of 70 participants (66%; $\chi^2=20.2$, $P<.001$). There was no significant difference between both groups before playing EMERGE ($\chi^2=0.0$, $P=.87$).
Table 2. Descriptive results of Objective Structured Clinical Examination before and after playing EMERGE for students presented and not presented a sigmoid diverticulitis case while playing EMERGE.

<table>
<thead>
<tr>
<th>Questions</th>
<th>Before playing EMERGE, n (%)</th>
<th>After playing EMERGE, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No sigmoid case (n=70)</td>
<td>Sigmoid case (n=70)</td>
</tr>
<tr>
<td>Correct diagnosis</td>
<td>61 (87)</td>
<td>65 (93)</td>
</tr>
<tr>
<td>Correct diagnostic procedures</td>
<td>5 (7)</td>
<td>2 (3)</td>
</tr>
<tr>
<td>Correct treatment of perforated sigmoid diverticulitis</td>
<td>66 (94)</td>
<td>67 (96)</td>
</tr>
<tr>
<td>Correct treatment of nonperforated sigmoid diverticulitis</td>
<td>40 (57)</td>
<td>39 (56)</td>
</tr>
<tr>
<td>Correct treatment of appendicitis</td>
<td>67 (96)</td>
<td>66 (94)</td>
</tr>
</tbody>
</table>

Table 3. Mean ratings of the experience of using EMERGE (1=fully agree to 6=fully disagree).

<table>
<thead>
<tr>
<th>Items</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using EMERGE is fun</td>
<td>1.68 (0.84)</td>
</tr>
<tr>
<td>EMERGE teaches new knowledge</td>
<td>1.98 (0.96)</td>
</tr>
<tr>
<td>EMERGE prepares me for clinical practice</td>
<td>1.95 (0.98)</td>
</tr>
<tr>
<td>I would use EMERGE regularly</td>
<td>2.22 (1.18)</td>
</tr>
<tr>
<td>My overall impression of EMERGE</td>
<td>1.91 (0.69)</td>
</tr>
<tr>
<td>EMERGE is easy to learn</td>
<td>1.61 (0.89)</td>
</tr>
<tr>
<td>EMERGE is easy to use</td>
<td>1.64 (0.91)</td>
</tr>
<tr>
<td>I use computers on a daily basis</td>
<td>1.63 (1.25)</td>
</tr>
<tr>
<td>Computers, consoles, and cell phones are my hobby</td>
<td>2.86 (1.52)</td>
</tr>
<tr>
<td>I mostly learn with books</td>
<td>3.49 (1.32)</td>
</tr>
</tbody>
</table>

Table 4. Correlations (Kendall tau) between semester and mean ratings of EMERGE.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Overall impression of EMERGE</th>
<th>Usability of EMERGE</th>
<th>Attitudes toward e-learning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Tau</td>
<td>P value</td>
<td>Tau</td>
</tr>
<tr>
<td>Semester</td>
<td>.195</td>
<td>.002</td>
<td>0.058</td>
</tr>
<tr>
<td>Overall impression of EMERGE</td>
<td>0.218</td>
<td>.001</td>
<td>0.122</td>
</tr>
<tr>
<td>Usability of EMERGE</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To test whether the treatment suggestions for the appendicitis case improved after playing EMERGE, we conducted an analysis using GEEs in a 2×2 design (time×sigmoid case). There were no significant main or interaction effects (all Wald $\chi^2<0.2$), which suggests that there was no difference in the proportion of students who listed the correct treatment for appendicitis before (133/140, 95.0%) and after (135/140, 96.4%) playing EMERGE.

**Impression Ratings of EMERGE**

Students rated their overall impression of EMERGE (Cronbach alpha=.805; mean 1.95, SD 0.71) the usability of EMERGE (Cronbach alpha=.851; mean 1.62, SD 0.84), and their attitude toward e-learning (Cronbach alpha=.518; mean 2.66, SD 0.98). Table 3 shows the mean ratings of the students’ impressions of EMERGE. As Table 4 shows, there was a significant correlation between semester and overall impression of EMERGE, indicating that students in lower semesters had a more positive impression of EMERGE than students in higher semesters. There was also a significant correlation between the overall impression of EMERGE and the perceived usability of EMERGE.

**Discussion**

The main goal of this study was to test EMERGE as an educational tool and to determine its effect on student motivation and knowledge gain. A pretest-posttest comparison yielded a significant increase in declarative knowledge. An exploratory analysis showed that this increase in declarative knowledge was smaller in more advanced students. With regard to procedural knowledge, a significant effect was noted for one of six categories, and this effect was only present in students who had actually been exposed to the respective case while playing the
game. Only students that worked with the perforated sigmoid diverticulitis case also showed an increase in performance in the OSCE questions about nonperforated sigmoid diverticulitis, which is a sign for positive concordance validity [26].

There are several potential shortcomings that limit the generalizability of the results. For instance, the results may be influenced by the fact that participation in this study was on a voluntary basis and included mainly motivated students. Because motivated students are more likely to participate in experiments [27], future studies should investigate if less motivated students could also benefit from playing EMERGE. One way to study EMERGE’s effects on both motivated and less motivated students would be to integrate EMERGE into the regular curriculum.

Another limitation of our study is the fact that we did not study any long-term effects. Because our study focused only on a single intervention, future studies are needed to investigate if the effects of playing EMERGE can also be found after a longer period of time.

Also, we cannot disentangle the effects of EMERGE and other curricular activities because this was not a randomized controlled trial. It is important to note that the surgical curriculum is spread out over several semesters. Therefore, it is possible that some of the participants may have been exposed to some of the clinical conditions included in this study. However, because we did find a positive effect of playing EMERGE on knowledge gain, prior exposure was not strong enough to cancel out all potential learning effects.

For this study, case scenarios were limited to three diseases. Although students playing EMERGE did show an increase in declarative and—to a lesser extent—procedural knowledge, this was only true for the four clinical vignettes included in the study and procedural knowledge was only assessed in relation to one condition (ie, sigmoid diverticulitis). Thus, future studies should include a wider range of clinical cases.

However, it is important to identify the effect that leads to the knowledge gain. First, because this study lacks a control group, it is not clear what part of the knowledge gain can be attributed to playing EMERGE and what part is due to testing effects (ie, a gain in test scores because of familiarity with the questions). Second, at the end of the game students were debriefed by a virtual doctor who informed them about the correct diagnosis, so we cannot untangle the effect of playing EMERGE and the debriefing on knowledge gain.

Overall, students’ impression of EMERGE was positive. However, students in lower semesters had a more positive impression of EMERGE than students in higher semesters. Future studies should include more detailed measures of students’ impressions to further untangle which groups of students are most likely to benefit from using EMERGE.

Beneath technical feasibility, curricular content of serious games plays a crucial role. Prior research on the impact of serious games on knowledge gain has been mixed—especially with regard to different levels of pedagogical strategies [15]. Therefore, using a third-party game in the curriculum may be risky because combining established pedagogical strategies with those used in serious games may be challenging. Hence, clinical teachers must adapt the cases in the game to their own curriculum to get a positive effect on knowledge gain.

Several prior studies have also shown positive effects of training with virtual patient simulators. For instance, Haubruck et al [28] showed that students who trained in chest tube insertions with the app Touch Surgery performed better than students who trained an unrelated skill. Additionally, Kowalewski et al [29] showed that virtual patient simulators have face and content validity. Also, Schwarz et al [30] showed that students generally have positive attitudes toward virtual platforms that train adaptive algorithms.

EMERGE shares certain features with immersive patient simulators [21]. However, it offers a higher degree of immersion because it expands on immersive patient simulators by adding features found in modern computer games, such as a high-fidelity game environment and the challenge of treating several patients in a limited amount of time. This opens the possibility of teaching nontechnical skills that are important in high-risk environments, such as the emergency department [22], but are hard to develop in an academic setting.

Lastly, working with EMERGE was widely accepted among students and was rated to be highly enjoyable by the students, which is an important precursor for motivated learning. Although this finding is commonly reported when evaluating serious games [31], it is still an important finding because, in addition to considerations of validity, creating enjoyable and motivating teaching interventions may be a goal in and of itself.

In this nonrandomized trial, we found significant gains in declarative and procedural knowledge in students self-selecting to use a serious game. Future studies need to determine whether this gain is attributable to the game itself.

Acknowledgments
We would like to thank Ms Claire Cahm for proofreading and Mr Juan Soriano Nunez for overseeing the experiment.

Conflicts of Interest
None declared.

References
None declared.


Abbreviations

GEE: generalized estimating equation
OSCE: Objective Structured Clinical Examination
Energy Expenditure and Enjoyment During Active Video Gaming Using an Adapted Wii Fit Balance Board in Adults with Physical Disabilities: Observational Study

Laurie A Malone¹, PhD, FACSM; Mohanraj Thirumalai¹, PhD; Sangeetha Padalabalanarayanan¹, MS; Whitney N Neal¹, MA, CHES; Sean Bowman¹, MPH; Tapan Mehta¹, PhD

University of Alabama at Birmingham/Lakeshore Foundation Research Collaborative, Birmingham, AL, United States

Corresponding Author:
Laurie A Malone, PhD, FACSM
University of Alabama at Birmingham/Lakeshore Foundation Research Collaborative
1720 2nd Avenue South
Birmingham, AL, 35294-1212
United States
Phone: 1 205 934 0142
Email: lamalone@uab.edu

Abstract

Background: Individuals with physical disabilities have fewer opportunities to participate in enjoyable physical activity. One option for increasing physical activity is playing active video games (AVGs); however, many AVGs are inaccessible or offer limited play options.

Objective: This study aimed to examine energy expenditure and enjoyment in adults with mobility impairment during AVG play using off-the-shelf (OTS) and adapted versions of the Wii Fit balance board (Nintendo).

Methods: During visit 1, participants completed a functional assessment and the familiarization period. During visit 2, metabolic data were collected during a 20-minute baseline and four 10-minute bouts of Wii Fit Plus game play, with two bouts on each of the boards. During the resting period, participants completed the Physical Activity Enjoyment Scale (PACES). Statistical analyses were computed using SPSS software. Data were analyzed separately for individuals who were able to play while standing on both boards (StdStd); those who could not play while standing on the OTS board, but were able to play while standing on the adapted board (aStd); and those who could only play while sitting on the adapted board (aSit).

Results: Data were collected for 58 participants (StdStd, n=17; aStd, n=10; aSit, n=31). The sample included 31 men and 27 women with a mean age of 41.21 (SD 12.70) years. Energy expenditure (metabolic equivalent [MET]) during game play was significantly greater than that during rest for all players. Only 17 participants (StdStd group) were able to play using the OTS board. During game play on the adapted board, the average MET values for the two game sets were 2.261 (SD 0.718) kcal/kg/hour and 2.233 (SD 0.751) kcal/kg/hour for the aSit group, 3.151 (SD 1.034) and 2.990 (SD 1.121) for the aStd group, and 2.732 (SD 0.655) and 2.777 (SD 0.803) for the StdStd group. For game play on the adapted board, self-reported ratings of perceived exertion on a 0-10 scale suggested greater exercise intensity levels, with median scores ranging from moderate (3) to very hard (7). The PACES scores indicated that all players enjoyed using the adapted board, with a median score of 4 on a 5-point scale.

Conclusions: The adapted Wii Fit balance board provided an opportunity for individuals with mobility impairments, including wheelchair users, to engage in AVG. All participants were able to utilize the adapted controller and enjoyed the AVG activity. Although the average MET values achieved during AVG represented light-intensity exercise (<3 METs), 16% of sitting participants and 41% of standing participants achieved moderate-intensity exercise (3-6 METs) in at least one of the games. Factors not accounted for, which may have influenced the intensity of exercise, include game selection, limited familiarization period, and discomfort wearing the COSMED portable metabolic system for measurement of oxygen consumption. Accessible AVG controllers offer an innovative approach to overcome various barriers to participation in physical activity. The next steps include assessment of an AVG intervention using an adapted board gaming controller on health and fitness outcomes.


doi:10.2196/11326

JMIR Serious Games 2019;7(1):e11326

http://games.jmir.org/2019/1/e11326/
**KEYWORDS**

exergaming; video games; exercise; physical activity; disability; energy expenditure; enjoyment

**Introduction**

Physical inactivity is significantly higher in people with disabilities due to fewer opportunities and countless barriers to engaging in leisure-time physical activity [1-7]. Lack of transportation, inaccessible fitness facilities, absence of staff trained to work with people with disabilities, and boredom associated with the use of standard exercise equipment contribute to higher sedentary behaviors in this population [1-7].

Replacing sedentary behaviors with active video games (AVGs) holds promise as a way to reduce those barriers and increase leisure-time physical activity in people with disabilities [8-12]. Moreover, AVGs have the potential to be a “gateway experience” to physical activity, suggesting that such games may open the door to interest and participation in other forms of physical activity for persons with disabilities [10].

AVGs, also called “exergames,” are video games that require actions of large body parts (like the trunk or upper or lower extremity) or the whole body to control game play as opposed to games that only require hand or finger movements to play. Gaming systems that have AVG capability include the Nintendo Wii, Sony PlayStation Move, and Microsoft Xbox Kinect. Depending on the design of games and control systems used, AVGs have been shown to be an enjoyable leisure-time physical activity option to replace sedentary screen time [13-17], with the potential to increase cardiorespiratory fitness and enhance balance and functional mobility [15,16,18-28]. Furthermore, AVGs show promise as an enjoyable physical activity alternative for individuals with disabilities, largely because they overcome certain common barriers to physical activity such as transportation and access to adequate facilities [10,11].

It is important to determine whether the level of physical activity offered by AVGs is high enough to achieve the same fitness and health benefits offered by traditional exercise. Several studies have reported that increases in energy expenditure are sufficient for maintaining and improving health in individuals with mobility impairments such as cerebral palsy [12,29-32], spinal cord injury [33,34], and stroke [29,35-37]. A literature review by Deutsch and colleagues [29] sought to determine whether evidence existed to support the use of video games for the promotion of fitness and wellness for adults after stroke (moderate severity) and those with cerebral palsy (mild severity). It was found that both groups were able to achieve moderate energy expenditure by playing Nintendo Wii, Sony Playstation, and Xbox Kinect games. In another study by Hurkmans et al [31], adults with cerebral palsy who were able to stand without support were asked to play Wii Sports tennis and boxing while their energy expenditure was measured using a portable gas analyzer. The results of the study showed that participants achieved moderate-intensity exercise during both games. A similar study by Robert et al [32] evaluated exercise intensity in children with cerebral palsy who were able to stand and children without cerebral palsy. They found similar exercise intensity levels for both groups, suggesting that children with cerebral palsy who are able to play video games while standing can obtain exercise-related benefits that are similar to those obtained by children without cerebral palsy. A systematic review by Mat Rosly et al [33] found that adults with spinal cord injury were able to achieve the recommended moderate-to-vigorous physical activity levels proposed by the American College of Sports Medicine. A case study of two young adults showed that persons with spinal cord injury are able to achieve moderate intensity–level heart rates when playing boxing on the Nintendo Wii while seated [34]. Furthermore, when compared with conventional boxing, Mat Rosly et al [38] found that individuals with spinal cord injury were able to achieve the same moderate-intensity level as that when playing exergame boxing. Another study determined showed poststroke, individuals were able to play Wii tennis and Wii boxing at a moderate-intensity level (classified using the American College of Sports Medicine and the American Heart Association guidelines as 3-6 metabolic equivalents [METs]) [35]. Another study compared the performance of playing Nintendo Wii and Sony Xbox 360 in adults poststroke and individuals without a disability [36]; the study found that AVG game play while standing produced moderate-intensity exercise levels and poststroke participants approached anaerobic metabolism when playing in the sitting position.

Although AVGs hold promise as an opportunity to increase physical activity in people with disabilities, physical limitations such as decreased motor control, range of motion, muscle strength, ambulatory status, and balance limit AVG accessibility for a large portion of this population [10,12,39]. For example, the majority of the studies noted above for individuals with cerebral palsy and poststroke comprised individuals with mild-to-moderate mobility impairments who were able to play the AVGs without adapted equipment. Similarly, limited AVG play options are available for people who are unable to stand, have balance problems or poor motor control, or cannot use their lower body to perform game movements [10]. For instance, floor-pad game controllers used by AVGs (eg, Dance, Dance Revolution and Wii Outdoor Challenge) have obvious accessibility limitations. Although motion-controlled AVGs offer slightly greater access, accelerometer-based hand controllers like those used by the Sony PlayStation Move and Nintendo Wii platforms often require rapid and precise movements for successful play, and AVGs using the camera-based controller of Microsoft Kinect typically require the player to be standing for proper game function. Therefore, AVG adaptations to game controllers are essential in order to offer people with disabilities options for AVGs for moderate-to-vigorous exercise in environments such as their home and communities [11]. Offering adapted controllers for AVG play to people with disabilities is also a necessary first step toward examining the feasibility of such controllers for increasing energy expenditure.

The development of adapted game hardware not only offers an innovative approach to overcoming numerous barriers to exercise in people with disabilities, but also provides an enjoyable form of exercise to this population. Successful
adaptations to game controllers and interfaces that allow people with disabilities to play video games have been developed [40-43]; however, there has been limited research and development efforts focused on improving the accessibility of commercially available gaming controllers for use with AVGs. The Rehabilitation Engineering Research Center on Interactive Exercise Technologies and Exercise Physiology for People with Disabilities at the University of Alabama at Birmingham/Lakeshore Foundation Research Collaborative examined the accessibility of video game controllers including the Wii Fit balance board system. Data on game play, participants’ ability to use the controllers, user feedback, and research staff qualitative observations indicated that the Wii balance board was in need of adaptation for successful game play. These data were fed to the engineering team for development of an adapted gaming balance board.

The primary deficiencies in the balance board included a small platform area (19.5"×12"), the inability to use a stabilization assistive device (ie, walker or cane), and the requirement of a full range of motion for responsive game play. To address these deficiencies and increase accessibility, the balance board was redesigned to feature a much larger platform area (40"×38"), built-in lateral stabilization supports (ie, handrails), and adjustable sensitivity for center of balance support [44]. The adaptations were selected to not only enable wheelchair users to use the balance board, but to make it a universal device with enhanced safety for all users (Figure 1).

Usability of the off-the-shelf (OTS) and adapted balance board controllers was evaluated in individuals with mobility impairments using the System Usability Scale (SUS) [44]. The user-centered design approach resulted in an adapted version of the Wii Fit balance board, which met the needs of a variety of users. Results demonstrated a successful adaptation and increase in usability of the adapted balance board, with the adapted board scoring significantly greater mean SUS scores than those of the off-the-shelf board.

Despite the fact that a variety of successful adaptations to AVG game controllers and interfaces that allow people with disabilities to play AVGs have been developed by rehabilitation engineers and assistive technology specialists, there have been limited research and development efforts focused on measuring energy expenditure. Modifications of these game controllers must be successful in allowing people with disabilities to not only play AVGs, but also achieve levels of energy expenditure that are similar to the levels of people without disabilities.

Figure 1. Adapted Wii Fit balance board with a ramp for wheelchair access, adjustable-height handrails, and control box. Usable space on the adapted balance board measures 91.5 cm x 91.5 cm, with the load cells placed at each of the four corners. The electrical components of the off-the-shelf Wii Fit balance board were reconfigured and integrated into the new form factor and electrical design of the adapted board.
The purpose of this study was to examine energy expenditure and enjoyment in adults with physical disabilities, specifically those with mobility impairments (ie, inability to stand, balance issues, poor motor control, and inability to use lower extremity for game play), during AVG play using OTS and adapted versions of the Wii Fit balance board.

Methods

Design and Setting

The study was conducted in the Exercise and Sport Science Laboratory at Lakeshore Foundation (Birmingham, AL), a community organization that provides physical activity, sport, and recreation opportunities to individuals with physical disability and chronic health conditions (trial registration: ClinicalTrials.gov NCT02994199). For the purposes of this study, participants generally visited the laboratory a total of 3 times within a 3-week period.

Participants

Eligibility criteria included age ≥18 years, a confirmed diagnosis of lower extremity-mobility limitation (eg, spina bifida, cerebral palsy, muscular dystrophy, 1 year after spinal cord injury, multiple sclerosis, stroke, or limb loss) with partial or full use of upper extremities and use of an assistive device (eg, cane, walker, or wheelchair) or problems with gait, balance, or coordination. Participants were excluded if they had an unstable cardiovascular condition, a visual impairment that interfered with playing video games, or weight over 350 lbs (159 kg) including their assistive device.

Procedures

Visit 1

During the first visit, informed consent/assent was obtained, and demographic and health history information was documented. An assessment of each participant’s functional ability was conducted as described below. In addition, participants were familiarized with the equipment (K4b2 portable metabolic system, COSMED, Rome, Italy) used for the study and the video games that would be played during subsequent visits. Participants played a portion of or the entire game for all games that were to be used during testing.

For the assessment of physical function during the first visit, each participant performed 18 functional movement tasks from the International Classification of Functioning, Disability and Health (ICF) [45,46]. Participants completed each task individually and were scored according to their difficulty in completing the task on a scale of 0 to 4. As defined in the ICF manual, the scoring was as follows: 0, “No difficulty”; 1, “Mild difficulty”; 2, “Moderate difficulty”; 3, “Severe difficulty”; and 4, “Complete difficulty.” The specific ICF tasks selected for use in this study were based on a consensus among the research staff about which mobility activities listed in the ICF would potentially be required for AVG play (eg, standing, reaching, throwing, and jumping) based on observations during the pilot testing. Scores on each of the 18 tasks were added together as a composite to represent participant physical function [9]. A lower composite score indicated greater functional ability on the selected tasks.

In addition to the functional assessment, participants also completed a series of questions from the HealthMeasures resources [47], which were used as an assessment of the individual’s own perspective regarding their functional ability. Questions were taken from the Patient Reported Outcomes Measurement Information System (PROMIS) [48]. The series comprised questions from PROMIS SFv1.0 Physical Function 20a and PROMIS SFv1.0 Physical Function Samples with Mobility Aid. Questions addressed issues of how difficult a variety of daily tasks (ie, vacuuming, yard work, walking, and bathing) were to complete (5 point scale, “without any difficulty” to “unable to do”; 14 questions), whether their health limited their ability to complete certain activities (ie, carry groceries, strenuous sports, walk a mile; 5 point scale, “not at all” to “cannot do”; 6 questions), and their ability to stand and move with and without support (“yes” or “no”; 1 question; 5 point scale, “without any difficulty” to “unable to do”, 10 questions).

Active Video Game Play

The second visit consisted of exercise testing during video game play. Upon arrival, participants were set up with the COSMED K4B2 portable metabolic system and a Polar (Kempele, Finland) heart rate monitor to assess pulmonary gas exchange and indirect calorimetry. Data collection began with a 20-minute rest period to measure the resting energy expenditure. For the rest period, participants sat quietly with no speaking or distractions besides light reading of a magazine or viewing their cellular phone. Next, game play began with continued gas exchange and heart rate data collection.

The Nintendo Wii video game console and a video game CD for the Wii Balance Board (Wii Fit Plus, Nintendo) were used for game play. Two game sets were created as outlined in Table 1. Game Set A included Rhythm Kung Fu, Rhythm Parade, Obstacle Course, and Bird’s Eye Bull’s Eye. Game Set B included Island Cycling, Penguin Slide, Hula Hoop, and Ski Slalom. Selected games were chosen in an effort to provide moderate-level physical activity during game play. Games represented popular genres (eg, fitness, sport, and adventure) and had an Entertainment Software Rating Board score of “E” (everyone).

Participants first played the two game sets on the OTS gaming board and then played the games on the adapted gaming board. The order of the game sets (Set A and Set B) played was randomly assigned to the participant. Each game set was played for 10 minutes with a rest period of 5 minutes afterward [9].
To assess the project aims, a series of analyses were conducted. To test whether energy expenditure during game play was different from the RPE (which could be readily taken for granted but was not assumed), resting METs were compared to game play METs for each subgroup. In general, 1 MET represents the amount of oxygen consumed and the number of calories burned at rest. In addition, for the subgroup StdStd, the change in METs (game play – rest) was compared between the OTS and adapted boards. The intensity of game play was defined as follows: <3 METs, light intensity; 3-6 METs, moderate intensity; >6 METs, vigorous intensity. Heart rate as a measure of exercise intensity was also recorded during all game play sessions. As a subjective rating of energy expenditure, RPE scores were analyzed for both OTS and adapted boards. To examine enjoyment, PACES scores were analyzed for both OTS and adapted boards for participants who were able to play both games; for participants who could only play using the adapted controller, scores were simply reported. Our primary comparison was for the StdStd group in which participants were able to play using both the OTS and adapted boards. For the aSit and aStd groups, where participants were unable to play using the OTS board, we chose to report only descriptive measures for the corresponding scores while playing the adapted boards. Furthermore, MET change (game - rest) comparisons between the OTS and adapted board in aSit and aStd groups were not computed as the change score for OTS would be equal to zero.

Data Analysis
Analyses were performed using the Statistical Package for the Social Sciences (IBM Corp, Chicago, IL). Due to the failure of groups to meet normality assumptions as well as a relatively small sample size for some comparisons, we reported the median and interquartile range for all measures and applied the Wilcoxon signed rank nonparametric test to detect significant differences between measures when comparisons were merited [50].

For each of these analyses, participants were divided into one of three unique groups based on their method of game play and level of access: Sitting (aSit), wherein individuals were seated for game play, were unable to access the OTS board, and played only on the adapted board; Standing, wherein individuals used the adapted board only (aStd), as they were able to stand, but due to mobility or balance issues could not access the OTS board, and therefore, they played only on the adapted board; and Standing with both boards (StdStd), wherein individuals were able to stand for game play on both the OTS and adapted boards.

To assess the project aims, a series of analyses were conducted. To test whether energy expenditure during game play was different from the RPE (which could be readily taken for granted but was not assumed), resting METs were compared to game play METs for each subgroup. In general, 1 MET represents the amount of oxygen consumed and the number of calories burned at rest. In addition, for the subgroup StdStd, the change in METs (game play – rest) was compared between the OTS and adapted boards. The intensity of game play was defined as follows: <3 METs, light intensity; 3-6 METs, moderate intensity; >6 METs, vigorous intensity. Heart rate as a measure of exercise intensity was also recorded during all game play sessions. As a subjective rating of energy expenditure, RPE scores were analyzed for both OTS and adapted boards. To examine enjoyment, PACES scores were analyzed for both OTS and adapted boards for participants who were able to play both games; for participants who could only play using the adapted controller, scores were simply reported. Our primary comparison was for the StdStd group in which participants were able to play using both the OTS and adapted boards. For the aSit and aStd groups, where participants were unable to play using the OTS board, we chose to report only descriptive measures for the corresponding scores while playing the adapted boards. Furthermore, MET change (game - rest) comparisons between the OTS and adapted board in aSit and aStd groups were not computed as the change score for OTS would be equal to zero.

Results
A total of 58 participants (aStd, n=10; aSit, n=31; StdStd, n=17) completed two 10-minute bouts of select Wii Fit Plus games on the OTS and adapted boards. The sample included 31 men and 27 women with a mean age of 41.21 (SD 12.70) years. Disabilities included spinal cord injury (n=11), multiple sclerosis (n=9), cerebral palsy (n=8), stroke (n=6), spina bifida (n=5), traumatic brain injury (n=4), limb loss (n=2), transverse myelitis (n=2), and others (n=11).

Heart rate and MET data are reported in Tables 2 and 3. For all conditions, the heart rate was significantly higher during game play than at rest (Table 2). Analysis of the metabolic data indicated that energy expenditure (METs) during game play...
was significantly greater than resting MET values for all players during both game sets on the adapted and OTS boards (Table 3).

Changes in energy expenditure (game play METs – rest METs) for each group are shown in Figure 2. When comparing MET change values between the OTS and adapted boards for the StdStd group, no significant differences ($P<.001$) were observed. The OTS board data are not reported for aSit and aStd, as these groups could not play on the OTS board and a MET change of zero was assumed for each participant.

As a subjective rating of exercise intensity, RPE values were recorded as shown in Figure 3. Median scores were displayed for each of the three groups. For the StdStd group, scores were compared between the two boards. For Set A, RPE did not differ significantly between game play on the OTS and adapted boards. For Set B, RPE values for the adapted board were significantly higher ($P<.05$) than those for the OTS board.

With regard to enjoyment, there were no significant differences in the PACES scores between the OTS and adapted boards (Figure 4). Median PACES scores for all conditions were 4 or “agree.”

Table 2. Resting and game play heart rate on the off-the-shelf and adapted boards.

<table>
<thead>
<tr>
<th>Play style groups and game set</th>
<th>Game HR, mean (SD)</th>
<th>Game HR, median (IQR)</th>
<th>Resting HR, mean (SD)</th>
<th>Resting HR, median (IQR)</th>
<th>Wilcoxon signed rank P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Off-the-shelf board</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aSit (n=27)</td>
<td>94.720 (19.783)</td>
<td>86.706 (81.270-106.294)</td>
<td>84.288 (12.278)</td>
<td>85.408 (74.985-90.446)</td>
<td>—</td>
</tr>
<tr>
<td>aStd (n=8)</td>
<td>109.959 (10.248)</td>
<td>105.866 (102.975-117.275)</td>
<td>83.659 (10.861)</td>
<td>86.999 (74.539-90.791)</td>
<td>—</td>
</tr>
<tr>
<td>StdStd (n=14)</td>
<td>96.677 (0.568)</td>
<td>92.436 (78.966-109.271)</td>
<td>73.118 (11.161)</td>
<td>71.247 (67.099-79.265)</td>
<td>.001</td>
</tr>
<tr>
<td><strong>Adapted board</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aSit (n=27)</td>
<td>95.844 (15.423)</td>
<td>95.108 (84.836-103.678)</td>
<td>84.288 (12.278)</td>
<td>85.408 (74.985-90.446)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>aStd (n=8)</td>
<td>109.547 (102.975-117.275)</td>
<td>105.866 (102.975-117.275)</td>
<td>83.659 (10.861)</td>
<td>86.999 (74.539-90.791)</td>
<td>.01</td>
</tr>
<tr>
<td>StdStd (n=14)</td>
<td>97.074 (21.027)</td>
<td>93.499 (81.296-108.745)</td>
<td>73.118 (11.161)</td>
<td>71.247 (67.099-79.265)</td>
<td>.001</td>
</tr>
</tbody>
</table>

*aHR: heart rate.
*bIQR: interquartile range.
caSit: participants who only played seated on the adapted board.
dNot available.
eU: Unable to utilize off-the-shelf board for game play.
fN/A: not applicable.
gaStd: participants who only played standing on the adapted board.
hStdStd: participants who were able to play standing on both boards.
Table 3. Resting and game play energy expenditure on the off-the-shelf and adapted boards.

<table>
<thead>
<tr>
<th>Play style groups and game set</th>
<th>Game METs, mean (SD)</th>
<th>Game METs, median (IQR)</th>
<th>Resting METs, mean (SD)</th>
<th>Resting METs, median (IQR)</th>
<th>Wilcoxon signed rank P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Off-the-shelf board</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aSit</td>
<td>1.157 (0.375)</td>
<td>1.123 (0.916-1.307)</td>
<td>—</td>
<td>—</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Set A (n=31)</td>
<td>U</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Set B (n=31)</td>
<td>U</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>aStd</td>
<td>1.063 (0.346)</td>
<td>0.902 (0.812-1.269)</td>
<td>—</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>Set A (n=10)</td>
<td>—</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Set B (n=10)</td>
<td>—</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>StdStd</td>
<td>1.014 (0.235)</td>
<td>1.032 (0.781-1.164)</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set A (n=17)</td>
<td>2.839 (0.646)</td>
<td>2.777 (2.623-3.048)</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Set B (n=17)</td>
<td>2.702 (0.568)</td>
<td>2.707 (2.265-2.979)</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td><strong>Adapted board</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aSit</td>
<td>1.157 (0.375)</td>
<td>1.123 (0.916-1.307)</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set A (n=31)</td>
<td>2.261 (0.718)</td>
<td>2.143 (1.847-2.701)</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Set B (n=31)</td>
<td>2.233 (0.751)</td>
<td>2.203 (1.762-2.567)</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>aStd</td>
<td>1.063 (0.346)</td>
<td>0.902 (0.812-1.269)</td>
<td>.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set A (n=10)</td>
<td>3.151 (1.034)</td>
<td>2.827 (2.482-3.605)</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Set B (n=10)</td>
<td>2.990 (1.121)</td>
<td>3.028 (2.204-3.675)</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>StdStd</td>
<td>1.014 (0.235)</td>
<td>1.032 (0.781-1.164)</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set A (n=17)</td>
<td>2.732 (0.655)</td>
<td>2.714 (2.443-2.974)</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Set B (n=17)</td>
<td>2.777 (0.803)</td>
<td>2.572 (2.253-3.136)</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

aMETs: metabolic equivalents.
bIQR: interquartile range.
caSit: participants who only played seated on the adapted board.
dNot available.
eU: Unable to utilize off-the-shelf board for game play.
fN/A: not applicable.
gaStd: participants who only played standing on the adapted board.
hStdStd: participants who were able to play standing on both boards.
Figure 2. Change in energy expenditure (metabolic equivalents [METs], gameplay – rest) for each group with comparison of MET change between boards (adapted vs off-the-shelf [OTS]) for the StdStd group. Play style groups: aSit - participants who only played seated on the adapted board; aStd - participants who only played standing on the adapted board; StdStd - participants who were able to play standing on both boards. MET-change data for the OTS board are not reported for aSit and aStd groups, as these groups could not play on the OTS board. Computing P values for these comparisons with the assumption of OTS=0 would yield highly statistically significant differences (<.01).
Figure 3. Rating of perceived exertion (RPE) following game play for each group with comparison of RPE between boards (adapted vs off-the-shelf [OTS]) for the StdStd group. Play style groups: aSit - participants who only played seated on the adapted board; aStd - participants who only played standing on the adapted board; StdStd - participants who were able to play standing on both boards.

Game Set A

<table>
<thead>
<tr>
<th></th>
<th>aSit</th>
<th>aStd</th>
<th>StdStd</th>
<th>StdStd</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>![Boxplot for Game Set A aSit]</td>
<td>![Boxplot for Game Set A aStd]</td>
<td>![Boxplot for Game Set A StdStd]</td>
<td>![Boxplot for Game Set A StdStd]</td>
</tr>
<tr>
<td></td>
<td>![Adapted]</td>
<td>![Adapted]</td>
<td>![Adapted]</td>
<td>![OTS]</td>
</tr>
<tr>
<td></td>
<td>![Adapted]</td>
<td>![Adapted]</td>
<td>![Adapted]</td>
<td>![OTS]</td>
</tr>
</tbody>
</table>

P = .59

Game Set B

<table>
<thead>
<tr>
<th></th>
<th>aSit</th>
<th>aStd</th>
<th>StdStd</th>
<th>StdStd</th>
</tr>
</thead>
<tbody>
<tr>
<td>RPE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>![Boxplot for Game Set B aSit]</td>
<td>![Boxplot for Game Set B aStd]</td>
<td>![Boxplot for Game Set B StdStd]</td>
<td>![Boxplot for Game Set B StdStd]</td>
</tr>
<tr>
<td></td>
<td>![Adapted]</td>
<td>![Adapted]</td>
<td>![Adapted]</td>
<td>![OTS]</td>
</tr>
<tr>
<td></td>
<td>![Adapted]</td>
<td>![Adapted]</td>
<td>![Adapted]</td>
<td>![OTS]</td>
</tr>
</tbody>
</table>

P = .05

Legend: Adapted = light blue, OTS = dark red.
**Discussion**

**Overview**

Development of AVGs that are accessible to people with disabilities offers an innovative approach to overcoming a number of barriers to participation in physical activity. To date, little effort has been made to design commercially available AVG controllers in a way to promote active play in persons who require a wheelchair for mobility. Although several studies have documented increases in energy expenditure during AVG by individuals with mobility impairments who played standing [30,31,35,36], limited evidence is available regarding outcomes associated with AVG play while sitting. A feasibility study conducted by Rowland and Rimmer [12] modified the parameters for Dance, Dance Revolution game play to allow for arm use by placing the mat on a table and found increases in energy expenditure among three nonambulatory young adults with cerebral palsy. Mat Rosly et al [51] examined game play MET and RPE levels by persons with spinal cord injury utilizing the PlayStation 3 console with two PlayStation Move hand-held controllers. Other attempts have been made to supplement existing exercise equipment with gaming for various populations (ie, spinal cord injury and cerebral palsy) with mobility impairments. Adding game play to wheelchair ergometry makes exercise more enjoyable [52-54], similar to the case with leg cycling [55], while also achieving recommended exercise intensity levels for health-related outcomes.

Recognizing the potential for AVG play to increase energy expenditure among people with more severe mobility impairments led our team to develop an adapted balance board gaming controller [44]. Thus, the objective of the current study was to examine energy expenditure and enjoyment during AVG in persons with mobility impairments when using both OTS and adapted versions of the Wii Fit balance board.

Resting METs were compared to game play METs for each subgroup (based on the degree of mobility impairment) to determine if players actually experienced an increase in energy expenditure during AVG play above the resting values. Given

---

**Figure 4.** Assessment of game play enjoyment based on Physical Activity Enjoyment Scale (PACES) scores. Play style groups: aSit - participants who only played seated on the adapted board; aStd - participants who only played standing on the adapted board; StdStd - participants who were able to play standing on both boards. PACES data for the off-the-shelf (OTS) board are not reported for aSit and aStd, as these groups could not play on the OTS board.
the severity of mobility impairment of some participants, this could not be assumed. The significant differences between game play and resting MET values for all players suggest that AVG play using the adapted controller may provide a leisure-time physical activity option that reduces sedentary time for individuals with a variety of mobility limitations. The increased usability of the gaming board opens the door to AVG play for many people who were previously unable to participate \[10, 44\].

For players who were unable to utilize the OTS board, change in the METs from rest to game play equaled to zero (ie, METs remained at resting values). However, on the adapted board, participants who played while sitting were able to achieve average MET values of 2.261 (SD 0.718) and 2.233 (SD 0.751) kcal/kg/hour on the two game sets each. Average MET values for those who played standing were a bit higher (3.151 [SD 1.034] and 2.990 [SD 1.121] kcal/kg/hour), as expected, given the ability to engage greater muscle mass (ie, lower extremity). Although these average values represent light-to-moderate-intensity exercise, 16% of the seated participants and 41% of the standing participants achieved moderate-intensity exercise on at least one of the games. However, these standard categories based on MET values do not take into account the effect of an individual’s impairment level on intensity of exercise.

In some cases, a self-reported rating of perceived exertion (RPE) on a 0–10 scale suggested greater exercise intensity levels than MET levels recorded, which is similar to the results of an exergaming study in persons with spinal cord injury \[51\]. Seated players reported a median RPE of 4 (moderate) for both game sets. Players who could only play standing on the adapted board reported somewhat hard-to-moderately hard RPE levels (5-6), while players who played standing on both boards reported moderate-to-somewhat hard RPE levels (4-5).

In a study on healthy young adults, in which participants played various Wii games, exercise intensity varied by the game, ranging from light to moderate \[24\]. The authors suggested that for games that require controller skill, exercise intensity may be influenced by the player’s prior gaming experience. Furthermore, the benefits of light-intensity exercise are acknowledged even if moderate levels are not reached, which was the case for some participants in our study. As noted in a recent study, replacement of sedentary time with light-intensity physical activity is associated with less mortality in the general population of adults aged ≥40 years \[56\], with beneficial effects on health outcomes such as blood glucose levels \[57\].

In addition to the potential for increased energy expenditure, many people perceive AVGs as fun, providing an enjoyable option to perform recommended daily amounts of physical activity. In this study, the PACES scores indicated that all players, sitting and standing, enjoyed playing each of the game sets, with a median score of 4 on a 5-point scale. In a study comparing heavy-bag boxing to AVG boxing in a seated position among persons with spinal cord injury, participants reported more enjoyment during AVG boxing \[38\]. In another study, game play performance and exercise intensity were positively correlated with AVG enjoyment in youth with mobility impairments \[58\]. Enjoyment may serve as a determinant of physical activity, further suggesting the need to develop AVG interventions and examine the role of enjoyment on the level of engagement, exercise intensity, and adherence.

**Limitations**

This study was not a randomized controlled trial; therefore, no claims can be made regarding causality or efficacy. As an observational study, inherent limitations existed and thereby limit the generalizability of results to the broader community. All participants were recruited from the membership of a community physical activity and recreation center for individuals with physical disabilities. Individuals were, to some degree, physically active with varied AVG experience. Factors that may have influenced intensity of exercise but were not accounted for include game selection, limited familiarization period, and discomfort wearing the COSMED system for oxygen consumption measurement. Although a familiarization period was provided, some degree of game play learning may have occurred during data collection. In addition, given that participants played only a select group of AVGs, potential differences in enjoyment and energy expenditure between OTS and adapted controllers may not have been fully captured. Moreover, the standard categories of exercise intensity based on MET values do not take into consideration the effect of impairment level on exercise intensity. Furthermore, the intrinsic nature of measures examining subjective aspects of exercise such as enjoyment and perceived exertion prevented comparison of these aspects across board type for the participants who were unable to utilize the OTS board. Future studies should expand the participant recruitment pool, examine a broader range of AVGs, provide a more extensive familiarization period, and compare AVG play while utilizing the adapted controllers for other leisure-time physical activities. Further analyses that explore the multivariate correlation and variance-covariance matrices in the outcome measures (RPE, heart rate, MET, PACES) would provide a better understanding of the relationship between exercise intensity and enjoyment and of clinically meaningful differences during AVG in this subpopulation.

**Conclusions**

Engagement in physical activity is an important component of a healthy lifestyle and reduces the risk for many serious health conditions such as cancer, heart disease, stroke, and obesity \[59\]. Individuals with physical disabilities are at an even greater risk for many of these conditions. With limited options for physical activity, finding alternative means in order to stay active is critical for achieving optimal health. The adapted board improved access and allowed participants of all mobility levels to engage in AVG play, removing barriers associated with the OTS board. The adapted game controller provided AVG game play options for individuals unable to stand during play. Players were able to achieve MET values above the resting level, thereby reducing sedentary time. Furthermore, light-to-moderate intensity exercise levels were reached by some participants, providing an enjoyable option for engagement in health-related physical activity.
Acknowledgments

The content of this article was developed under a Rehabilitation Engineering Research Center on Interactive Exercise Technologies and Exercise Physiology for People with Disabilities RecTech grant from the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR, grant numbers H133E120005 and 90REGE0002-01-00). The NIDILRR is a center within the Administration for Community Living (ACL), Department of Health and Human Services (HHS). The contents of this article do not necessarily represent the policy of the NIDILRR, ACL, and HHS, and endorsement by the Federal Government should not be assumed. We thank Justin McCroskey, Brandon Kane, and Lieu Thompson, who were part of the research team, for their assistance with this project. We also express our gratitude to Samuel R Misko, Brandon Kirkland, and Jud Dunlap, as employees of the UAB Engineering and Innovative Technology Development group, for design and development of the adapted board.

Conflicts of Interest

None declared.

References


47. HealthMeasures: Transforming How Health is Measured.: Northwestern University URL: http://www.healthmeasures.net/ [accessed 2019-01-09] [WebCite Cache ID 75JHa0AHM]
58. Sainte-Maurice PF, Troiano RP, Berrigan D, Kraus WE, Matthews CE. Volume of light versus moderate-to-vigorous physical activity: Similar benefits for all-cause mortality? J Am Heart Assoc 2018 Apr 02;7(7) [FREE Full text] [doi: 10.1161/JAHA.118.010885] [Medline: 29610219]
IQR: interquartile range
MET: metabolic equivalent
OTS: off-the-shelf
PACES: Physical Activity Enjoyment Scale
PROMIS: Patient Reported Outcomes Measurement Information System
RPE: rating of perceived exertion
StdStd: participants who were able to play standing on both boards
SUS: System Usability Scale
Effectiveness of a Behavior Change Technique–Based Smartphone Game to Improve Intrinsic Motivation and Physical Activity Adherence in Patients With Type 2 Diabetes: Randomized Controlled Trial

Christoph Höchsmann¹,², PhD; Denis Infanger¹, PhD; Christopher Klenk¹, MD; Karsten Königstein¹, MD; Steffen P Walz³, PhD; Arno Schmidt-Trucksäss¹, MD

¹Department of Sport, Exercise and Health, University of Basel, Basel, Switzerland
²Pennington Biomedical Research Center, Baton Rouge, LA, United States
³Centre for Design Innovation, Swinburne University of Technology, Melbourne, Australia

Corresponding Author:
Arno Schmidt-Trucksäss, MD
Department of Sport, Exercise and Health
University of Basel
Birsstrasse 320 B
Basel, 4052
Switzerland
Phone: 41 612074740
Email: arno.schmidt-trucksaess@unibas.ch

Abstract

Background: Regular physical activity (PA) is an essential component of a successful type 2 diabetes treatment. However, despite the manifest evidence for the numerous health benefits of regular PA, most patients with type 2 diabetes remain inactive, often due to low motivation and lack of PA enjoyment. A recent and promising approach to help overcome these PA barriers and motivate inactive individuals to change their PA behavior is PA-promoting smartphone games. While short-term results of these games are encouraging, the long-term success in effectively changing PA behavior has to date not been confirmed. It is possible that an insufficient incorporation of motivational elements or flaws in gameplay and storyline in these games affect the long-term motivation to play and thereby prevent sustained changes in PA behavior. We aimed to address these design challenges by developing a PA-promoting smartphone game that incorporates established behavior change techniques and specifically targets inactive type 2 diabetes patients.

Objective: To investigate if a self-developed, behavior change technique-based smartphone game designed by an interdisciplinary team is able to motivate inactive individuals with type 2 diabetes for regular use and thereby increase their intrinsic PA motivation.

Methods: Thirty-six inactive, overweight type 2 diabetes patients (45-70 years of age) were randomly assigned to either the intervention group or the control group (one-time lifestyle counseling). Participants were instructed to play the smartphone game or to implement the recommendations from the lifestyle counseling autonomously during the 24-week intervention period. Intrinsic PA motivation was assessed with an abridged 12-item version of the Intrinsic Motivation Inventory (IMI) before and after the intervention. In addition, adherence to the game-proposed PA recommendations during the intervention period was assessed in the intervention group via the phone-recorded game usage data.

Results: Intrinsic PA motivation (IMI total score) increased significantly in the intervention group (+6.4 (SD 4.2; P<.001) points) while it decreased by 1.9 (SD 16.5; P=.623) points in the control group. The adjusted difference between both groups was 8.1 (95% CI 0.9, 15.4; P=.029) points. The subscales “interest/enjoyment” (+2.0 (SD 1.9) points, P<.001) and “perceived competence” (+2.4 (SD 2.4) points, P=.001) likewise increased significantly in the intervention group while they did not change significantly in the control group. The usage data revealed that participants in the intervention group used the game for an average of 131.1 (SD 48.7) minutes of in-game walking and for an average of 15.3 (SD 24.6) minutes of strength training per week. We found a significant positive association between total in-game training (min) and change in IMI total score (beta=0.0028; 95% CI 0.0007-0.0049; P=.01).
Conclusions: In inactive individuals with type 2 diabetes, a novel smartphone game incorporating established motivational elements and personalized PA recommendations elicits significant increases in intrinsic PA motivation that are accompanied by de-facto improvements in PA adherence over 24 weeks.

Trial Registration: ClinicalTrials.gov NCT02657018; https://clinicaltrials.gov/ct2/show/NCT02657018

KEYWORDS
behavior change; exercise adherence; gamification; intrinsic motivation; mhealth; mobile phone game; physical activity; type 2 diabetes

Introduction
Diabetes mellitus affects over 400 million adults worldwide [1], and 90%-95% of all cases are attributed to type 2 diabetes [2]. The disease is associated with numerous complications and comorbidities that drastically increase direct and indirect medical costs and considerably contribute to the disease’s enormous financial strain worldwide [3].

Regular physical activity (PA), with its proven positive effects on glucose and lipid metabolism, blood pressure, cardiovascular complications, and quality of life, is a key component of successful diabetes treatment [4]. Despite these health benefits and the fact that patients with diabetes are usually encouraged to increase PA by their physicians, long-term adherence to PA-promoting programs is generally poor, and the level of regular PA remains low [2,5]. PA has been estimated to be responsible for 6%-10% of the world’s type 2 diabetes prevalence [6] and increase the risk of all-cause mortality by an estimated 60% [7]. Lack of infrastructure, missing social support, health concerns, and especially low motivation and lack of PA enjoyment are the main deterrents that keep patients with type 2 diabetes from effectively changing their PA behaviors [8,9]. As intrinsically motivated individuals have been shown to have higher PA engagement and better PA adherence than those who are primarily motivated by external factors [10], PA-promoting programs should aim at particularly helping patients increase their PA enjoyment, and consequently, intrinsic motivation for regular PA [11].

A promising approach that has increasingly been examined in recent years to promote regular PA in unmotivated, inactive target groups is exergames. Through the enjoyable game experience, console-based exergames have been shown to motivate inactive patients with type 2 diabetes to voluntarily engage in more regular PA and thereby improve their glycemic control and overall health status [12]. Very recently, PA-promoting game apps such as Pokémon GO (Niantic Labs, San Francisco, CA) have entered the market and likewise aim at sustaining PA habits through gamified incentives. Although Pokémon GO certainly has the potential to increase daily PA, it should be noted that the game-related initial increases of daily PA of 25% in the first week have been shown to gradually diminish in subsequent weeks and return to baseline after only 6 weeks [13]. It is possible that the game design does not include a sufficient degree of narrative, gameplay, or storytelling, which are required to sustainably motivate the player to play the game and consequently make a PA-promoting game effective in the long term [14]. To address these design challenges, with an interdisciplinary team featuring sports scientists, gamification researchers, professional game developers, and clinical professionals, we developed a novel smartphone game that incorporates established motivational elements [15] and behavior change techniques [16] to encourage inactive patients with type 2 diabetes to adopt a healthier, more active lifestyle.

The purpose of this study was to investigate if the behavior change technique–based smartphone game can motivate inactive individuals with type 2 diabetes for regular use and thereby increase their intrinsic PA motivation. We hypothesized that use of the game would lead to greater improvements in intrinsic PA motivation than a control intervention consisting of a one-time lifestyle counseling and thereby increase PA adherence.

Methods
Study Design
This 24-week randomized controlled trial was conducted in accordance with the Declaration of Helsinki [17] between August 2016 and April 2018 at the Department of Sport, Exercise and Health of the University of Basel, Switzerland (trial registration: NCT02657018) and was approved by the local ethics committee (EKNZ 2015-424). Written informed consent was obtained from all study participants prior to inclusion in the study. The primary aim of this study was to investigate the effect of a novel PA-promoting smartphone game on daily PA in inactive patients with type 2 diabetes, measured as steps per day with the previously validated Garmin Vivofit 2 activity wristband [18] for 1 week before and after the intervention period. A significant increase in daily PA was found in the intervention group with average postintervention step counts corresponding to established PA recommendations [19]. The increases in daily PA were accompanied by significant improvements in aerobic capacity and stabilization of the glycemic control, measured as hemoglobin A1c (C. Höchsmann, personal communication, June 2018). Additional aims were to assess the effect of the game on the predefined [20] further outcomes “intrinsic PA motivation” and “PA adherence.” Participants were allocated at random to either the intervention or the control group using R version 3.2.3 (R Foundation for Statistical Computing, Vienna, Austria) and the R add-on package “blockrand” version 1.3 to apply permuted block randomization with randomly varying block sizes. The randomization list was generated in advance and transmitted by a person not involved in the study by using serially numbered, sealed, opaque envelopes. All outcome assessors were blinded with respect to the group allocation. It has been shown that after
the initial 6 weeks, adherence to an exercise program in patients with type 2 diabetes decreases steadily, with a dropout of almost 80% after 6 months but is relatively stable beyond the 6-month timeframe [21]. Therefore, we chose an intervention duration of 24 weeks, which matches this critical 6-month timeframe and would consequently allow assessment of the longer-term PA adherence. The detailed study protocol can be found in a previous publication [20].

Recruitment
Physically inactive overweight (body mass index ≥ 25 kg/m²) patients with type 2 diabetes (noninsulin-dependent) between the ages of 45 and 70 years were recruited in cooperation with various hospitals, doctor’s offices, and diabetes care centers in the Basel metropolitan area and via online and newspaper advertising. Eligible participants were required to have used a smartphone regularly during the year before the study to ensure that they were technologically proficient enough to represent a realistic target audience for a PA-promoting smartphone game. During the eligibility screening, participants underwent a medical examination including height, body mass, body fat content (by bioelectrical impedance analysis), and resting blood pressure measurements as well as resting electrocardiography. To verify insufficient levels of PA before the study (<150 minutes of moderate-intensity PA per week), participants filled out the Freiburg Questionnaire of PA [22] as part of the eligibility screening. Present health risks that contraindicate exercise testing [23] as well as an impaired physical mobility, acute infections, or injuries excluded participants from enrolling and continuing their participation in the study.

Intervention
After the baseline assessment, the self-developed and commercially released smartphone game “Mission: Schweinehund” was installed on the phones of participants in the intervention group. The game was designed to be self-explanatory and motivate for regular use through its game mechanics. Therefore, participants only received basic instructions regarding the game controls, which were not related to an intended frequency or duration of use during the intervention period. The game uses the self-determination theory as the theoretical framework [24]. The self-determination theory is a widely researched theory of motivation that addresses both intrinsic and extrinsic motives for acting and has shown its utility in explaining processes that underpin exercise behavior [25] as well as motivation to play video games [26]. The goal of the game is to restore a decayed garden by planting trees and flowers. In doing so, the player attracts animals that used to live in the garden to come back and help with the restoration process. At the same time, the Schweinehund, the game’s adversary, is kept in check. In German, “innerer Schweinehund” (inner swine hound) refers to the weak or lazy part of one’s nature, often regarding PA, that has to be overcome to get one’s self going. The garden setting was deliberately chosen, as its restoration stands metaphorically for the restoration of the player’s body through regular PA. In addition, it has been shown that gardening is among the target group’s preferred forms of PA [27] and that gardening-themed games are quite popular and comprehensible to a wide range of players because of their straightforward mechanics [28,29]. In the game, regular PA is rewarded with water or building materials that are needed to restore the garden and proceed in the storyline. When designing the game, close attention was paid to the inclusion of established motivational elements [15] and behavior change techniques [16]. In addition to the rewards for successful PA behavior, goal setting, action planning, feedback on performance, and prompts and cues were incorporated into the game mechanics to support sustained changes in PA behavior. During the development phase of the game, all design choices, behavioral/motivational elements, and game mechanics were tested extensively with the target group regarding usability and motivational efficacy. Overall, four user studies (unpublished data) with a total of 44 participants and various thematic foci (e.g., motivational efficacy of game concept and storyline, suitability of in-game workout regimen and baseline tests, usability of sensor tracking, and suitability of motivational elements) were conducted during the 26-month development phase.

The game’s PA content includes in-game workouts as well as the promotion of daily PA that follow the American College of Sports Medicine and European Association for Cardiovascular Prevention & Rehabilitation principles of exercise training [30,31]. In-game workouts consist of 130 variations of strength, endurance, balance, and flexibility exercises whose execution, as well as daily PA, is tracked via the phone’s sensors (camera, accelerometer, and gyroscope). To allow individualization of the PA-related content of the game, exercise tests such as the 1-minute Sit-to-Stand Test [32] and the 6-Minute Walk Test [33] assess the player’s fitness level at baseline and periodically during play. Based on the results, an algorithm selects appropriate entry levels and tailored rates of intensity progression for the exercise regimens and the daily PA goals, which could also be manually adjusted by the player to fit personal preferences and potential physical limitations. An individualization of the exercise regimen is crucial because unrealistic, overwhelming targets often reduce patients’ motivation and thereby directly affect adherence [34]. Regularity of PA and relative improvements rather than high absolute values determine the progression in the game and thereby make game success independent of the individual fitness level. Participants in the control group received a one-time lifestyle counseling to promote baseline activities of daily life [35]. Further, control group participants were provided with a structured exercise plan consisting of strength and endurance exercises with moderately increasing intensity and duration, comparable to the content of the game that was to be implemented autonomously during the intervention period. A detailed description of the game including screenshots can be found in a previous publication [20].

Outcome Measures

Intrinsic Physical Activity Motivation
Intrinsic PA motivation was measured using an abridged 12-item version of the Intrinsic Motivation Inventory (IMI) at baseline and after the 24-week intervention. The IMI has gained widespread acceptance as a multidimensional measure of intrinsic motivation in the context of sports and physical activity [36-41]. The questionnaire included four subscales:
“interest/enjoyment,” “perceived competence,” “perceived choice,” and “value/usefulness.” These subscales have been previously used to assess participants’ subjective experience regarding participation in a television exercise program [36] and to examine the motivational pull of video games [26]. The “interest/enjoyment” subscale is considered the true self-report measure of intrinsic motivation, whereas “perceived competence,” “perceived choice,” and “value/usefulness” are viewed as positive predictors of intrinsic motivation [36]. The items of each subscale were modified to fit the content of this study and rated on a 7-point Likert scale ranging from 1 (not at all true) to 7 (very true). This yielded total scores between 3 and 21 for each subscale and between 12 and 84 for the entire questionnaire. Higher scores indicated more internally motivated, self-regulated PA behavior.

Physical Activity Adherence

In the intervention group, adherence to the game-proposed PA recommendations was assessed during the intervention period via the recorded usage data from the participants’ phones. Usage data included daily PA (steps per day), completed and canceled in-game workouts, and patterns and total duration of game use. The accuracy of iPhones and Android phones to detect steps during various walking conditions independent of the placement on the body has been confirmed in our previous study [18]. For in-game walking, stride cadence was measured to assess periods of moderate-to-vigorous-intensity walking defined as ≥100 steps/minute [19].

Statistical Analysis

Summary statistics were calculated to characterize the study sample and for pre- and postintervention data, as appropriate. Continuous data were summarized using mean (SD) and median (interquartile range [IQR]). Intrinsic PA motivation (IMI total score and scores for all four subscales) after the intervention was analyzed by analysis of covariance [42]. Results are presented as differences in outcome (with 95% CI) between participants in the intervention group and those in the control group, adjusted for the corresponding values at baseline. PA adherence was analyzed descriptively using the median, IQR, and range to illustrate game-related and overall-recorded daily PA per week during the 24-week intervention period. A linear regression model was used to assess the relationship between total in-game training (minutes) and change in IMI total score. Assumptions of the analysis of covariance were checked visually and 21 for each subscale and between 12 and 84 for the entire questionnaire. Higher scores indicated more internally motivated, self-regulated PA behavior.

Changes in Intrinsic Physical Activity Motivation

Figure 2 shows the pre- and postintervention data as mean and IQR for the IMI total score and all four subscales. Intrinsic PA motivation (IMI total score) increased significantly by an average of 6.39 (SD 4.19; \( P<.001 \)) points in the intervention group and decreased by an average of 1.94 (SD 16.46; \( P=.62 \)) points in the control group with an adjusted difference of 8.15 points (95% CI 0.90-15.39; \( P=.03 \)) between the two groups. Similarly, we observed significant increases in scores for the subscales “interest/enjoyment” (by 2.00 [SD 1.94] points, \( P<.001 \)) and “perceived competence” (by 2.44 [SD 2.36] points, \( P<.001 \)) in the intervention group but no significant change in the control group. The adjusted difference between the two groups was 2.03 points (95% CI 0.04-4.09; \( P=.049 \)) for “interest/enjoyment” and 2.88 points (95% CI 0.59-5.17; \( P=.02 \)) for “perceived competence.” The value/usefulness subscale showed a significant adjusted difference of 2.72 points (95% CI 0.28-5.16; \( P=.03 \)) in favor of the intervention group despite nonsignificant changes in either group. The score for the “perceived choice” subscale increased significantly by 1.22 (SD 2.44) points in the intervention group, but with a nonsignificant adjusted difference of 0.67 points (95% CI 1.27 to 2.61; \( P=.91 \)) between the two groups.

Sample Size

An \( a \) priori sample size calculation based on the primary outcome was conducted for this study. Since intrinsic PA motivation and PA adherence are secondary outcomes, the corresponding analyses presented in this report should be considered explorative, and we reported the evidence in the data for the hypothesized effect of the game use on intrinsic PA motivation and PA adherence.

Results

Participant Flow and Characteristics

Figure 1 illustrates the participants’ flow through the study. A total of 68 subjects were assessed for eligibility, of which 19 did not meet the inclusion criteria and 13 declined to participate, leading to the exclusion of 32 subjects. All the remaining participants (\( n=36 \)) were randomly assigned to either the intervention group (\( n=18 \)) or the control group (\( n=18 \) (Figure 1). Baseline characteristics of study participants were balanced between the two groups (Table 1). One participant dropped out of the study before follow-up due to medical reasons not related to the study. No study-related or other adverse events were reported during the intervention period. Further, 35 participants completed the study and were included in the analyses.

Sample Size

An \( a \) priori sample size calculation based on the primary outcome was conducted for this study. Since intrinsic PA motivation and PA adherence are secondary outcomes, the corresponding analyses presented in this report should be considered explorative, and we reported the evidence in the data for the hypothesized effect of the game use on intrinsic PA motivation and PA adherence.

Results

Participant Flow and Characteristics

Figure 1 illustrates the participants’ flow through the study. A total of 68 subjects were assessed for eligibility, of which 19 did not meet the inclusion criteria and 13 declined to participate, leading to the exclusion of 32 subjects. All the remaining participants (\( n=36 \)) were randomly assigned to either the intervention group (\( n=18 \)) or the control group (\( n=18 \) (Figure 1). Baseline characteristics of study participants were balanced between the two groups (Table 1). One participant dropped out of the study before follow-up due to medical reasons not related to the study. No study-related or other adverse events were reported during the intervention period. Further, 35 participants completed the study and were included in the analyses.

Changes in Intrinsic Physical Activity Motivation

Figure 2 shows the pre- and postintervention data as mean and IQR for the IMI total score and all four subscales. Intrinsic PA motivation (IMI total score) increased significantly by an average of 6.39 (SD 4.19; \( P<.001 \)) points in the intervention group and decreased by an average of 1.94 (SD 16.46; \( P=.62 \)) points in the control group with an adjusted difference of 8.15 points (95% CI 0.90-15.39; \( P=.03 \)) between the two groups. Similarly, we observed significant increases in scores for the subscales “interest/enjoyment” (by 2.00 [SD 1.94] points, \( P<.001 \)) and “perceived competence” (by 2.44 [SD 2.36] points, \( P<.001 \)) in the intervention group but no significant change in the control group. The adjusted difference between the two groups was 2.03 points (95% CI 0.04-4.09; \( P=.049 \)) for “interest/enjoyment” and 2.88 points (95% CI 0.59-5.17; \( P=.02 \)) for “perceived competence.” The value/usefulness subscale showed a significant adjusted difference of 2.72 points (95% CI 0.28-5.16; \( P=.03 \)) in favor of the intervention group despite nonsignificant changes in either group. The score for the “perceived choice” subscale increased significantly by 1.22 (SD 2.44) points in the intervention group, but with a nonsignificant adjusted difference of 0.67 points (95% CI 1.27 to 2.61; \( P=.91 \)) between the two groups.

Sample Size

An \( a \) priori sample size calculation based on the primary outcome was conducted for this study. Since intrinsic PA motivation and PA adherence are secondary outcomes, the corresponding analyses presented in this report should be considered explorative, and we reported the evidence in the data for the hypothesized effect of the game use on intrinsic PA motivation and PA adherence.

Results

Participant Flow and Characteristics

Figure 1 illustrates the participants’ flow through the study. A total of 68 subjects were assessed for eligibility, of which 19 did not meet the inclusion criteria and 13 declined to participate, leading to the exclusion of 32 subjects. All the remaining participants (\( n=36 \)) were randomly assigned to either the intervention group (\( n=18 \)) or the control group (\( n=18 \) (Figure 1). Baseline characteristics of study participants were balanced between the two groups (Table 1). One participant dropped out of the study before follow-up due to medical reasons not related to the study. No study-related or other adverse events were reported during the intervention period. Further, 35 participants completed the study and were included in the analyses.

Changes in Intrinsic Physical Activity Motivation

Figure 2 shows the pre- and postintervention data as mean and IQR for the IMI total score and all four subscales. Intrinsic PA motivation (IMI total score) increased significantly by an average of 6.39 (SD 4.19; \( P<.001 \)) points in the intervention group and decreased by an average of 1.94 (SD 16.46; \( P=.62 \)) points in the control group with an adjusted difference of 8.15 points (95% CI 0.90-15.39; \( P=.03 \)) between the two groups. Similarly, we observed significant increases in scores for the subscales “interest/enjoyment” (by 2.00 [SD 1.94] points, \( P<.001 \)) and “perceived competence” (by 2.44 [SD 2.36] points, \( P<.001 \)) in the intervention group but no significant change in the control group. The adjusted difference between the two groups was 2.03 points (95% CI 0.04-4.09; \( P=.049 \)) for “interest/enjoyment” and 2.88 points (95% CI 0.59-5.17; \( P=.02 \)) for “perceived competence.” The value/usefulness subscale showed a significant adjusted difference of 2.72 points (95% CI 0.28-5.16; \( P=.03 \)) in favor of the intervention group despite nonsignificant changes in either group. The score for the “perceived choice” subscale increased significantly by 1.22 (SD 2.44) points in the intervention group, but with a nonsignificant adjusted difference of 0.67 points (95% CI 1.27 to 2.61; \( P=.91 \)) between the two groups.
Figure 1. Flow diagram of study participants.

Table 1. Baseline characteristics of study participants.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Intervention (n=18)</th>
<th>Control (n=18)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>8</td>
<td>___</td>
</tr>
<tr>
<td>Male</td>
<td>10</td>
<td>___</td>
</tr>
<tr>
<td>Age (years)</td>
<td>18</td>
<td>56 (5)</td>
</tr>
<tr>
<td>Height (m)</td>
<td>18</td>
<td>171 (7)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>18</td>
<td>93 (12)</td>
</tr>
<tr>
<td>Body mass index (kg/m^2)</td>
<td>18</td>
<td>32 (4)</td>
</tr>
<tr>
<td>Fat mass (%)</td>
<td>18</td>
<td>38 (7)</td>
</tr>
<tr>
<td>Moderate-to-vigorous physical activity (min/week)</td>
<td>18</td>
<td>39 (34)</td>
</tr>
<tr>
<td>Diabetes duration (years)</td>
<td>18</td>
<td>4 (4)</td>
</tr>
</tbody>
</table>

^aNot applicable.
Figure 2. Illustration of the group-dependent change in total Intrinsic Motivation Inventory score and for the subscales "interest/enjoyment", "perceived competence", "perceived choice", and "value/usefulness". Pre: value at baseline; Post: value after the 24-week intervention period; n.s., not significant. *P<.05, **P<.001.

Physical Activity Adherence

Figure 3 illustrates the usage data regarding daily PA as well as patterns and total duration of in-game training as median, IQR, and range. On an average, the participants’ phones recorded 6559 (SD 1182) steps per day during the 24-week intervention period. The weekly steps during in-game walking, averaged per day, are also shown in Figure 3, with an average of 1893 (SD 723) daily steps. Participants performed an average of 4.9 (SD 1.3) in-game walking trainings per week (average duration of 26.8 [SD 8.2] minutes per training), leading to a total average duration of 131.1 (SD 48.7) minutes of in-game walking per week. The analysis of the stride cadence revealed that on an average, 83.7% (SD 3.5%) of all in-game walking was done at a cadence of ≥100 steps/minute, indicating an
average weekly amount of 109.8 (SD 43.4) minutes of moderate-to-vigorous-intensity in-game walking. One participant stopped playing the game after the third week and consequently did not produce any usage data beyond that time. We did not exclude this participant from the analyses because we consider nonuse as important of an outcome as regular use. Participants engaged in an average of 6.8 (SD 6.4) strength workouts per week with a total average duration of 15.3 (SD 24.6) minutes per week. In total, participants used the game for an average of 143.1 (SD 59.1) minutes of training per week. Overall 70.4% of strength workouts and 96.5% of walking workouts were completed, and 82.6% of all in-game workout reminders led to a completed work out on the same day.

The linear regression model (Figure 4) showed a significant positive association between total in-game training (minutes) and change in IMI total score (beta=0.0028; 95% CI 0.0007-0.0049; \( P=0.01; R^2=0.34 \)). This indicates that for every additional 30 minutes of in-game training per week during the 24-week intervention period, the total IMI score increased by 2.03 points (95% CI 0.54-3.51).

**Figure 3.** Illustration of phone-recorded physical activity data during the intervention period. A: weekly average of total steps per day, B: weekly average of steps per day during in-game training, C: average duration (minutes) of in-game strength training per week, D: total duration (minutes) of in-game training per week.
Discussion

Principal Findings

The results of this randomized controlled trial show that a novel smartphone game, whose storyline is based on established motivational elements and behavior change techniques, can significantly improve intrinsic PA motivation and lead to stable increases in PA in inactive patients with type 2 diabetes over a 24-week period. Participants in the intervention group showed significant increases in the IMI total score and the subscales “interest/enjoyment,” “perceived competence,” and “perceived choice,” indicating improvements in true PA-related intrinsic motivation and factors that predict intrinsic motivation. There was no significant change in the “value/usefulness” subscale in the intervention group; however, it is noteworthy that the baseline value of this intrinsic motivation-predicting subscale was relatively high (18.7 points) and thus did not leave a lot of room for improvement in a scale with a maximum score of 21 points. The average recorded time of 143 minutes of total in-game training per week during the intervention phase underlines the game’s strong potential in motivating formerly inactive patients with type 2 diabetes (39 minutes of moderate-to-vigorous PA per week at baseline) to meet and sustainably adhere to established PA recommendations [4]; this level also confirmed our hypothesis that the game-induced increases in intrinsic PA motivation would lead to an improved PA adherence.

Interpretation of Results

Intrinsic PA motivation increased significantly in the intervention group during the 24-week intervention period. A comparison of the magnitude of these improvements with other studies is difficult, since no studies used IMI to assess changes in intrinsic motivation in game-based interventions that promote PA. However, a cross-sectional study [36] showed that enjoyment and perceived competence of performing a television exercise program (ie, higher intrinsic motivation) are the most predictive factors of more frequent participation in the program. The recorded in-game training data in our study support this finding. By illustrating the direct impact of the game-induced increase in intrinsic PA motivation on the weekly PA behavior, the usage data provide a more tangible meaning to the shown increase in the relatively abstract IMI scores, extending beyond the primarily predictive value of the IMI.

Analysis of the recorded usage data further shows that participants walked an average of 1893 steps per day during in-game training. This is distinctly higher than the amount reported for other PA-promoting smartphone games such as Pokémon GO, even when assuming that the reported maximal increase of 955 steps per day was entirely attributable to Pokémon GO-related walking [13]. Further, in contrast to Pokémon GO, which showed a gradual decline in daily steps back to baseline values after only 6 weeks of the abovementioned initial increase in daily PA, the amount of both in-game steps and overall recorded daily steps in our study was considerably more consistent. Average daily in-game steps between 1619 and 2206 throughout the intervention period indicate a substantially better PA adherence that is especially meaningful when considering the distinctly longer timeframe of 24 weeks. The stable, objectively measured increases in weekly PA during the intervention period, along with the positive association between total in-game training (minutes) and change in IMI score, further confirm the previously indicated [25,36] importance of pursuing improvements in intrinsic motivation through increased PA enjoyment especially in behavioral interventions targeting inactive and unmotivated individuals. As this smartphone game is a motivating and enjoyable experience, the subjectively perceived cost of PA is
reduced and people are encouraged to engage in regular PA based on their personal intrinsic motivation [14,43]. Thus, our game’s narrative and gameplay may be more elaborate and the motivational elements incorporated into our storyline may be more successful in motivating players for long-term use as compared to Pokémon GO. It is further possible that our approach of tailoring the in-game PA recommendations and exercises to the fitness level of each player prevented feelings of incapability and failure that result from overwhelming PA volumes and intensities and instead enabled players to experience PA-related competence by allowing them to complete suitable workouts and meet appropriate and realistic PA goals.

The effectiveness of our game is highlighted by the fact that 83.7% of all in-game walking (110 minutes/week) was of moderate-to-vigorous intensity (≥100 steps/minute) and therefore most likely suitable to improve aerobic fitness [44] and prevent morbidity and premature mortality [45]. This weekly amount of moderate-to-vigorous-intensity walking is equal to an average duration of 15.6 minutes per day, which has been shown to be associated with a 14% reduction in all-cause mortality and has therefore been proposed as the minimum amount of PA required to extend life expectancy [46].

The average amount of time spent in in-game strength exercises was distinctly lower than that spent in in-game walking throughout the intervention period. Although the completion rate of strength workouts was considerably lower than that of walking workouts (70.4% vs 96.5%), the shorter average engagement in strength workouts was a factor of how these workouts were designed. In contrast to walking workouts, which were designed to be of sufficient intensity and duration to improve aerobic fitness [4,44] while avoiding a demotivating physical overload, strength workouts were designed as brief bouts of activity to interrupt prolonged times of sedentary behavior that could be easily integrated into the daily routine and performed anywhere without extensive equipment. This design choice is supported by recent findings, showing that interrupting prolonged sitting by brief bouts (2-5 minutes) of PA every 20-30 minutes can yield improvements in glycemic control in inactive individuals with an impaired glucose regulation for up to 22 hours [47,48]. On an average, participants made use of this design feature approximately seven times per week, with an average duration of approximately 2.5 minutes per workout, and thereby made clinically relevant changes to their sedentary daily routine.

Overall, the game encouraged an average of 143 minutes of in-game activity per week and supported patients with type 2 diabetes who were physically inactive for many years, in adopting and adhering to a physically active lifestyle that corresponds to established PA guidelines [4].

**Limitations and Implications for Future Research**

A limitation of this study is that we have no objectively measured record of any additional PA beyond the phone-recorded PA. Although participants used the game extensively as a training tool and accrued close to the guideline-recommended amount of 150 minutes of moderate-intensity PA per week during in-game training alone, it would have been interesting to see if participants engaged in any additional structured, moderate-intensity PA outside of the game. Further, although the phones recorded daily steps with a likely high accuracy when they were placed on the body [18], we have no record of the number of steps that were taken when the phones were not placed on the body. Although participants were encouraged to carry their phones with them as much as possible (ie, reminder function of the game), and based on the average daily step counts, it is likely that they did most of the time, it is quite conceivable that especially when participants were at home or work, they did not always carry their phones on them. These periods without phone wear likely led to an underestimation of unknown magnitude of the true number of daily steps. Therefore, future smartphone-based studies should consider measuring daily steps additionally with an accurate accelerometer that has the high potential for good wear-time compliance such as an activity wristband [18]. A further limitation is that we are not certain of which conceptual ideas and motivational elements incorporated into the game have indeed caused the increased intrinsic PA motivation and led to the improved PA behavior. Although it is justifiable to argue that the ensemble of all behavior change mechanics was crucial for the success of the game, a more detailed analysis would have provided important knowledge for future game designs. Finally, a follow-up assessment after an appropriate interim period should be considered to evaluate the effectiveness of the game regarding the sustainability of the improvements in PA adherence beyond the 24-week intervention period.

**Conclusions**

In summary, our randomized controlled trial shows that a novel smartphone exergame that incorporates established motivational elements and personalized PA recommendations in the storyline can generate significant increases in intrinsic PA motivation in inactive individuals with type 2 diabetes. The clinical relevance of the game-induced increases in intrinsic PA motivation is highlighted by the associated de facto and stable improvements in PA adherence during the 24-week intervention period that demonstrate the game’s suitability for successfully encouraging persistently sedentary individuals to meet and adhere to established PA guidelines. The combination of playful elements and an individualized PA promotion has high potential for success in other inactive, unmotivated target groups with and without chronic diseases.

**Acknowledgments**

We thank Promotion Software GmbH for providing technical support in programming the game application “Mission: Schweinehund” and during the intervention period. We are grateful to Novartis Pharma GmbH for the financial support during the development of the game application and RMIT University for providing support. We also thank all the participants for their
time and effort, which was critical to the successful completion of this research. This research was funded by the Swiss National Science Foundation (SNSF grant no. 166214).

**Authors' Contributions**

CH, AST, and SPW specified the research plan for the development and evaluation of the game concept. CH and AST initiated, planned, and conceptualized the present study and applied for the research grant. CH, CK, and KK were involved in the recruitment of participants and data collection and contributed to the discussion. CH and DI performed statistical analyses and created the tables and figures. CH, AST, and SPW wrote the manuscript. All authors reviewed and edited the manuscript and approved the final version to be submitted for publication.

**Conflicts of Interest**

None declared.

**Multimedia Appendix 1**

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 149KB - games_v7i1e11444_app1.pdf]

## References


17. WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects Internet. URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [accessed 2018-12-19] [WebCite Cache ID 74mhDVPJz]


Abbreviations

IMI: Intrinsic Motivation Inventory
IQR: interquartile range
PA: physical activity
Active Video Games for Rehabilitation in Respiratory Conditions: Systematic Review and Meta-Analysis

Joshua Simmich1,2, BSc (Hons), MPhtySt; Anthony J Deacon1,3, BSc (Biomed), BIT, GCELead, MBBS, MIT, MBus; Trevor G Russell1,2, BPhty, PhD (Physiotherapy)

1Centre for Research Excellence in Telehealth, University of Queensland, Brisbane, Australia
2School of Health and Rehabilitation Sciences, University of Queensland, Brisbane, Australia
3School of Electrical Engineering and Computer Science, Queensland University of Technology, Brisbane, Australia

Corresponding Author:
Joshua Simmich, BSc (Hons), MPhtySt
Centre for Research Excellence in Telehealth
University of Queensland
St Lucia
Brisbane,
Australia
Phone: 61 7 3346 9633
Email: joshua.simmich@uqconnect.edu.au

Abstract

Background: Exercise and physical activity are key components of treatment for chronic respiratory diseases. However, the level of physical activity and adherence to exercise programs are low in people with these diseases. Active video games (AVGs) may provide a more engaging alternative to traditional forms of exercise.

Objective: This review examines the effectiveness of game-based interventions on physiological outcome measures, as well as adherence and enjoyment in subjects with chronic respiratory diseases.

Methods: A systematic search of the literature was conducted, with full texts and abstracts included where they involved an AVG intervention for participants diagnosed with respiratory conditions. A narrative synthesis of included studies was performed. Additionally, meta-analysis comparing AVGs with traditional exercise was undertaken for 4 outcome measures: mean heart rate (HR) during exercise, peripheral blood oxygen saturation ($S_pO_2$) during exercise, dyspnea induced by the exercise, and enjoyment of the exercise.

Results: A total of 13 full-text papers corresponding to 12 studies were included in the review. Interventions predominantly used games released for the Nintendo Wii (8 studies) and Microsoft Xbox Kinect (3 studies). There were 5 studies that examined the acute effects of a single session of AVGs and 7 studies that examined the long-term effects after multiple sessions of AVGs. Trials conducted over more than 1 session varied in duration between 3 and 12 weeks. In these, AVG interventions were associated with either similar or slightly greater improvements in outcomes such as exercise capacity when compared with a traditional exercise control, and they also generally demonstrated improvements over baseline or nonintervention comparators. There were a few studies of unsupervised AVG interventions, but the reported adherence was high and maintained throughout the intervention period. Additionally, AVGs were generally reported to be well liked and considered feasible by participants. For outcome measures measured during a single exercise session, there was no significant difference between an AVG and traditional exercise for HR (mean difference 1.44 beats per minute, 95% CI –14.31 to 17.18), $S_pO_2$ (mean difference 1.12 percentage points, 95% CI –1.91 to 4.16), and dyspnea (mean difference 0.43 Borg units, 95% CI –0.79 to 1.66), but AVGs were significantly more enjoyable than traditional exercise (Hedges $g$ standardized mean difference 1.36, 95% CI 0.04-2.68).

Conclusions: This review provides evidence that AVG interventions, undertaken for several weeks, can provide similar or greater improvements in exercise capacity and other outcomes as traditional exercise. Within a single session of cardiovascular exercise, an AVG can evoke similar physiological responses as traditional exercise modalities but is more enjoyable to subjects with chronic respiratory diseases. However, there is very limited evidence for adherence and effectiveness in long-term unsupervised trials, which should be the focus of future research.

(JMIR Serious Games 2019;7(1):e10116) doi:10.2196/10116

http://games.jmir.org/2019/1/e10116/
Introduction

Background

Chronic respiratory diseases such as asthma [1], chronic obstructive pulmonary disease (COPD) [2], and cystic fibrosis (CF) [3] represent a significant disease burden globally, accounting for approximately 7% of all global deaths [4]. Recent systematic reviews have shown physical activity and exercise to be beneficial for pulmonary function and health-related quality of life in both asthma [5] and CF [6]. Physical activity is also negatively associated with mortality and exacerbations in COPD [7]. Exercise is a key component of pulmonary rehabilitation for COPD, which a recent systematic review has shown to improve quality of life and exercise capacity in subjects [8]. Exercise-based pulmonary rehabilitation can also be an effective intervention for other respiratory conditions [9,10]. Exercise is prescribed in clinical guidelines for both CF [11] (grade B recommendation) and COPD [12] (strong recommendation).

Physical activity is defined as any bodily movement resulting in energy expenditure, and exercise is a subset of physical activity that is planned, structured, repetitive, and aimed at improving or maintaining physical fitness [13]. Despite the clear evidence for physical activity and exercise in the management of respiratory conditions, adherence in these populations is generally low. Objective measurements indicate that almost half of adults with CF do not achieve 30 minutes of physical activity per day [14]. Similarly, in COPD, pulmonary rehabilitation has attendance and completion rates as low as 70% and 40%, respectively [15,16], and a recent meta-analysis has shown people with COPD are generally severely physically inactive [17]. Low motivation has been identified as a key internal barrier to physical activity in asthma [18], COPD [19], and CF [20], and enjoyment of physical activity has been shown to be a key motivator of exercise in these populations [20-22].

Video games have attracted attention as a novel way of delivering health care interventions with the potential to motivate change in health behaviors [23]. Video games can be designed to incorporate a physical activity or exercise component, and these are known as active video games (AVGs) or exergames [24]. The use of AVGs has been demonstrated in many age groups and clinical conditions as an enjoyable form of rehabilitation [25,26] and can also be designed to encompass other treatments such as airway clearance or spirometry [27-29]. Games used in rehabilitation may either be designed specifically for this purpose or be commercial games designed for recreational exercise that have been adapted by clinicians and researchers for use in rehabilitation.

Reviews of the use of games for purposes such as improving balance in the elderly, rehabilitation of motor function in stroke, or rehabilitation of upper limb functioning in cerebral palsy, have found small effects in favor of games over traditional therapies [25,26]. Several advantages of video games over traditional rehabilitation have been posited, such as objective measurement of performance and progress (eg, time spent playing the game, number of repetitions performed), low cost, and improved motivation and adherence [30]. The results of individual studies have reported that a single session of an AVG was considered more enjoyable than traditional exercise in subjects with CF [31,32] and COPD [33], in keeping with a systematic review that concluded that AVGs are considered enjoyable by many different populations of subjects undergoing rehabilitation [26]. Given the significant positive correlation between emotional judgment (such as enjoyment) of physical activity and frequency of physical activity in healthy adults [34,35], it is likely that interventions that improve enjoyment of exercise will improve adherence to exercise. In healthy populations, studies have found a correlation between enjoyment of an AVG and exercise adherence [36] or energy expenditure during play [37]. However, the relationship among games, enjoyment, adherence, and overall effectiveness is poorly understood, especially in chronic respiratory conditions.

To date, only 1 systematic review specifically examining the use of video games in respiratory conditions exists. This review was limited in scope as it only aimed to review the exercise intensity of AVGs for people with CF [38], and therefore it only included 5 studies. Another broad review of video games in rehabilitation did not include respiratory conditions and instead focused on aging and pathologies such as stroke and Parkinson disease for which there was more evidence [39]. Other broad reviews have focused only on pathologies for which there are randomized controlled trials looking at changes in rehabilitation outcomes over time [25] or only on games aimed at improving motor function rather than cardiorespiratory function [26].

Reviews that have included respiratory conditions have been limited to scoping reviews rather than systematic reviews [40,41].

Objectives

By comparing multiple diseases, this review provides a broader picture of the use of video games for rehabilitation in respiratory conditions than previous reviews. Moreover, it examines evidence more specific to the posited advantages of video games in rehabilitation by investigating not only clinical effectiveness of video games but also their economic feasibility and reported measures of adherence and enjoyment. Additionally, it should provide clinicians and researchers with a unified view of the literature on this newly emerging field of game-based pulmonary rehabilitation. The primary aim of this review was to evaluate the effectiveness of AVGs for 3 categories of outcomes: (1) clinical outcomes, (2) economic feasibility, and (3) patient enjoyment and adherence. The secondary aim was to compare AVGs with traditional exercise for the aforementioned outcome measures. This review will examine the hypotheses that (1) AVGs improve clinical outcomes and are economically feasible and enjoyable and (2) that AVGs produce equivalent clinical outcomes as traditional exercise programs but are more cost-effective, more enjoyable, and have higher adherence.
Methods

Search Strategy
Studies were identified by searching the following electronic databases: PubMed, Scopus, Web of Science, EMBASE, Cumulative Index to Nursing and Allied Health Literature (CINAHL), and Institute of Electrical and Electronics Engineers (IEEE) Xplore. Medical subject headings terms and equivalent subject headings were used for PubMed, EMBASE, CINAHL, and IEEE Xplore. Search terms for AVGs (eg, “video games,” “computer games,” “digital games,” and “gamification”) were combined with terms for respiratory diseases (eg, “lung,” “pulmonary,” “respiratory,” “airway,” “cystic fibrosis,” and “COPD”) derived from key terms for respiratory diseases listed in the Cochrane Airways Group Search Strategies document [42]. For IEEE Xplore, broader search terms were found to be suitable. There were no date or language restrictions placed on the search, though all search terms were in English. Full details of the final search strategies for each database are available in Multimedia Appendix 1.

The search was completed on October 30, 2017, and each database was searched from inception to the date of search. Search results were imported into Zotero (Version 4.0, George Mason University) and EndNote (Version 18.0, Clarivate Analytics) to automatically remove duplicate citations, and the results were further searched by hand to remove remaining duplicates. The combined list of citations was then imported into Covidence, a Web-based review manager (Veritas Health Innovation) for screening of titles, abstracts, and full texts. Hand-searches of reference lists and citation searches were also conducted for literature citing included studies.

Study Selection
The inclusion and exclusion criteria are outlined in Textboxes 1 and 2. Two reviewers (JS and AD) independently screened the titles and abstracts of identified papers according to inclusion and exclusion criteria, and potentially relevant papers were identified for full-text screening. Full-text papers were retrieved and screened independently by these 2 reviewers, and reasons were given for all excluded citations. In both stages of screening, disagreements were resolved by consensus or in consultation with a third reviewer (TR) if disagreement persisted.

Data Extraction and Quality Assessment
One reviewer (JS) extracted data and assessed the quality of each included study. The extracted data and the results of quality assessment were then independently checked by a second reviewer (AD) for accuracy and completeness. Quality was assessed using the Checklist for Measuring Quality assessment tool by Downs and Black [43]. Where multiple papers reported data from the same study and cohort, they were considered as 1 study for the purposes of data extraction and quality assessment.

Data Synthesis
Only studies that compared AVGs with a control exercise rather than an exercise test were included in the meta-analysis (quantitative synthesis) and only for outcome measures reported at similar time points in at least 3 different studies. A narrative synthesis was undertaken for all other studies.

A random effects model was used to assess all outcome measures. The method of Hartung-Knapp-Sidik-Jonkman (HKSJ) [44,45] was employed. This method is recommended over the more common DerSimonian and Laird (DSL) method [46,47] for meta-analyses where only few studies are available (the DSL method was also calculated for comparison). Pooled effect sizes were reported as weighted mean differences among treatments in the original units when possible, with a 95% CI. If studies used different scales, effect sizes were expressed as standardized mean differences using Hedges g [48], also with a 95% CI. Statistical heterogeneity was estimated using the point estimate and 95% CI of the $I^2$ statistic.

Textbox 1. Inclusion criteria.

- Is an experimental or quasi-experimental study of an intervention including quantitative data. This includes studies of the following designs: randomized controlled trials, quasi-randomized trials, interrupted time series, controlled before-and-after studies, uncontrolled before-and-after studies
- Participants are predominantly people diagnosed with respiratory conditions. Respiratory conditions include any disease under the National Library of Medicine Medical Subject Headings term “Respiratory Tract Diseases” (C08), including lung cancer (C08.381.540), cystic fibrosis (C08.381.187), asthma (C08.127.108), and chronic obstructive pulmonary disease (C08.381.495.389).
- Includes an intervention involving any form of video game, console game, virtual reality game, or gamified mobile app or computer program that requires players to perform a physical action that forms part of their rehabilitation, such as perform cardiovascular exercise, use a spirometer, perform airway-clearance techniques

Textbox 2. Exclusion criteria.

- Study is not experimental or does not include quantitative data. Therefore, publications of the following designs are excluded from descriptive and observational studies (case series, case-control, cohort studies, and qualitative research designs), secondary research studies (reviews and meta-analyses), and protocol papers
- Includes only interventions that do not require participants to physically perform rehabilitative actions in order to progress in the game, such as educational games that merely encourage or teach the importance of consistent medication use or physical activity
- Is not a full-text paper published in a peer-reviewed journal
Additionally, 90% CI were calculated as above for those outcome measures where included studies attempted to demonstrate equivalence among interventions (heart rate, dyspnea, and peripheral blood oxygen saturation \([S_P{\text{O}}_2]\)). To demonstrate equivalence with \(\alpha = 0.05\), the 90% CI must lie between the predetermined equivalence margins [49]. For heart rate (HR) evoked during exercise, we used a difference of 10 beats per minute (bpm) (ie, \(-10\) to \(+10\) bpm) for our equivalence margins as there is no agreed-upon minimum important difference, and 2 previous studies nominated a difference of 10 bpm as important when performing their initial sample size calculations [31,32]. For dyspnea, equivalence margins of 1 unit (ie, \(-1\) to \(+1\)) were used, as it has been suggested that the minimum clinically important difference for the modified 11-point (0-10) Borg scale is 1 unit [50]. For \(S_P{\text{O}}_2\), a drop of 4% or more has been used as a definition for exertional oxygen desaturation in CF [51,52] and COPD [53]; therefore, the width of our equivalence margins was set at 4% (ie, \(-2\)% to \(+2\)%).

Data from crossover trials were treated as paired data in the meta-analysis where raw mean differences could be calculated. An estimate of the within-subject SD was estimated where possible from the reported CIs of the difference among treatments or \(P\) values from a paired \(t\) test of the difference among treatments. From the within-subject SD, a within-subject correlation was calculated using methods outlined by Elbourne et al [54]. Where a within-subject correlation could not be calculated, the correlation from another study of similar design was used as an estimate. A sensitivity analysis, where this correlation was altered to either 0 (uncorrelated) or 1 (perfectly correlated), was conducted to test the influence this assumption had on the pooled estimate. Where results had to be reported as standardized mean differences, because of comparison across different scales, data from crossover trials were treated as if it came from independent groups.

Data reported only as median and interquartile range were converted to mean and SD using formulas from Wan et al [55]. Ordinal data, such as that from Likert or Likert-type items, were treated as interval data for the purposes of this meta-analysis. Authors of the studies included in the meta-analysis were contacted for additional information, and, when provided, these additional data were used in place of the summary data reported in the publications.

Data conversions and statistical analysis were performed using R software (Version 3.4.0, R Foundation for Statistical Computing), with meta-analysis using the metafor package (Version 1.9-9, Viechtbauer) [56].

Code used to generate the analysis and sections of the paper was written as a dynamic manuscript, using the R package rmarkdown, version 1.3 (RStudio Inc) [57], and it is available in Multimedia Appendix 2 for reproducing all data conversions and statistical analysis. Data used in this manuscript are presented in Multimedia Appendices 3 and 4.

**Results**

**Study Selection**

Figure 1 shows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram [58], which illustrates the screening and selection process. The initial literature search gave a total of 1040 results, with 612 unique records remaining after duplicates were removed—PRISMA flow diagram [58] shown in Figure 1. After screening titles and abstracts, 91 results that met the inclusion criteria remained for full-text assessment. A total of 13 papers passed both the inclusion and exclusion criteria after full-text assessment and were included in the narrative synthesis, which corresponded to 12 studies. The interventions and outcomes studied were heterogeneous, with the meta-analysis including only 4 studies for HR, 3 studies for mean \(S_P{\text{O}}_2\), 3 studies for reported dyspnea, and 3 studies for enjoyment of the exercise interventions. Manual screening of titles listed in the reference lists of included full-text papers did not yield any additional papers. One potential paper was found by citation tracking but did not pass full-text screening.

**Study Characteristics**

There were a total of 292 participants in the 12 full-text papers. Sample sizes ranged from a small observational study of 7 to larger studies with 40 participants (parallel-groups design), 30 participants (crossover-trial), and 32 participants (observational study). In total, 6 studies included participants with CF [27,31,32,59-61], 3 studies included participants with COPD [33,62,63], and 3 with other or mixed respiratory conditions [64-66]. Ages were varied because of the respiratory conditions of interest: studies of asthma included only children, studies of CF included children or young adults, and studies of lung cancer and COPD included only older adults. The randomized experimental studies often used crossover designs [27,31-33,59], with only 3 studies [60,64,66] employing 2 parallel-groups design. Nonrandomized and observational studies represented the remaining 4 studies [61-63,65]. Additional details about each included study, including the specific games played on each device, can be found in Multimedia Appendix 5.

Interventions were predominantly commercially available off-the-shelf AVGs, namely games released for the Nintendo Wii and Microsoft Xbox Kinect. The Wii was used in 8 out of 12 studies [31,33,59,60,62,63,65,66], whereas the Kinect was used in 3 out of 12 studies [32,61,64]. Only 1 study reported the testing of custom-designed hardware and games for respiratory conditions [27].
A total of 4 out of 12 included studies reported randomized studies examining the long-term effects of AVGs versus traditional training in respiratory conditions [27,60,64,66]. One study examined the effect of AVGs versus treadmill training in asthma, with training sessions twice a week over 8 weeks [64]. Another examined the effect of AVGs in combination with usual pulmonary rehabilitation compared with usual pulmonary rehabilitation alone, with 3 weeks of pulmonary rehabilitation in both groups but 1 week of AVGs in the final week of the experimental group [66]. A third compared a custom spirometer game with a control device over a 2- to 3-week period, with participants prompted but not required to use the device [27]. There were also 3 studies [62,63,65] that looked at longer-term effects of AVGs but were not randomized and controlled experiments.

There were 5 studies [31-33,59,61] that examined acute effects of a single session of an AVG. For most of these, the AVG was compared with traditional exercises [31-33,64], though some compared AVGs with maximal [61] or submaximal exercise tests [59].

Quality Assessment
Multimedia Appendix 6 describes the quality of the included studies as assessed on the Downs and Black checklist.

Due to the nature of the intervention, no study blinded participants to treatment and only 3 blinded those assessing all outcome measures [27,60,64]. Additionally, none attempted to demonstrate that their participants were representative of the population or that their participants were recruited randomly. Only 4 studies reported concealing patient allocation until the experiment was complete [31,32,60,64]. Only 6 studies reported adequate statistical power [31,32,60,61,64,66], though 1 study did perform a power calculation but was underpowered because of dropouts [62].

Meta-Analysis
A meta-analysis comparing AVGs with traditional exercise was only able to be conducted for 4 outcome measures: mean HR during exercise, $S_pO_2$ during exercise, dyspnea induced by the exercise, and enjoyment of the exercise.

Mean Heart Rate
The mean HR of AVGs and traditional exercise was assessed in 4 studies with a total of 95 participants. As shown in Figure 2, the mean difference between mean HR during AVGs and mean HR during traditional exercise was 1.44 (95% CI –14.31 to 17.18) bpm, indicating there was no significant difference between the 2 interventions. Statistical heterogeneity by $I^2$ was found to be 93% (95% CI 79%-99%).

The lower and upper bounds of the 90% CI for the difference in HR between interventions are –10.21 bpm and 13.08 bpm, respectively, which are not within the equivalence margins of –10 bpm and 10 bpm. Therefore, these treatments cannot be considered equivalent at $\alpha=.05$ level.
Mean Peripheral Blood Oxygen Saturation

The mean $S_pO_2$ of AVGs and traditional exercise was assessed in 3 studies with a total of 59 participants. As shown in Figure 2, the mean difference between mean $S_pO_2$ during AVGs and mean $S_pO_2$ during traditional exercise was 1.12 (95% CI –1.91 to 4.16) raw percentage points, indicating there was no significant difference between the 2 interventions. Statistical heterogeneity by $I^2$ was found to be 93% (95% CI 74%-100%).

The lower and upper bounds of the 90% CI for the difference in mean $S_pO_2$ between interventions are –0.94% and 3.18%, respectively, which are not within the equivalence margins of –2% and 2%. Therefore, these treatments cannot be considered equivalent at $\alpha = .05$ level.

Dyspnea

The mean dyspnea of AVGs and traditional exercise was assessed in 3 studies with a total of 59 participants. As shown in Figure 2, the mean difference between reported dyspnea during AVGs and during traditional exercise was 0.43 (95% CI –0.11 to 0.97) points on the modified 11-point (0-10) Borg scale, indicating there was no significant difference between the 2 interventions. Statistical heterogeneity by $I^2$ was found to be 47% (95% CI 0%-99%).

In the absence of other information, Cohen proposed that standardized effect sizes (such as Cohen $d$ or Hedges $g$) of greater than or equal to 0.8 be considered as large effects [67]. The effect of enjoyment of AVGs observed in this meta-analysis would therefore be considered a large effect.

Sensitivity Analysis

A sensitivity analysis was conducted by altering the correlation coefficients of crossover studies where the correlation could
not be calculated, which had been assumed to be equal to the correlation calculated in another similar study. This assumption had negligible influence on the outcome measures of HR and enjoyment and little influence on the results of dyspnea (see Multimedia Appendix 7 for full analysis).

**Narrative Synthesis**

**Active Video Games Versus Exercise Test Protocols**

Results from 2 studies examining the physiological responses to a single session of active video gaming compared with maximal exercise tests (eg, cardiopulmonary exercise testing and incremental shuttle walk test) suggest that video games usually provide a less intense exercise stimulus than a maximal exercise test. Holmes et al [61] compared the intensity of exercise elicited from the Microsoft Xbox Kinect with the maximal exercise capacity of participants with CF and found the average HR achieved during the 10-min video game session was 86% (95% CI 81-92) of the peak HR achieved during a cardiopulmonary exercise test (using a modified Godfrey Cycle Ergometer Protocol). From this, they conclude the game represented an exercise intensity of 6.1 (SD 1.8) metabolic equivalents (METS), just above the 6METS threshold for vigorous intensity activity in healthy adults [68]. They also found the game produced lower reported dyspnea, lower rating of perceived exertion, and caused less oxygen desaturation than the maximal exercise test.

Another study compared AVGs with a submaximal exercise test, such as a 6-min walk test (6MWT). Del Corral et al (2014) [59] performed a randomized cross-over observational trial to compare the physiological responses with 5 min of 3 different AVGs using the Nintendo Wii console and a 6MWT (repeated twice). They found the 2 games that were focused on aerobic exercise (EA Sports Active and Family Trainer Extreme Challenge) both elicited a higher volume of oxygen consumption (VO₂) than the 6MWT. However, the game that included balance exercises in addition to aerobic exercises (Wii FitPlus) elicited a lower VO₂ during the final 3 min of exercise than the 6MWT. HR was also lower in the Wii FitPlus game than the 6MWT but did not differ among the 6MWT and the other 2 games, and there was no statistically significant difference in dyspnea among any intervention.

**Active Video Games Versus Rest Only**

One study compared physiological responses with AVGs to rest without comparison with a control exercise. The average response to several Nintendo Wii gaming sessions in subjects with COPD was presented by Wardini et al, showing that games significantly increased HR (by a mean of 13.8 bpm) and dyspnea and decreased S̄O₂ compared with rest.

**Long-Term Physiological Effects of Active Video Games**

In total, 7 studies included more than 1 session of an AVG intervention. When compared with those seen in a control intervention, game interventions were associated with either similar or slightly greater improvements in measured outcomes. For example, in a randomized controlled trial with parallel groups, Gomes et al [64] reported both Microsoft Xbox Kinect and treadmill training for 8 weeks improved aerobic capacity and asthma control, but only the game intervention showed a statistically significant improvement in exhaled fraction of nitric oxide. Similarly, Bingham et al [27] reported forced expiratory volume in 1 second was better maintained during the game intervention than the control intervention. Mazzoleni et al [66] reported both 6MWT distance and dyspnea improved more in the participants performing pulmonary rehabilitation with an adjunct AVG than pulmonary rehabilitation alone.

AVG interventions generally demonstrated improvements over baseline or nonintervention comparators. In a randomized 6-week trial, del Corral et al [60] reported the Nintendo Wii group, but not the control group receiving no intervention, improved both modified shuttle walk test and 6MWT distances, as well as several measures of strength. Similarly, in a single-arm study, Albores et al (2013) [62] reported a statistically significant improvement in endurance shuttle walk test, number of sit-to-stands in 30 seconds and number of arm lifts in 30 seconds after 12 weeks of unsupervised Nintendo Wii training in subjects with COPD. In another single-arm study, Hoffman et al (2013) [65] reported an increase in functional performance, as measured by the number of steps per day, during the game intervention period. No studies reported AVGs were inferior to control or were associated with a negative change to an outcome measure relative to baseline. Wardini et al [63] performed a 3- to 4-week intervention but did not report changes in outcome measures over time.

**Adherence and Patient Preference**

A total of 5 studies examined adherence to unsupervised AVG treatments and generally reported adherence was higher than 70% and maintained throughout the intervention period [27,60,62,63,65]. The controlled trial by del Corral [60] reported the adherence to the home intervention was 95% during the 6-week trial, but adherence dropped in the 12-month follow-up with 65% no longer using the AVG at all. A single-arm study by Hoffman et al (2013) [65] reported a rate of adherence of 96.6% (SD 3.4%) to a Nintendo Wii intervention in subjects who had undergone surgery for suspected lung cancer. Moreover, the mean time spent performing a walking exercise with the Wii increased during the first 6 weeks’ postsurgical recovery period and was maintained for the following 10 weeks [69]. In their uncontrolled (single-arm) study, Albores et al (2013) [62] instructed subjects with COPD to use a Nintendo Wii for exercise for at least 30 minutes on most days of the week. Subjects self-reported exercising at the prescribed frequency and duration throughout the 12-week intervention, with a frequency of 5.7 (SD 1.0) days per week and a weekly total of 3.0 (SD 0.9) hours. Bingham et al [27] did not prescribe the use of their intervention; therefore, adherence could not be calculated. Nonetheless, their game intervention was used for significantly more minutes per day than the control (4.8 [SD 4.4] vs 1.6 [SD 1.8]), with no difference in the number of days each intervention was used. In contrast, in a study by Wardini et al [63], the attendance rate to the game intervention (64% [SD 35%]) was lower than the attendance rate for the pulmonary rehabilitation program (88% [SD 13%]), though the game was delivered as an adjunct rather than an alternative program and no statistical comparison was performed.
Measures of patient preferences, likeability, or perceived feasibility were measured in several studies (in addition to enjoyment, which was summarized in the meta-analysis), showing that games were generally considered likeable and feasible interventions [31,33,63,66]. For example, Kuys et al [31] reported that both the Nintendo Wii intervention and the treadmill or bicycle control exercise were rated highly effective and feasible without significant differences between the 2 interventions. Similarly, LeGear et al [33] reported similar responses between game exercise and treadmill exercise for perceived safety and whether the participants could see themselves performing that exercise at home. Using a 7-item, 7-level Likert-type scale, Mazzoleni et al [66] reported no significant difference between the acceptability of pulmonary rehabilitation alone (43.9 [SD 3.0]) and with active video gaming as an adjunct (42.4 [SD 3.5]). Wardini et al [63] reported the overall enjoyment of the adjunct game intervention to be 8.0 cm (SD 2.6 cm) on a 10-cm visual analog scale. Moreover, participants gave a mean score of 8.0 (SD 2.6 cm when asked if they would recommend the adjunct program to another patient with COPD and a rating of 7.0 cm (SD 3.7 cm) when asked if they would consider purchasing a game system of their own.

Adverse Events

No studies reported the occurrence of adverse events linked to AVGs. Wardini et al [63] reported 1 patient with COPD and coronary artery disease requiring the use of nitroglycerin and 5 subjects experiencing transient but asymptomatic oxygen desaturation below 85%, but they did not report whether these were a result of the AVG intervention or the control intervention.

Economic Analysis

No studies performed an economic analysis comparing AVGs with alternative forms of exercise or rehabilitation.

Discussion

Summary of Findings

This systematic review aimed to examine the hypotheses that (1) AVGs improve clinical outcomes and are economically feasible and enjoyable and (2) that AVGs produce equivalent clinical outcomes as traditional exercise programs but are more cost-effective, more enjoyable, and have higher adherence. AVGs generally demonstrated improvements in exercise capacity or functional performance over baseline or nonintervention comparators, and they were generally reported to be well liked by participants. There were no differences among AVGs in physiological outcomes when compared with traditional exercises (such as treadmill or stationary cycling) in the laboratory over a single session, but AVGs were enjoyed more by participants. Despite a few long-term trials, the limited evidence suggests AVGs may provide similar or greater effects on clinical outcome measures compared with traditional forms of exercise. None of the included studies provided any evidence for the economic feasibility of the interventions.

The results of this review corroborate the findings of a systematic review into the effects of AVGs on people with CF [38], which found that the HR response to active video gaming was similar to the traditional methods of physical training. One of the studies included in that review was excluded from this review as it did not compare AVGs with traditional exercise or with rest [70]. This study found that the HR target for moderate physical activity (64% of predicted maximum HR) was met during the boxing activity of the Nintendo Wii game Wii Sports and exceeded by the free jogging activity in the Nintendo Wii game Wii Fit in children with CF.

Similarly, the results of this review are consistent with a meta-analysis of AVGs for healthy adults and children by Peng et al [71], who concluded that the HR, energy expenditure, and oxygen consumption were no different between AVGs and traditional physical activities. In contrast, a more recent meta-analysis on AVGs in healthy children [72] found that AVGs had a large positive mean effect on HR, small positive mean effect size on enjoyment, and a similar rate of perceived exertion compared with traditional exercise in a laboratory setting.

Physiological Responses to Active Video Games

The results of the meta-analysis indicate that the physiological response to AVGs was not significantly different from traditional exercises (Figure 2). However, statistical equivalence was not demonstrated for any of these physiological outcomes, possibly because of there being so few included studies and the heterogeneity among studies.

The physiological response to AVGs may partly derive from psychological arousal rather than the physical workload of the activity. An increase in the HR and respiratory rate has been observed in healthy adults and children playing inactive video games with traditional controllers [73,74]. Similarly, Sherman et al [75] found that the HR changes were similar whether the game was controlled with a joystick or a motion sensor detecting the movement of the player’s body. However, the HR changes reported in the studies included in this review are greater than those seen in response to inactive video games. Additionally, the pooled effect of AVGs on HR and dyspnea was no different from traditional moderate-intensity exercise; therefore, psychological arousal likely had only a small or negligible contribution to the observed physiological responses.

Furthermore, in each study included in the meta-analysis, the participants were instructed to perform both AVGs and traditional exercises at similar self-perceived intensities and supervised to ensure compliance. For example, in both Kuys et al [31] and Salomini et al [32], the participants in each intervention were instructed to exercise at an intensity corresponding to 3 to 5 on the modified Borg scale. Therefore, it is not known whether these interventions would produce similar workloads in a situation where subjects were unsupervised or allowed to exercise at a self-selected intensity. Additionally, 1 included study [32] did not use the 11-point (0-10) modified Borg scale used by other included studies but rather a visual analog scale that was also 0 to 10, though the results were nonetheless pooled as if they are equivalent scales. Dyspnea assessed on a visual analog scale correlates strongly with the modified Borg scale [76,77], but the correlation between these 2 assessment scales is not perfect, and it therefore may have been a source of additional variability. Furthermore,
Salonini et al [32] assessed dyspnea during exercise, whereas other studies [31,33] assessed dyspnea at the end of the exercise bout. Both of these differences in assessment among studies may have further contributed to the heterogeneity observed in the meta-analysis.

Additionally, the meta-analysis pooled values for mean \(S_O_2 \) as this was the most commonly reported measure, but minimum \(S_O_2 \) may be more relevant to detecting the occurrence of exertional oxygen desaturation that can occur in CF [51,52] and COPD [53,78]. Only 2 papers reported minimum \(S_O_2 \). Kuys et al [31] found that minimum \(S_O_2 \) was 2% lower in traditional exercise compared with game-based exercise, but this was not a significant difference (95% CI –2 to 6). LeGear et al [33] reported that minimum \(S_O_2 \) was 91.3% during the game versus 88.7% in the treadmill, but they did not perform a statistical comparison between these interventions (nor report sufficient data to enable one to be calculated).

Other Responses to Active Video Games

The pooled estimate of enjoyment of AVGs was significantly higher than traditional exercises; however, this estimate was based upon a small number of studies with heterogeneous outcome measures. Each included study assessed enjoyment using different measures and none used an experimentally validated measure of enjoyment, which limits comparison among included studies and with the literature as a whole. Established, validated instruments for assessing enjoyment of exercise include the Physical Activity Enjoyment Scale [79] or Feeling Scale [80]. In healthy sedentary adults, an increase of 1 unit on the 11-point bipolar (–5 to +5) Feeling Scale during moderate-intensity treadmill walking is associated with an increase of 15 minutes of weekly physical activity at 6-month follow-up [81], though this result was observational, and causation therefore could not be demonstrated. Nonetheless, it is possible that the 2-unit difference on the 0 to 10 cm visual analog scale in response to AVGs compared with traditional exercise reported by Kuys et al [31], which corresponds with the pooled estimate of the meta-analysis of this study (as shown in Figure 2), could result in at least a similar increase in physical activity in subjects with respiratory diseases.

The included studies assessed enjoyment of exercise by assessing participants after exercise rather than during. Enjoyment assessed after exercise may be biased because of either the time between exercise and giving an enjoyment response or to an emotional response to completion of the task.

Some studies used a continuous mode of exercise as a control (continuous cycling or treadmill walking), whereas participants playing AVGs performed periods of exercise interspersed with rest periods between game levels [32,33]. Interval training has been demonstrated to be more enjoyable than continuous exercise in healthy active populations [82,83], which may explain at least part of the enjoyment effect of AVGs over traditional exercise in the included studies. Additionally, the video game intervention often involved different exercises and a greater variety of exercises compared with the control [31,33]. Thus, it is possible that a control intervention, which also comprised a variety of exercises (eg, dancing, boxing, and running) may have been just as enjoyable as the AVG. In future, studies should compare AVGs with interval training comprising a routine or circuit with an equivalent variety of exercises.

Additionally, several included studies noted that the greater enjoyment found in AVGs may be because participants found them more novel than traditional exercises. Studies of longer duration must be performed to see if enjoyment is maintained in subjects with respiratory diseases using AVGs as part of their rehabilitation and if this corresponds to a greater adherence to exercise. Most studies used supervised exercise training and therefore could not assess the effect of video games on adherence to unsupervised home exercise or self-chosen duration of exercise. The greater enjoyment seen in video games could fail to translate to greater adherence or longer durations of exercise, perhaps because the additional technological requirements of AVGs compared with other modes of exercise (such as jogging) may decrease adherence in a real-world situation.

No adverse events linked to the AVG intervention were reported in any included studies, though the small sample sizes of included studies may have limited their ability to detect rare adverse events.

Strengths and Limitations

This review includes evidence from multiple respiratory conditions, which provides a broader picture using a greater quantity of evidence than is currently available for any condition individually. Additionally, despite the different methodologies and reported outcome measures, this review transformed and standardized the data as best as possible to enable a meta-analysis to be conducted, using the conservative HKSJ approach [44,45].

The main limitation of this review is the small number and significant heterogeneity of the included studies. This review mostly included relatively small studies that examined multiple different respiratory diseases, within different age groups and using different interventions (eg, types of software and technology). A subgroup analysis was not performed for age groups or disease populations as the evidence was limited to just 1 or 2 studies in each group. This clinical heterogeneity was the reason for performing a narrative synthesis for several outcomes. The statistical heterogeneity was also high; however, because of the low number of studies, even the estimates of statistical heterogeneity were highly uncertain.

An additional limitation is the inclusion of only studies published in full, which ensured data came only from high-quality studies but may have caused the review to be vulnerable to publication bias, which also may have limited previous reviews [26]. Language bias may also have been present as all search terms were in English, although indexed terms were used, which may have minimized potential language bias. Finally, the HKSJ estimation method, though chosen because it provides a more conservative estimate than more common approaches, may overestimate uncertainty with so few included studies [46].
Conclusions
The results of this systematic review and meta-analysis indicate that in a single session, AVGs, when used for cardiovascular exercise, can evoke similar exercise intensities to those produced by traditional exercise modalities but are considered more enjoyable by subjects with respiratory diseases. However, little evidence exists from unsupervised long-term trials of AVGs used for any rehabilitative purpose in respiratory conditions, though some limited evidence from supervised longer-term trials indicates that AVGs may be equal or slightly superior to traditional exercise training in some outcome measures.

Future research is needed regarding the long-term effects of AVGs, the use of AVGs for other rehabilitative purposes such as strength training or airway clearance techniques, and the economic benefits of utilizing AVGs rather than traditional supervised or unsupervised rehabilitation programs. Additionally, future work could explore the differences between technologies (game consoles) and software (individual games) to enable the creation of more effective gaming interventions for use in rehabilitation. Finally, when assessing the enjoyment of game interventions, validated outcome measures should be used to facilitate comparison to established research on enjoyment of exercise interventions.

Acknowledgments
The authors thank Dr Michael David (University of Queensland, Brisbane, Australia) for his advice regarding the statistical methods used in this meta-analysis.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Search terms used for each database.

[PDF File (Adobe PDF File), 34KB - games_v7i1e10116_app1.pdf ]

Multimedia Appendix 2
Compressed archive containing source files (in RMarkdown format), plus necessary associated formatting templates and citation files, to generate the submitted manuscript. Requires raw data files (see other appendices).

[ZIP File (Zip Archive), 749KB - games_v7i1e10116_app2.zip ]

Multimedia Appendix 3
Raw data for results of database searches and screening process.

[ZIP File (Zip Archive), 342KB - games_v7i1e10116_app3.zip ]

Multimedia Appendix 4
Raw data extracted from included studies, for characteristics of included studies and outcome variables for meta-analysis.

[ZIP File (Zip Archive), 2KB - games_v7i1e10116_app4.zip ]

Multimedia Appendix 5
Characteristics of included studies.

[PDF File (Adobe PDF File), 44KB - games_v7i1e10116_app5.pdf ]

Multimedia Appendix 6

[PDF File (Adobe PDF File), 666KB - games_v7i1e10116_app6.pdf ]

Multimedia Appendix 7
Sensitivity analysis for meta-analysis.

[PDF File (Adobe PDF File), 23KB - games_v7i1e10116_app7.pdf ]
References


40. Costa TH, Soares NM, Reis WA, Bublitz FM. A systematic review on the usage of games for healthcare. 2015 Presented at: IEEE 5th International Conference on Consumer Electronics-Berlin (ICCE-Berlin); September 6-9, 2015; Berlin, Germany. [doi: 10.1109/ICCE-Berlin.2015.7391316]


Abbreviations

AVG: active video game
bpm: beats per minute
CF: cystic fibrosis
CINAHL: Cumulative Index to Nursing and Allied Health Literature
COPD: chronic obstructive pulmonary disease
DSL: DerSimonian-Laird
HKSJ: Hartung-Knapp-Sidik-Jonkman
HR: heart rate
IEEE: Institute of Electrical and Electronics Engineers
METS: metabolic equivalents
6MWT: 6-min walk test
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SpO2: peripheral blood oxygen saturation
VO2: volume of oxygen consumption
Creating a Theoretically Grounded Gaming App to Increase Adherence to Pre-Exposure Prophylaxis: Lessons From the Development of the Viral Combat Mobile Phone Game

Laura Whiteley¹*, MD; Leandro Mena²*, MD, MPH; Lacey K Craker³*, MPH; Meredith Garver Healy³*, BA; Larry K Brown¹,³*, MD

¹Department of Psychiatry, Warren Alpert Medical School, Brown University, Providence, RI, United States
²University of Mississippi Medical Center, University of Mississippi, Jackson, MS, United States
³Young Adult Behavioral Health, Department of Psychiatry and Human Behavior, Rhode Island Hospital, Providence, RI, United States
* all authors contributed equally

Corresponding Author:
Laura Whiteley, MD
Department of Psychiatry
Warren Alpert Medical School
Brown University
Residency Training Program-Brown University-Butler Hospital
345 Blackstone Boulevard
Providence, RI, 02906
United States
Phone: 1 401 455 6375
Email: laura_whiteley@brown.edu

Abstract

Background: In the United States, young minority men who have sex with men (MSM) are most likely to become infected with HIV. The use of antiretroviral medications to reduce the risk of acquiring HIV infection (pre-exposure prophylaxis, PrEP) is an efficacious and promising prevention strategy. There have been significant advances regarding PrEP, including the definitive demonstration that PrEP reduces HIV acquisition and the development of clinical prescribing guidelines. Despite these promising events, the practical implementation of PrEP can be challenging. Data show that PrEP’s safety and effectiveness could be greatly compromised by suboptimal adherence to treatment, and there is concern about the potential for an increase in HIV risk behavior among PrEP users. Due to these challenges, the prescribing of PrEP should be accompanied by behavioral interventions to promote adherence.

Objective: This study aimed to develop an immersive, action-oriented iPhone gaming intervention to improve motivation for adherence to PrEP.

Methods: Game development was guided by social learning theory, taking into consideration the perspectives of young adult MSM who are taking PrEP. A total of 20 young men who have sex with men (YMSM; aged 18-35 years) were recruited from a sexually transmitted infection (STI), HIV testing, and PrEP care clinic in Jackson, Mississippi, between October 2016 and June 2017. They participated in qualitative interviews guided by the information-motivation-behavioral skills (IMB) model of behavior change. The mean age of participants was 26 years, and all the participants identified as male. Acceptability of the game was assessed with the Client Service Questionnaire and session evaluation form.

Results: A number of themes emerged that informed game development. YMSM taking PrEP desired informational game content that included new and comprehensive details about the effectiveness of PrEP, details about PrEP as it relates to doctors’ visits, and general information about STIs other than HIV. Motivational themes that emerged were the desire for enhancement of future orientation; reinforcement of positive influences from partners, parents, and friends; collaboration with health care providers; decreasing stigma; and a focus on personal relevance of PrEP-related medical care. Behavioral skills themes centered around self-efficacy and strategies for adherence to PrEP and self-care.

Conclusions: We utilized youth feedback, IMB, and agile software development to create a multilevel, immersive, action-oriented iPhone gaming intervention to improve motivation for adherence to PrEP. There is a dearth of gaming interventions for persons
on PrEP. This study is a significant step in working toward the development and testing of an iPhone gaming intervention to decrease HIV risk and promote adherence to PrEP for YMSM.

**KEYWORDS**
cell phone; HIV; young adult; sexual and gender minorities

**Introduction**

**Background**
The primary prevention of HIV infection remains a crucial priority. In 2016, there were 39,782 new HIV diagnoses in the United States, and 67% of these new diagnoses were of gay and bisexual men [1]. HIV also disproportionately affects communities of color. Despite only comprising 12.6% of the US population, nearly 45% of all new HIV infections occur among African Americans. The use of antiretroviral medications to reduce the risk of acquiring HIV infection (pre-exposure prophylaxis, PrEP) is an efficacious and promising new prevention strategy [1,2]. There have been significant advances regarding PrEP including the definitive demonstration that PrEP reduces HIV acquisition. Despite promising data, the practical implementation of PrEP is challenging [3-10]. Unfortunately, individuals who are often at the highest risk of HIV infection and eligible for PrEP come from populations that historically have been underserved by health care [11,12]. Therefore, engaging diverse persons on PrEP in care is challenging, and reinforcement and support for patients and doctors is required [11,13-15]. Behavioral Interventions promoting adherence to comprehensive PrEP treatment and healthy sexual behaviors will need to be tailored to diverse, underserved, and at-risk populations and will need to reinforce the clinician-patient relationship.

In addition, the behavioral interventions accompanying PrEP need to be scalable, cost-effective, and easily integrated into clinical settings [2,16]. Without these necessary components, integration of behavioral interventions into clinical settings cannot be realistically sustained. Utilizing gaming technology to deliver behavioral interventions for young men who have sex with men (YMSM) can improve intrinsic motivation and information and build behavioral skills for adherence. The use of an intervention that utilizes gaming technology is particularly compelling for use with younger adults, as this age group is at the highest risk of acquiring HIV and this age group has most actively inquired about PrEP in clinical settings. In addition, 72% of young male adults in the United States say they play video games often or sometimes [17]. Gaming technology is also popular among minority men who have sex with men (MSM), the subgroup most at risk for acquiring HIV. In the past, gaming was mistakenly identified as a primarily adolescent and heterosexually dominated activity, but current data support that gaming is actually quite diverse [18]. MSM gamers, referred to as gaymers in pop culture and mainstream articles, are highly represented on the web and increasingly shape the market [19-21]. In addition, within the United States, African Americans aged 18 to 35 years represent the most active and fastest growing user group of mobile phones [22,23], and more than half of the reported video gaming occurs on portable devices [18,21,22].

Interactive game play has been shown to enhance players’ motivation to improve health behaviors and self-care in a variety of clinical settings and populations. Games can attract and maintain attention, a key component for effective behavior change. Compelling interactive phone-based games can expose players to essential health-related content thousands of times and also give players unlimited opportunities to rehearse new skills and receive personalized feedback on health choices made within the game [28-30]. Games have been shown to be efficacious in promoting fitness, improving weight management, and improving safer sex skills [27,28,30,31]. For example, an HIV and AIDS prevention computer game called Life Challenge was developed by the New York State Department of Health to enhance safer sex negotiation by adolescents and young adults. The game showed significant improvement in self-efficacy for partner negotiation and condom skills for those who started with the least self-efficacy [30]. Two pregnancy prevention games, The Baby Game and Romance, designed for sexually active young adults, showed trends in improving knowledge and attitudes about parenting and unprotected sexual behaviors [29]. Video games have also been applied to improve self-management skills and healthy behaviors in those living with asthma, diabetes, and cancer [32-36]. For example, a video game named Re-Mission (tested with adolescents and young adults aged 13-29 years) was designed as an action-adventure game named Re-Mission (tested with adolescents and young adults aged 13-29 years) was designed as an action-adventure within the game [28-30]. Games have been shown to be efficacious in promoting fitness, improving weight management, and improving safer sex skills [27,28,30,31]. For example, an HIV and AIDS prevention computer game called Life Challenge was developed by the New York State Department of Health to enhance safer sex negotiation by adolescents and young adults. The game showed significant improvement in self-efficacy for partner negotiation and condom skills for those who started with the least self-efficacy [30]. Two pregnancy prevention games, The Baby Game and Romance, designed for sexually active young adults, showed trends in improving knowledge and attitudes about parenting and unprotected sexual behaviors [29]. Video games have also been applied to improve self-management skills and healthy behaviors in those living with asthma, diabetes, and cancer [32-36]. For example, a video game named Re-Mission (tested with adolescents and young adults aged 13-29 years) was designed as an action-adventure within the game [28-30]. Games have been shown to be efficacious in promoting fitness, improving weight management, and improving safer sex skills [27,28,30,31]. For example, an HIV and AIDS prevention computer game called Life Challenge was developed by the New York State Department of Health to enhance safer sex negotiation by adolescents and young adults. The game showed significant improvement in self-efficacy for partner negotiation and condom skills for those who started with the least self-efficacy [30]. Two pregnancy prevention games, The Baby Game and Romance, designed for sexually active young adults, showed trends in improving knowledge and attitudes about parenting and unprotected sexual behaviors [29]. Video games have also been applied to improve self-management skills and healthy behaviors in those living with asthma, diabetes, and cancer [32-36].
intervention to improve antiretroviral treatment (ART) adherence for youth and young adults living with HIV that informs the development of this game [40]. LeGrand et al have published a description of the development phase of a game entitled Epic Allies. Epic Allies is designed to improve ART uptake, engagement in care, and adherence among HIV positive YMSM and transgender women who have sex with men. While playing Epic Allies, users can earn medals and tokens for taking medications and reading health-related studies. [42]. Another intervention called PlayForward aims to reduce the risk for HIV among at-risk, ethnic, and racial minority adolescents. This tablet-based game provides an interactive world using an avatar where players face challenges such as peer pressure to drink alcohol or engage in other risky sexual behaviors [43]. A mobile phone-optimized intervention entitled healthMpowerment is designed to reduce sexual risk behaviors among YMSM. In this intervention, YMSM can acquire reputation points through reading information about HIV, playing sexually transmitted infection (STI)–related games, and through positive interactions with other users. Despite the promise and popularity of digital games, there are no published or presented abstracts related to gaming for persons on PrEP that we could find.

**This Study**

We developed a gaming intervention to improve adherence to PrEP for YMSM aged 18 to 35 years. This intervention integrates a smart pill bottle cap (that measures adherence) with an iPhone game. The gaming intervention was informed by the information-motivation-behavioral skills (IMB) theory of learning. The IMB model, consistent with social learning theory, is broadly applicable and can be used to guide game development and create theoretically consistent gaming content [46-50]. The iPhone gaming app was designed for participants to experience absorbing action-oriented adventures that increase information about their health (eg, knowledge about PrEP and HIV prevention and adherence), improve motivation (eg, action figures experience health benefits of adherence to PrEP), and build skills (action figures interact with clinicians at appointments, take medications as prescribed, and practice safe sex; see Multimedia Appendices 1-13). Adherence (measured by the smart pill bottle cap and game-related text messages) are integrated into the gaming intervention. Multiple reviews have demonstrated that behavioral interventions shown to be most efficacious are those tailored for the target population and preceded by formative research to inform intervention development [51,52]. The aim of this study was to describe the development of this iPhone gaming app entitled ViralCombat.

**Methods**

**Gaming App Development**

Development of ViralCombat was accomplished using iterative and collaborative procedures to integrate the clinical experiences of YMSM taking PrEP, academic researchers, and technology partners. Game development was guided by qualitative interviews with a group of MSM aged between 18 and 35 years taking PrEP. Guided by the principles of agile software development [53], the qualitative interviews and the game and app programming were synergistic. Agile software development aims for continuous design testing and adaptation based on continuous feedback [53].

**Sample and Recruitment**

Males aged between 18 and 35 years were eligible for enrollment in the study according to the following criteria: (1) English-speaking, (2) currently taking PrEP, (3) reported having sex with other men, and (4) able to give consent and not impaired by cognitive or medical limitations as per clinical assessment. Those who did not meet the abovementioned inclusion criteria were excluded. We recruited 20 YMSM for qualitative interviews to guide game development after institutional review board’s (IRB) approval. Subjects were from a convenience sample recruited from a PrEP clinic in Jackson, Mississippi, between October 2016 and June 2017. Subjects were approached by research staff with an IRB-approved flyer, and written consent was obtained upon meeting with study staff for the qualitative interview. Overall, 20 subjects were approached over the course of the interviews, and all of them consented and completed the interview. Subjects were recruited until data saturation was achieved, and a relative balance in the sample was achieved based on age greater than 26 years versus younger. The mean age of participants was 26 years (range: 18-35 years; 11 out of the 20 were younger than 26 years). A majority of the participants (17/20, 85%) identified as African American, 100% (20/20) identified as MSM, 70% (13/20) completed 12th grade, 75% (15/20) had been taking PrEP longer than 6 months, 75% (15/20) reported missing a dose of PrEP in the previous week, 45% (9/20) reported missing 3 doses or more in the previous week, and 75% (5/20) had been taking PrEP for 6 months or longer.

**Procedures**

The study was approved by the hospital’s IRB, and participant consent and interviews were conducted in a private room located in a PrEP clinic in Jackson, Mississippi. Interviews were conducted by 1 of the 2 MDs (psychiatrists) with support from a trained research assistant, and subjects were reimbursed US $50 per survey (for a total of US $150) for their time. The psychiatrists who conducted the interviews did not provide medical or clinical services in the HIV clinic. Interviews lasted between 45 and 60 min and were digitally recorded. As we adapted our gaming intervention from games that were already developed (Dr. Nano X and BattleViro), the system and the framework (eg, code, database, and design) were already in place at the beginning of the project. Adaptations to the game occurred as themes emerged from the interviews [41]. Biweekly meetings were held with the programmers to discuss all adaptations, including changes to content, game graphics framework, and game messaging.

**Adaptation of the Information-Motivation-Behavioral Skills Adherence Gaming Intervention**

The IMB gaming app for persons on PrEP, entitled ViralCombat, was adapted from the popular Mission Critical Studios game entitled Dr. Nano X as well as our previously developed game for persons on ART entitled BattleViro [40,41]. Dr. Nano X is a 5-star-rated mobile game (the highest rating possible) in the iTunes app store and is available on both Android and iPhone.
The ART adherence game, named BattleViro, has been previously described in the Journal of Medical Internet Research [41]. We worked directly with the development team at Mission Critical Studios to develop ViralCombat using Dr. Nano X and BattleViro as a basic framework. Adapting our game to promote PrEP adherence from already existing games greatly decreased the cost of the project. Characters, actions, and IMB messaging about PrEP were built specifically for ViralCombat; however, we were able to reuse backgrounds, mechanisms of game play and controls, and many sound effects from Dr. Nano X and BattleViro [54,55]. ViralCombat game levels are distinct from those in BattleViro, and they include organ systems such as the penis, anus, and mouth.

We worked iteratively with youth living with HIV and Mission Critical Studios to create the app. The game starts with a 45-second narrative movie that explains the storyline and game objectives. The player is told by the narrator, in a deep and dramatic voice, “you have been chosen, due to your smart decisions and healthy choices, to be cloned and shrunken in order to enter your own body to destroy attacking viruses and infections.” The narrator emphasizes that “in order to acquire ammunition and strength, you must take you medication in real life and also pick up pills during play.” Players are given a tutorial on game actions and can also design their individual and diverse characters. As players successfully battle HIV, engage with providers, and take medication, they move to new, distinctive levels (arterial system, penis, anus, and mouth). Messages from the doctors, nurses, and friends encourage and provide clues during difficult twists and turns in the battle. Answering quiz questions from clinician avatars allows each player to earn strength and points; wrong answers are corrected and explained. Players find medication, strength, and points by acting on positive suggestions. During each mission, the player’s score (pill count and health) is shown. All character control and gaming is done by touch screen technology on the phone; no additional accessories are needed for play. Throughout the game, the terms “HIV” and “PrEP” are seen on the screen, but they are never part of an audio feature, so stigmatizing information cannot be overheard by another person. The game-related text messages have gaming graphics, similar to other apps and games, and the app uses push notifications to engage gamers in play. If players were less than 90% adherent during the week, gaming is done by touch screen technology on the phone; no additional accessories are needed for play. Throughout the game, the terms “HIV” and “PrEP” are seen on the screen, but they are never part of an audio feature, so stigmatizing information cannot be overheard by another person. The game-related text messages have gaming graphics, similar to other apps and games, and the app uses push notifications to engage gamers in play. If players were less than 90% adherent during the week, they include organ systems such as the penis, anus, and mouth.

**Information Needed for Adherence**

Participants were asked about knowledge and information that influences their adherence to PrEP and engagement in medical appointments. Questions included:

- **What type of information from doctors or friends makes it easier to take PrEP?**
- **What information makes it easier to come to PrEP related appointments?**

This part of the interview aimed to understand the specific knowledge about HIV and PrEP that promotes adherence behaviors. For example, some probes focused on how side effects and other health-related information can influence adherence to medication and care (for more examples, see Textbox 1).

**Motivation for Adherence**

Participants were queried about motivational issues related to adherence to PrEP with probes such as:

- **I would like to hear about what you think the serious issues are surrounding PrEP and taking medications to prevent HIV.**
- **What are the things that make it hard to take PrEP?**

This part of the interview was dedicated to understanding both personal and social motivations for adherence to PrEP. Queries were focused on the positive and negative attitudes toward taking PrEP, perceived negative effects of nonadherence to PrEP, and the individual’s perceptions of social support from significant others, family, friends, and medical care providers (for more examples, see Textbox 1).

**Behavioral Skills for Adherence**

Participants were also asked about their ability to perform necessary adherence and HIV preventative tasks and their perceived self-efficacy for these tasks. Questions included:

- **What are the ways that you stay safe from HIV?**
- **What are the ways that you remember to take PrEP and remember your PrEP related appointments?**
- **What events in your life make it harder to remember to take PrEP? Or remember your appointments?**
Textbox 1. Qualitative interview guide based on the information-motivation-behavioral (IMB) skills model.

<table>
<thead>
<tr>
<th>Information</th>
<th>Motivation</th>
<th>Behavioral skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>What knowledge or information about HIV and pre-exposure prophylaxis (PrEP) is helpful to know?</td>
<td>What are the main issues in coming for PrEP related medical appointments?</td>
<td>Do you use alarms, your phone, or reminders to remember to take PrEP?</td>
</tr>
<tr>
<td>What knowledge or information helps you to take PrEP daily?</td>
<td>What are the things that make it hard to take PrEP?</td>
<td>What do you do if you miss a dose of PrEP?</td>
</tr>
<tr>
<td>Does knowing about side effects change decision making around taking PrEP?</td>
<td>What are the attitudes or feelings that people like you have that make it harder to take PrEP? Or easier to take PrEP?</td>
<td>What are the strategies for adherence to PrEP over time and across different situations?</td>
</tr>
<tr>
<td></td>
<td>How do partners, your family, and your community play a role in adherence to PrEP related care?</td>
<td>Are there things that you do such as eating, or avoiding certain substances, that make taking PrEP easier?</td>
</tr>
</tbody>
</table>

General gaming attitudes

- What is your reaction to getting some PrEP related and HIV related information and skills in a game?
- Do you ever play games that teach you facts or in which you learn something?
- Do you go online or use your phone to learn information about your health?
- Have you ever played a health-related game before on your phone or at a computer?

Reactions to ViralCombat

- What did you like and not like about the game?
- What do you think this activity is trying to teach you?
- How much did the material look like the other games that you play?
- How could this activity or content be improved for others like you?
- Now that you have seen this game, would you want to play it?
- Before you came here today, did you ever find anything like this on a game on a phone or on a computer?
- Would you be worried about playing the game when others could see it?
- What would you say if someone asked you about the game?

We also asked participants about strategies for self-reinforcement for adherence over time and across different situations. We asked questions such as:

Do you consciously think about your PrEP medication schedule on a long-term basis?

What strategies have you used or developed to remember to take PrEP or to go to your appointments based on your activities?

This part of the interview aimed to assess perceived abilities and strategies to store, obtain, and self-cue the use of medications despite challenges and across situations (for more examples, see Textbox 1).

General Gaming Attitudes

Participants were also asked about their general attitudes and experiences with games. Participants were asked questions such as:

What games do you, or people you know, play on the cellphone?

What types of graphics, avatars, and rewards do you like? And what do you not like?

How are games useful? Do you develop any skills when you play games?

These queries elicited descriptions of popular game activities and attitudes about gaming. The responses were used to make the format and game mechanics of ViralCombat engaging and immersive (for more examples, see Textbox 1).
**ViralCombat Storyboard and iPhone Game**

Participants were asked for feedback about the storyboard or the iPhone game (once the mobile game was ready) with the probes such as:

- What was the main point of this activity?
- What could you learn from this activity?
- Would you recommend this type of game to your friends?
- What is your reaction to having some HIV and PrEP related information and skills in an iPhone game?

After the first version of the game was developed on the iPhone, participants were asked additional and modified probes such as:

- Is the game easy to navigate and easy to understand?
- Did any part of the game not work?
- Are there other topics that the game should cover that it does not?

Answers to these questions guided the iterative development of the game levels, actions, characters, and graphics (for more examples, see Textbox 1).

**Medication Adherence Monitoring Tracking and Game-Related Text Messages**

Participants were also asked about the electronic pill monitoring organizers and adherence-related text messages. We queried participants about the smart pill bottle cap that can electronically monitor, measure, and securely relay adherence pill bottle openings to our research team. Participants were asked about the use of reminder messages with a game-related graphic. Feedback was elicited about game-related messages if players missed a dose using textual phrases such as “Missing you in Combat” and “Get in the game.” Feedback was also received about SMS text messages if doses are taken on time such as “Great job in battle” and “You are fighting off virus well!” (see Multimedia Appendix 1).

**Quantitative Feasibility and Acceptability Data**

The game was developed iteratively. A total of 11 participants were shown the storyboard of the game for approximately 30 min; qualitative feedback was elicited, and game development on the iPhone occurred. After the development of the iPhone version of ViralCombat, 9 of the 20 participants played the game for 45 to 50 min with the interviewer in the room (see Multimedia Appendices 1-13). These 9 participants were shown each of the game levels on the phone by the interviewer and played each level. After playing the game, these 9 participants completed qualitative interviews, and then, written or quantitative feedback was obtained. Quantitative feedback was collected from these 9 participants using versions of the Client Service Questionnaire (CSQ) and the session evaluation form (SEF) [56,57]. Both these instruments were developed to measure client satisfaction and perspectives on intervention aspects. The CSQ consists of 8 items that assess general satisfaction with the game. An example query from the CSQ is “In an overall, general sense, how satisfied are you with the game?” (for which the response options are 4=“Very satisfied”; 3=“Mostly satisfied”; 2=“Indifferent or mildly dissatisfied”; and 1=“Quite dissatisfied”) [56]. We did not calculate an alpha coefficient for this sample because of the small number of participants completing the CSQ (n=9). In other studies, the internal consistency of the CSQ-8 is high, with alpha coefficients ranging from .84 to .93 for the CSQ [58]. The SEF contains 13 items that assess the feasibility and perceived utility of activities in the game. For example, the SEF states: “I will be able to apply what I learned from this game in my life” (for which the response options are 1=“Strongly agree”; 2=“Agree”; 3=“Disagree”; and 4=“Strongly disagree”) [57]. As the SEF asks about specific the utility of separate elements of the intervention, and not about singular constructs, alpha coefficients for the SEF are not commonly calculated.

**Data Analysis**

**Qualitative Data**

Trained research assistants transcribed verbatim the digital audio recordings of each interview. Then, the MD- or PhD-level research team member reviewed the transcripts with the digital recording for accuracy. Qualitative data analysis followed the tenets of thematic analysis, which consisted of sequential steps [59,60], and interviews continued until data saturation was achieved. The research team familiarized themselves with the data, reviewing each transcription. Next, the research team met weekly and generated a list of codes as they emerged. The team generated a thematic table of the analyses and checked the extent to which the emerging themes reflected the coded data [59,61]. The team grouped the themes under the general categories of the interview guide (ART information, ART motivation, ART behavioral skills, general game attitudes, and reactions to ViralCombat). Themes were examined in their relation to perceived utility of the game and for factors that would improve or detract from the game’s impact. Team discussion and interviews continued until discrepancies were resolved.

**Quantitative Data**

Participant responses on the CSQ and SEF were entered into an Excel file, and responses were verified with a second entry. Categorical response frequencies were calculated for each item of both scales. General acceptability of the intervention is illustrated using individual items from the scales. CSQ items are reported using the proportion of participants endorsing satisfaction with the intervention (response options “Very satisfied” and “Mostly satisfied” were combined). SEF items are reported using the proportion endorsing “agreement” with feasibility and utility of the game (response options “Strongly agree” and “Agree” were combined).

**Results**

**Reactions to ViralCombat Storyboard and iPhone Game**

Interviews from both the storyboard and iPhone game revealed a number of themes that guided game development. Participants desired informational game content that included new and comprehensive details about PrEP, details about PrEP as it relates to doctors’ visits, and general health information.
Motivational themes that emerged were the desire for enhancement of future orientation; reinforcement of positive influences from peers, partners, and friends; collaboration with health care providers; decreasing stigma; and increasing personal relevance of HIV prevention. Behavioral skills themes centered around self-efficacy and strategies for PrEP adherence and medical care (see Textbox 2 for qualitative interview themes and resulting game adaptations based on the IMB model).

Textbox 2 highlights the barriers and facilitators to adherence expressed by our participants and the corresponding gaming app action or message that was adapted to enhance facilitators or challenge barriers. Textbox 2 also includes general gaming attitudes that influenced the development of ViralCombat and specific reactions to the ViralCombat storyboard and iPhone game. In addition to the themes in Textbox 2, participants who played the game on the phone said that important gaming characteristics included directly destroying and fighting off HIV in game play, earning health points by taking pills that looked like PrEP, a prologue or introduction with a dramatic voice-over, and images about HIV and STIs that were relevant. Participants wanted levels that become increasingly difficult (for a sense of accomplishment). Participants did not want HIV or PrEP in the title of the game because of concerns about privacy and stigma but wanted to fight graphics labeled “HIV” and wanted to see the word PrEP during gaming. A 29-year-old black male participant said:

> I think there should be facts in the game about other health stuff, about smoking and drinking on PrEP.

Many participants also wanted more guidance through the levels. For example, a 29-year-old white male participant stated:

> I would like better hints on each level on how to beat the level.

An 18-year-old black male said (see Textbox 2):

> Clues or hints when it gets hard would make this better.

Monitoring Pill Bottle Opening and Game-Related Text Messages

We asked participants about the SMS text messages with gaming graphics and the use of a smart pill cap that measured adherence. During the interviews, we demonstrated how openings of the pill bottle were measured wirelessly, and we showed participants sample adherence-informed SMS text messages. When looking at the smart pill cap, an 18-year-old black male participant stated:

> It’s cool how it links with game. It’s awesome that there is a bottle that knows what you are doing.

The feedback about the cap was extremely positive; however, a 19-year-old black male participant stated:

> The pill cap is OK, I don’t love how big it is. People can see it in my bookbag more.

A 26-year-old white male described:

> It would be better if it was a bit smaller, but I like that it can record when I take pills, that helps me.

Although 2 participants had negative remarks about the size of the pill bottle, 18 out of 20 participants responded they liked the cap design and did not think the size of the cap was an impediment to use. During interviews, participants were also shown SMS text messages that corresponded to adherence data from the smart pill cap. Participants liked the proposed SMS text messages, and an 18-year-old black male described:

> These texts cue me to take my meds.

A 30-year-old black male stated:

> I like the game pictures. …the texts made me smile and also I felt like they were just for me.

A 35-year-old black male participant described:

> The texts got me in the mood to play the game again.

Of the participants, 3 described that they wanted more variation in the game-related texts:

> Seeing the same graphics again and again is boring. [35-year-old black male and a 19-year-old black male]

An 18-year-old black male stated:

> I would like more variety in the game texts.
**Textbox 2.** Qualitative interview themes and resulting game adaptations based on the information-motivation-behavioral skills (IMB) model.

<table>
<thead>
<tr>
<th>Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>• New and comprehensive details about pre-exposure prophylaxis (PrEP) and how PrEP prevents HIV</td>
</tr>
<tr>
<td>• Game includes complex and realistic information about PrEP. Participants fight off HIV in each organ. HIV is graphically represented. Facts about HIV and PrEP are imparted at every level. HIV is pictured.</td>
</tr>
<tr>
<td>• PrEP as it relates to doctors’ visits</td>
</tr>
<tr>
<td>• Terms and verbiage often used at PrEP-related doctor visits are used and defined in the game frequently. The importance of regular HIV testing and testing for other STIs is explained.</td>
</tr>
<tr>
<td>• General health information</td>
</tr>
<tr>
<td>• Participants learn about sexually transmitted infections that PrEP does not protect them from getting (gonorrhea, syphilis, herpes, genital warts, and human papillomavirus).</td>
</tr>
<tr>
<td>• Participants in the game receive messages about how exercise and healthy eating also affect health. Participants also receive messages about avoiding illicit substances and how illicit substances can increase risk behavior.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Motivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Enhancement of future orientation</td>
</tr>
<tr>
<td>• Messages about staying healthy for family, friends, and children scroll through game. As gaming participant takes more pills and builds more health, they are able to move through levels, receive more artillery, and have more success staying healthy.</td>
</tr>
<tr>
<td>• Personal relevance of HIV</td>
</tr>
<tr>
<td>• Participants are shrunken down to enter into their own body to fight HIV. During game play, participants see how PrEP works to prevent HIV infection.</td>
</tr>
<tr>
<td>• Collaborating with health care providers</td>
</tr>
<tr>
<td>• Throughout the game, participants partner with doctors to advance to the next level, build strength, and collect artillery.</td>
</tr>
<tr>
<td>• Reinforce influences from peers, partners, and friends</td>
</tr>
<tr>
<td>• Scrolling messages remind gamers that staying healthy for partners, friends, and family is meaningful for themselves and for the loved ones in their lives.</td>
</tr>
<tr>
<td>• Decrease stigma</td>
</tr>
<tr>
<td>• Participants are empowered to kill HIV before it enters the body and gain points with each healthy decision.</td>
</tr>
<tr>
<td>• Adherence to PrEP is valued as healthy and smart decision-making.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavioral skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Self-efficacy for PrEP adherence and PrEP-related medical care</td>
</tr>
<tr>
<td>• Solving problems and collecting pills or swallowing pills in the game leads to more points, which leads to more strength, more health, and more artillery. This leads to more game play. Perseverance throughout levels leads to success in the game.</td>
</tr>
<tr>
<td>• Strategies for medication adherence and self-care</td>
</tr>
<tr>
<td>• Scrolling messages encourage the participant to use PrEP, schedule routine doctors’ appointments, and ask providers or doctors questions about topics relevant to HIV prevention.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>General gaming attitudes</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Desire for games with levels, sound effects, and colorful graphics. Ability to earn points in the game and choose avatars</td>
</tr>
<tr>
<td>• Levels or organ systems become increasingly difficult (for a sense of accomplishment). Background music, sound effects, and dramatic voice-overs are included. Colorful graphics are included and change often. Choice of avatars is available. Participants earn points in the game by swallowing pills.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reactions to ViralCombat</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Desire for game action that is realistic with relevant information about PrEP and HIV prevention.</td>
</tr>
</tbody>
</table>
Acceptability and Feasibility

CSQ and SEF scores were available from participants who played the game on the iPhone for 45 to 50 min. The responses were generally quite positive, 88% (8/9) of the participants were satisfied with the activities in the game, 77% (7/9) felt game topics were interesting, 100% (9/9) felt they would recommend the game to a friend, 77% (7/9) felt game topics stimulated their interest in the material, 100% (9/9) felt that game topics were relevant to their lives, and 66% (6/9) felt they were able to do the activities in the game (see Table 1). The gaming intervention was improved based on the above acceptability and feasibility feedback from the CSQ and SEF and also on the feedback from the qualitative interviews (see Textbox 2). Specifically, game play was made easier with written messages and hints throughout each level on how to move forward. We also improved narrated instructions at the beginning of each level to assist players. General health facts about smoking, side effects of PrEP, and how substances such as drugs and alcohol can increase risky behavior were incorporated into the game to increase perceived relevance. We also incorporated more detailed information about HIV and how PrEP works to improve learning and interest. We also included a condom quiz game and information about other STIs (see Multimedia Appendices 7, 8, and 11). To improve SMS text messages, we included emojis and designed 5 different game-related SMS text messages utilizing a larger variety of game graphics based on participant feedback.

Table 1. iPhone game acceptability and feasibility scores (N=9).

<table>
<thead>
<tr>
<th>Questionnaire or form</th>
<th>Participants endorsing statement, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Client Service Questionnaire</strong></td>
<td></td>
</tr>
<tr>
<td>How would you rate the quality of the game?</td>
<td>8 (88)(^a)</td>
</tr>
<tr>
<td>Did you get desired information from the game?</td>
<td>6 (66)(^a)</td>
</tr>
<tr>
<td>To what extent does the game’s content meet your needs?</td>
<td>7 (77)(^a)</td>
</tr>
<tr>
<td>If a friend were interested in a similar program, would you recommend our game to him or her?</td>
<td>9 (100)(^a)</td>
</tr>
<tr>
<td>How satisfied are you with the amount of information you have received in the game?</td>
<td>7 (77)(^a)</td>
</tr>
<tr>
<td>Has the information you received helped you to deal more effectively with issues important to you?</td>
<td>8 (88)(^a)</td>
</tr>
<tr>
<td>In an overall, general sense, how satisfied are you with the game?</td>
<td>8 (88)(^a)</td>
</tr>
<tr>
<td>Would you come back to the game again?</td>
<td>9 (100)(^a)</td>
</tr>
<tr>
<td><strong>Session evaluation form</strong></td>
<td></td>
</tr>
<tr>
<td>I learned a lot from the game.</td>
<td>7 (77)(^b)</td>
</tr>
<tr>
<td>I will be able to apply what I learned from the game in my life.</td>
<td>6 (66)(^b)</td>
</tr>
<tr>
<td>I was able to do the activities in the game.</td>
<td>8 (88)(^b)</td>
</tr>
<tr>
<td>The game was well organized.</td>
<td>7 (77)(^b)</td>
</tr>
<tr>
<td>The topic of the game was interesting.</td>
<td>7 (77)(^b)</td>
</tr>
<tr>
<td>The presentation of the information stimulated my interest in the material.</td>
<td>9 (100)(^b)</td>
</tr>
<tr>
<td>The topics of the game were relevant to my life.</td>
<td>8 (88)(^b)</td>
</tr>
<tr>
<td>The game was enjoyable.</td>
<td>8 (88)(^b)</td>
</tr>
<tr>
<td>I would recommend the game to others.</td>
<td>9 (100)(^b)</td>
</tr>
<tr>
<td>I felt comfortable during game play.</td>
<td>9 (100)(^b)</td>
</tr>
</tbody>
</table>

\(^a\) Proportion of participants that endorsed “Very satisfied” and “Mostly satisfied.”

\(^b\) Proportion of participants that endorsed “Strongly agree” and “Agree.”
Discussion

Principal Findings

In this project, we utilized in-depth interviewing, focused by social learning theory (IMB), to create an iPhone gaming intervention to measure and improve treatment PrEP adherence and decrease HIV risk for YMSM [46-48,53,62]. A number of themes emerged through qualitative interviews with young men that informed game development. We found that YMSM desired informational game content that included comprehensive details about how HIV is transmitted, PrEP-related doctors’ visits, side effects of PrEP, and general health information. Motivational themes or findings that emerged were the desire for enhancement of future orientation; the need for reinforcement of positive influences from peers, partners, and friends; and the promotion of collaboration with health care providers. Motivational themes also included decreasing stigma and increasing personal relevance of HIV prevention. Behavioral skills themes or findings centered around self-efficacy and strategies for PrEP medication and appointment adherence including regular HIV and STI testing.

Using the IMB theory in the development of this game ensured that the intervention was informed by decades of prevention research. This study demonstrates that qualitative assessment, social learning theory, and agile software development can complement each other and are important components to the development of a tailored and clinically relevant app. Participant data were used throughout the development of the game and informed the informational, motivational, and behavioral skill-building components of the game. Using a storyboard provided the research team with opportunities to share concept models with participants early on in the design process and gather feedback with respect to necessary modifications. Sharing the iPhone game with participants as it was developed also allowed for necessary, incremental improvements. YMSM at risk of HIV infection provided key qualitative insights with respect to the content and design of the game. Tailored games that are informed by those who will use them have more potential for effective integration and uptake in clinical settings [41].

Although iPhone games are pervasive in popular culture, no other gaming apps have been developed for YMSM on PrEP. Findings of this study highlight several important barriers and facilitators to PREP adherence for YMSM. Mobile interventions have the potential to reinforce skills learned in the clinic and require fewer resources to deliver patient-centered, evidence-based interventions [63]. Furthermore, apps and mobile phone games have the potential to engage YMSM in interventions, who otherwise may not be willing or able to participate in intervention programs.

Gaming and mobile apps also have the potential to advance the delivery of information and promote healthy decision-making in disproportionately affected populations, including minority men who often have less access to medical care and support [60]. National data from the Pew Research Center indicate that younger, ethnic and racial minority populations use mobile phones frequently [64]. The adolescents and young adults in this study repeatedly expressed having access to, and familiarity with, iPhones. This widespread use of iPhones facilitates the uptake of gaming apps in clinical populations. Therefore, mobile technologies such as mobile phone games and apps have a great potential to enhance medical care and prevention interventions for populations who are disproportionately affected by HIV and other STIs.

Limitations

Findings should be interpreted in light of study limitations. First, our participants were recruited from a PrEP clinic in Jackson, Mississippi. This clinic may not be representative of all PrEP clinics in the United States or internationally. Therefore, the generalizability of the data collected to inform the development of the app is unknown and may be limited. Second, this study focused on individual PrEP user perspectives. It may be equally important to integrate the perspectives of clinicians, friends, or partners into the game. In the future, including friends and social networks into gaming prevention programs could be novel and effective. Perspectives of partners, friends, and clinicians could also lead to a more robust understanding of barriers and facilitators to adherence to PrEP and PrEP-related medical care. Therefore, future research could examine the utility of integrating feedback from clinicians and friends into the gaming app. Finally, this app was developed for the iPhone. Development of the app for Android devices could allow for greater availability of the game and could be a forthcoming step in the future phases of research.

Conclusions

This study is a significant step in the development and testing of an iPhone gaming app to promote adherence to PrEP and PrEP-related medical care. The long-term goal of this research program is to test ViralCombat in a randomized trial and examine if the game is effective in changing PrEP-related attitudes and adherence-related behaviors. If the results appear promising, the research team can distribute the technology procedures for the intervention to relevant and interested clinics, community-based organizations, the Centers for AIDS Research, and AIDS Trials Networks. The app could also be made available on iTunes and Google Play (Android) for a nominal fee.

There are many advantages to using newer interactive technology to improve adherence rather than traditional face-to-face counseling, including scalability, efficiency, and cost-effectiveness. As electronic games are highly appealing to young men [25], they are a natural opportunity to deliver health education during leisure time and outside of the clinic [25-27,52]. Games can attract and maintain attention, which is a key component for effective behavior change. Compelling interactive games can expose players to essential health-related content thousands of times and also give players unlimited opportunities to rehearse new skills and receive personalized feedback on health choices made within the game [28,35]. We are not aware of other adherence interventions that integrate medication adherence monitoring technology, SMS text messaging, and a theoretically informed game to improve information, motivation, and behavioral skills for PrEP adherence. An intervention with these components may
empower and engage YMSM, aid clinics, and result in improvements in health.

Acknowledgments
This publication was made possible with help from the Providence-Boston Center for AIDS Research (P30 AI 042853). The project described was supported by grant number R34MH104068 (MPI: LW, LKB) from the National Institute of Child Health and Human Development. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health.

Conflicts of Interest
LM is a member of the Gilead Sciences Advisory Board and receives compensation for consulting from Gilead Sciences, Inc. No competing financial interests exist for the remaining authors.

Multimedia Appendix 1
ViralCombat main menu graphic.

[PDF File (Adobe PDF File), 66KB - games_v7i1e11861_app1.pdf]

Multimedia Appendix 2
Example of a short message service text message to the participants.

[PDF File (Adobe PDF File), 62KB - games_v7i1e11861_app2.pdf]

Multimedia Appendix 3
ViralCombat game menu.

[PDF File (Adobe PDF File), 37KB - games_v7i1e11861_app3.pdf]

Multimedia Appendix 4
Short narrative movie at the beginning of the game.

[PDF File (Adobe PDF File), 57KB - games_v7i1e11861_app4.pdf]

Multimedia Appendix 5
Players can design and individualize their game character.

[PDF File (Adobe PDF File), 47KB - games_v7i1e11861_app5.pdf]

Multimedia Appendix 6
Players are shrunken down to be able to enter the body to fight off HIV.

[PDF File (Adobe PDF File), 29KB - games_v7i1e11861_app6.pdf]

Multimedia Appendix 7
Answering questions with allied doctors and building knowledge help each player successfully move to the next level or area of the body. Example of questions answered incorrectly.

[PDF File (Adobe PDF File), 46KB - games_v7i1e11861_app7.pdf]

Multimedia Appendix 8
Answering questions with allied doctors and building knowledge help each player successfully move to the next level or area of the body. Example of questions answered correctly.

[PDF File (Adobe PDF File), 42KB - games_v7i1e11861_app8.pdf]
Multimedia Appendix 9
Example of a player gaining strength in the anus level by collecting health pills and condoms.

[PDF File (Adobe PDF File), 23KB - games_v7i1e11861_app9.pdf]

Multimedia Appendix 10
As players travel through the bloodstream, they must fight off viruses and gain health pills.

[PDF File (Adobe PDF File), 51KB - games_v7i1e11861_app10.pdf]

Multimedia Appendix 11
Throughout ViralCombat, players engage in various games in order to move on to the next level. In the Condom Sequence game, players must go through the correct sequence of steps for correctly using a condom.

[PDF File (Adobe PDF File), 59KB - games_v7i1e11861_app11.pdf]

Multimedia Appendix 12
Players prevent HIV and other sexually transmitted infections in the penis by killing viruses and other bacteria.

[PDF File (Adobe PDF File), 54KB - games_v7i1e11861_app12.pdf]

Multimedia Appendix 13
Summary of points earned at the end of each level.

[PDF File (Adobe PDF File), 53KB - games_v7i1e11861_app13.pdf]

References


Abbreviations

ART: antiretroviral treatment
CSQ: Client Service Questionnaire
IMB: information-motivation-behavioral skills
IRB: institutional review board
MSM: men who have sex with men
PrEP: pre-exposure prophylaxis
SEF: session evaluation form
SMS: short message service
STI: sexually transmitted infection
TMP/SMX: trimethoprim-sulfamethoxazole
YMSM: young men who have sex with men

Edited by G Eysenbach; submitted 23.08.18; peer-reviewed by J Vandenberg, L Hightow-Weidman; comments to author 12.10.18; revised version received 05.12.18; accepted 09.12.18; published 27.03.19

Please cite as:
Whiteley L, Mena L, Craker LK, Healy MG, Brown LK
Creating a Theoretically Grounded Gaming App to Increase Adherence to Pre-Exposure Prophylaxis: Lessons From the Development of the Viral Combat Mobile Phone Game
JMIR Serious Games 2019;7(1):e11861
URL: http://games.jmir.org/2019/1/e11861/
doi:10.2196/11861
PMID:

©Laura Whiteley, Leandro Mena, Lacey K Craker, Meredith Garver Healy, Larry K Brown. Originally published in JMIR Serious Games (http://games.jmir.org), 27.03.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on http://games.jmir.org, as well as this copyright and license information must be included.