Serious Gaming in Medical Education

A Proposed Structured Framework for Game Development

Aleksandra E. Olszewski, MD;
Traci A. Wolbrink, MD, MPH

Abstract: Serious games are increasingly being used for medical education. However, the design and development of serious games for the education of health professionals is highly variable, and very few articles report the development process used for game development. There are many established processes for software development that can improve and streamline development, and incorporating the best practices from educational pedagogy and software development may enhance teamwork and communication, decrease development costs, and improve the quality of serious games. In this article, we review and summarize the literature for serious game development for medical education, and combining the best practices, we propose a structured three-phase iterative development framework for serious game development.
(*Sim Healthcare* 00:00–00, 2016)

Key Words: Game development, formative evaluation, medical education, serious gaming, virtual simulation, peritoneal dialysis.

Innovative online learning tools, including serious games, are increasingly being used for medical education. Despite the increasing use of serious games, the design and development of serious games for education of health professionals are highly variable.[1–5] Many studies describe clinical and educational frameworks for their serious game development, but the technical aspects of development are inconsistent and rarely reported.[5–10] A recent serious gaming review concludes that for serious gaming to continue growth within medicine, development, evaluation, and distribution frameworks need to be constructed.[1]

Serious games are games developed for a purpose other than entertainment, such as teaching a specific knowledge or skill.[5] Wang et al[1] describe that to be labelled a serious game, an activity must also include challenging goals, scoring, and an engaging design. There are many types of serious games, and Wang et al[1] recently classified serious game types as the following: adaptation, adventure, board game, management simulation, platform, puzzle, quiz, and virtual simulation. Some of these games types are intended to be delivered during face-to-face encounters, and others are delivered virtually. In this study, we report on only virtual serious games. Virtual simulation, a specific and sophisticated technology that incorporates the imitation of the operation of a real-world process, may also be characterized a serious game if the simulator includes gaming elements as stated previously.[1,2] Not all virtual simulators include gaming components, and thus, not all virtual simulators are also considered serious games.

Serious games have been shown to increase learner satisfaction and knowledge gains over traditional teaching methodologies.[8,9] Serious games provide a scalable, convenient method for learners to practice skills in a safe environment while incorporating interactivity and competition in a format well liked by millennial learners.[2,5] The gaming aspect introduces motivational factors and cognitive scaffolds to promote learning and to keep learners motivated and engaged.[5–7] Adult learning theory principles are optimized through self-pacing and capacity for repetition, learner-controlled and real-time feedback, and accessibility to education when it is convenient and clinically useful.[2–5,7] Automated scoring and action-specific feedback allow knowledge gains while decreasing demand on educators.

Despite the previous evidence and theory supporting serious game, the development of games for medical education is complex, and there are many deterrents to their widespread development.[5–10] Game development requires expertise in medicine, education, and technology development to create the complex design, modeling, and scoring required to make an effective game. Clinician teams often partner with software development teams to create serious games, each team bringing a different skill set and working language. Limited medical knowledge by the developer, poor transfer of medical concepts and management strategies, and misaligned expectations for game scope and functionality between clinicians and developers create important challenges.[6–10] In addition, game development can be costly and require lengthy development time. Established processes for software development can be formally applied to improve and streamline serious game development.[8–15] Incorporating the best practices from educational pedagogy and

From the Department of Pediatrics (A.E.O.), Seattle Children's Hospital, University of Washington, Seattle, WA; Harvard Medical School (T.A.W.); Division of Critical Care Medicine (T.A.W.), Department of Anesthesiology, Perioperative and Pain Medicine, Boston Children's Hospital, Boston, MA.

Reprints: Traci A. Wolbrink, MD, MPH, Boston Children's Hospital, 300 Longwood Ave, Bader 634, Boston, MA 02115 (e-mail: traci.wolbrink@childrens.harvard.edu).

A technology development grant given by the Boston Children's Hospital Technology and Innovation Development Office was used to support the development of the simulator, which also included generously matched funding from the Department of Anesthesiology, Perioperative and Pain Medicine, Boston Children's Hospital.

The authors declare no conflict of interest.

DOI: 10.1097/SIH.0000000000000212

software development may enhance teamwork and communication, decrease development costs, and improve the quality of serious games.[8–15]

In this article, we review and summarize the literature for serious game development for medical education, and using the best practices, we propose a structured development framework for serious game development. We will describe the development of a training simulation game, a virtual peritoneal dialysis (PD) simulator, to illustrate use of the proposed framework including design, development, and formative evaluation. The PD simulator is available for free use at https://www.openpediatrics.org/assets/simulator/peritoneal-dialysis-simulator, on OPENPediatrics, a global medical education Web site.[16]

METHODS

We performed a PubMed search from January 2000 through July 2016 using the following search terms: "serious gam*," OR "gaming," OR "virtual simulat*," OR "video game," AND "healthcare," OR "education." Two independent reviewers reviewed all abstracts to select appropriate articles. In the event of disagreement, both reviewers reviewed the full article, and conflicts were resolved by consensus. In addition, we searched reference lists of relevant articles. We included reviews, meta-analyses, and original articles reporting results of games designed to teach knowledge, behavior, or skills to healthcare providers, excluding augmented reality and surgical simulator tools that involved devices not commercially included in a standard computer or gaming console (ie, laparoscopic surgical trainers). We excluded articles describing serious gaming used for education of patients, K-12 students, or nonmedical fields. Figure 1 depicts our search strategy. We assessed each article for inclusion of details about the development process used in game development. Evaluation strategies were considered formative if the feedback results were used to assess usability and/or the authors used other methods to improve the game itself, not just a study of satisfaction with a finished product. We classified the degree to which the details were reported as either:

1. None: no development process described; only the name of software program used or development team reported; and/or the developer used already existing software to create new cases.

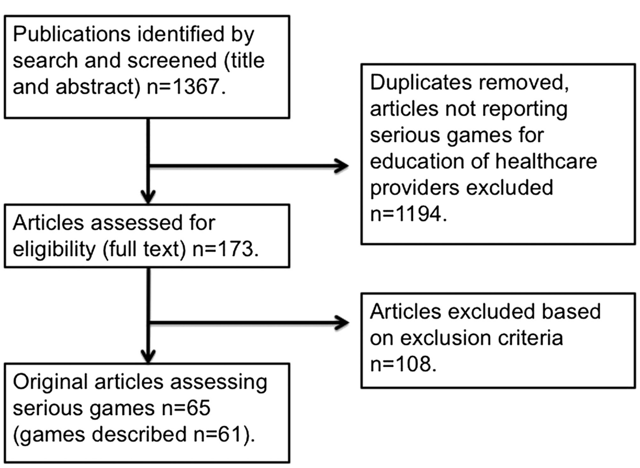

FIGURE 1. This figure outlines the search strategy and findings from the literature review.

2. Brief: development process described broadly; would be unable to replicate the development process with the reported information
3. Detailed: development process described in detail; would be likely to completely or partially replicate these processes, authors did not need publish the exact instruments or questionnaires.

This project and its component analyses were approved for exemption by the Boston Children's Hospital institutional review board.

RESULTS

A total of 1367 articles were identified, and after excluding those on the basis of our exclusion criteria, we included 65 original articles describing 61 serious games designed for teaching medical knowledge, behaviors, or skills.[8,9,17–79] Table 1 details the development process used for game development for each of the 61 serious games. Very few games described the development process in detail. Of those that reported any information about the development process, seven described an iterative development process, nine reported soliciting verbal feedback or using focus groups, and ten reported generating feedback obtained by questionnaires.

The two studies that went into most detail about the development were those by Davids et al[8] and Diehl et al.[36] Davids et al[8] describe a development process to create an acid-base game using wireframes, iterative prototyping, and a questionnaire containing a System Usability Scale (SUS) and qualitative feedback. The SUS is very frequently used across fields to assess a product's usability.[15] The SUS includes ten questions, each scored from 1 to 10, with responses summed to report a single number (range, 0–100), with a score of 70 or higher signifying acceptable product usability.[15] Diehl et al[36] describe a development process to create a diabetes virtual patient game using iterative development with periodic team meetings, structured usability testing followed by game modifications, and formal beta testing. During the structured usability testing, participants were video recorded using Think Aloud testing. Think Aloud Protocol is a validated tool where participants are observed thinking aloud while performing tasks and interacting with the game.[14] It provides a structured technique for qualitative data acquisition and analysis, helping reveal the thought processes of the learner and allowing for identification of usability issues.[14] Speech, facial expressions, and game actions were recorded using usability assessment software. At least two reviewers analyzed each video recording, using a standardized classification system to code all events observed in the videos. Events were classified as the following: system or user events; as negative, positive, or neutral; and as mild, moderate, or critical. Usability testing included the SUS survey, Likert scale questions, and open-ended questions. Beta testing allowed users to access the game at their own pace on their own computer for 15 days, and Google analytics monitored user activity in the game. After this, the SUS survey, Likert scale questions, and open-ended questions were administered.

Proposed Framework for Game Development

We propose a three-phase development and formative evaluation process on the basis of the results from our review (Fig. 2). We will use the development of a PD simulator to illustrate the process.

TABLE 1. The Results of Our Serious Game Review, Organized by Author and Publication Year, With Game Development Process Described. The Degree of Detail of the Description of Development and Evaluation was Classified According to Three Categories (None, Brief, Detailed)

				Processes Used for Development and/or Formative Evaluation						
Authors and Publication Year	Serious Game Title	Description of Development (None, Brief, Detailed)	Description of Formative Evaluation (None, Brief, Detailed)	None Described	Reported Name of Development Team	Reported Software Program Used	Iterative Design Process	Focus Group or Verbal Discussion	Usability Questionnaire	Detailed Information
Akl et al (2008)[17]	Guide-O-Game	Brief	Detailed	—	X	—	X	X	—	Iterative improvement process including pilot testing, feedback, and revision. Pilot tested to evaluate feasibility with four weekly sessions (session 45 m and feedback 15 m) with feedback including writing open-ended questions (used a nonstandardized nonvalidated measurement tool) followed by an open verbal group discussion
Amer et al (2011)[18]	Interactive Dental Video Game	None	None	—	—	X	—	—	—	—
Andreatta et al (2010)[19]	CAVE Triage Training	None	None	X	—	—	—	—	—	—
Antoniou et al (2014)[20]	SL Periodontology Virtual Patient	None	Detailed	—	—	X	—	X	—	Nine dental students used two different case environments and focus groups explored positives and negatives to each format.
Atack et al (2009)[21]	The Disaster Management Course	None	None	X	—	—	—	—	—	—
Bergeron et al (2008)[22]	Radiation Hazards Assessment Challenge	None	None	X	—	—	—	—	—	—
Bergeron et al (2008)[22]	Nuclear Event Triage Challenge	None	None	X	—	—	—	—	—	—
Boeker et al (2013)[23]	Uro-Island	None	None	—	X	X	—	—	—	—
Buttussi et al (2013)[24]	EMSAVE	Brief	Detailed	—	—	X	X	—	—	Pilot study performed using a prototype of game. Twelve emergency nurses evaluated the game's usability and perceived usefulness and acceptance. An iterative prototyping phase proceeded with informal evaluation by advanced life support providers and instructors.
Cavazza et al (2013)[25]	Interactive Storytelling	None	None	—	—	X	—	—	—	—
Cowan et al (2011)[26]	Off-pump Coronary Artery Bypass Game	None	None	X	—	—	—	—	—	—
Cowan et al (2010)[26]	TKA Game	Brief	Detailed	—	—	X	X	X	X	Iterative test-and-design methodology was used to develop game. Usability testing conducted by seven game development students and four orthopedic surgery residents in two focus groups. The testers freely explored game followed by completion of questionnaire for user interaction satisfaction and several open-ended questions.

Continued next page

TABLE 1. *(Continued)*

Authors and Publication Year	Serious Game Title	Description of Development (None, Brief, Detailed)	Description of Formative Evaluation (None, Brief, Detailed)	Processes Used for Development and/or Formative Evaluation: None Described	Reported Name of Development Team	Reported Software Program Used	Iterative Design Process	Focus Group or Verbal Discussion	Usability Questionnaire	Detailed Information
Craig et al[27] (2014)	V-PIN	None	None	—	—	X	—	—	—	Course faculty used already developed software to create branched narrative cases including scoring system.
Creutzfeldt et al (2008, 2010, 2012, 2014)[28–31]	Multiplayer Virtual World-CPR Game	None	None	—	—	X	—	—	—	Used an existing multiplayer environment to create cases.
Davids et al (2011)[8]	Electrolyte Workshop	Detailed	Detailed	—	—	X	X	—	X	Iterative design process used to develop game. Wireframes used to convey information between development and medical teams. Two iterations of wireframes were conducted and then software built. Three rounds of iterative testing conducted with prototype. User testing conducted by ten residents and six specialists in internal medicine, medicine, nephrology, and endocrinology. Users worked through game at own pace and at the end of session completed a satisfaction questionnaire on the basis of the SUS.
del Cura-González et al (2016)[32]	e-Learning (EDUCAGUIA)	None	None	—	X	—	—	—	—	—
De Leo et al (2014)[79]	GaMeTT	None	None	—	—	X	—	—	—	Used an existing multiplayer environment to create cases.
Delasobera et al (2010)[33]	Cardiac Arrest!	None	None	—	—	—	—	—	—	Used a previously developed game for the study.
Dev et al (2011)[34]	CliniSpace	Brief	Brief	—	—	—	—	—	—	Used feedback from previous studies of own game development to guide design of new game. No reports of specifics about how the feedback was previously collected. Incorporated the feedback into a structured framework that evaluated seven components of new technologies for medical learning.

Diehl et al (2013, 2015)[35,36]	InsuOnline	Brief	Detailed	—	X	X	X	—	X	Iterative development used with software testing and periodic team meetings, but specific details not mentioned. A pilot study of prototype to assess usability and playability. Alone and in single session, under concealed observation by researchers. Think Aloud protocol used. Speaking, facial expressions, and game actions were recorded in video using usability assessment software. Dual reviewers analyzed tapes and completed a standardized sheet to record and to classify all events observed in the videos. Events classified as system or user events, and as negative, positive, or neutral, and then as mild, moderate, or critical. Then users completed system usability scale tool, Likert scale questions, and open-ended questions. Critical corrections were made, and beta testing conducted. Beta testing included allowing users to access game at their own pace on own computer for 15 days. The same surveys were completed, and Google analytics were used to monitor user activity in game.
Duque et al (2008)[37]	Riskdom-Geriatrics	None	None	—	X	X	—	—	—	—
Evans et al (2014)[38]	Septris	None	Brief	—	X	X	—	—	—	Pilot study with five medical students and seventeen internal medicine residents was conducted to elicit feedback to improve game and scoring. Completed quizzes to assess medical knowledge and gave feedback on the game, but specifics of how and what game feedback were collected was not completely detailed.
Fonseca et al (2014, 2015)[39,40]	e-Baby	Brief	None	—	—	X	—	—	—	Describes the use of a structured framework of seven phases of model development: (1) the subject choice and motivational factors of the project, (2) evaluation and analysis of the user's needs, (3) solution identification, (4) goal articulation, (5) question analysis, (6) development and prototyping, and (7) user evaluation. Specifics of how these were implemented not completely detailed.
Foster et al (2009)[41]	3D Renal Care Learning Environment	Brief	None	—	—	X	—	—	—	Reports a scope for development for a future game.

Continued next page

TABLE 1. *(Continued)*

Authors and Publication Year	Serious Game Title	Description of Development (None, Brief, Detailed)	Description of Formative Evaluation (None, Brief, Detailed)	Processes Used for Development and/or Formative Evaluation						
				None Described	Reported Name of Development Team	Reported Software Program Used	Iterative Design Process	Focus Group or Verbal Discussion	Usability Questionnaire	Detailed Information
Gleason et al (2015)[9]	Research and Evidence Learning in Medicine	Brief	Brief	—	X	X	—	—	—	Described development in three broad phases: prototype with informal feedback from convenience sample at a conference where game was demonstrated; further case development with a clinician was brought in to guide medicine; and an experienced game developer and graphic designer completed game development. Informal usability testing by team directed improvements to game. Formal usability testing to be completed in future.
Graafland et al (2014)[43]	Medialis	None	Brief	—	—	—	—	—	X	Prototype was assessed fluency, reading time, and clarity of imaging, but details on this testing are not given. Questionnaire was given with questions on demographic characteristics, realism, educational and testing value, perceived desirability and preferred user groups, perceived user experience, and game implementation, but did not report whether this feedback was used to modify game.
Graafland et al (2013, 2014)[42,44]	Laparoscopic Equipment Failure Serious Game	None	None	X	—	—	—	—	—	—
Hannig et al (2012)[45]	eMedOffice	Brief	Detailed	—	X	X	X	X	X	Described an abstract model used to inform game development. Iterative development used to refine game steps. Summative evaluation from forty-one medical students in an elective that were required to use the game included a usability questionnaire, verbal feedback, and direct observations from an educational supervisor.
Huguet et al (2016)[46]	VICTEAMS	None	None	X	—	—	—	—	—	—
Johnsen et al (2016)[47]	Teaching Nursing Students Clinical Reasoning and Decision-Making Skills	Brief	Brief	—	—	—	—	X	X	Described the theories used to development the game very broadly. Six participants evaluated the prototype in usability laboratory, with usability evaluation methods including cognitive walkthrough, questionnaire, and individual interviews.
Juanes et al (2016)[48]	Anatomical-Ultrasound Visor for Regional Anaesthesia	None	None	—	—	X	—	—	—	Evaluation by eigthy anesthesiologists but did not report whether this feedback was used to modify game.

Kanthan and Senger (2011)[49]	The Path Is Right	None	None	—	X	—	—	—	—	—
Kanthan and Senger (2011)[49]	Path to Success	None	None	—	X	—	—	—	—	—
Katz and Demaria (2013)[50]	Central Venous Catheter Game	None	None	—	X	X	—	—	—	Report informal feedback on game and next step is for validation and revision phase, with no specific details about the feedback and revision given
Kerfoot et al (2012, 2014)[51–53]	Spaced-Education Game	None	None	X	—	—	—	—	—	—
Kizakevich et al (2007)[54]	Sim-Patient	None	None	—	—	—	—	—	—	Modified an existing software program to develop game.
Knight et al (2010)[55]	Triage Trainer	None	None	—	X	—	—	—	—	—
Kurenov et al (2009)[56]	Burn Center	None	None	—	—	X	—	—	—	—
Lagro et al (2014)[57]	GeriatriX	None	None	X	—	—	—	—	—	—
LeRoy Heinrichs et al (2008)[58]	Virtual ED II	None	Brief	—	X	X	—	X	X	Pilot study of game conducted with six nurses and seven physicians. After using the software, they completed a brief user satisfaction survey and participated in a focus group. Did not report whether this feedback was used to modify game.
LeRoy Heinrichs et al (2008)[58]	Peninsula City	None	Brief	—	—	—	—	X	X	Pilot study of game conducted with eight paramedics/emergency medical technicians and eight emergency medicine physicians. After using game, completed brief user satisfaction survey and participated in a focus group. Did not report whether this feedback was used to modify game.
Lim et al (2008)[59]	eNLG	None	None	X	—	—	—	—	—	—
Lin et al (2015)[60]	SICKO	None	None	—	X	—	—	—	—	—
Makransky et al (2016)[61]	Virtual Learning Environment in Medical Genetics Counseling	None	None	—	X	—		—	—	—
McKenzie (2013)[62]	Game Informed Online Learning About Managing Aggression in Health Settings	None	Brief	—	—	X	—	—	—	Reported that the content was reviewed by an external board, and the game was piloted on three providers. No details about what process was used to acquire feedback from providers. Edits to game made on the basis of pilot feedback.
Mohan et al (2014)[63]	ED Physician Decision Making	Brief	Brief	—	X	—	X	—	—	Reports iterative testing of game elements and beta-testing of a prototype by thirty emergency medicine clinicians. No details about what process was used to acquire feedback from beta-testers.

Continued next page

TABLE 1. *(Continued)*

Authors and Publication Year	Serious Game Title	Description of Development (None, Brief, Detailed)	Description of Formative Evaluation (None, Brief, Detailed)	Processes Used for Development and/or Formative Evaluation: None Described	Reported Name of Development Team	Reported Software Program Used	Iterative Design Process	Focus Group or Verbal Discussion	Usability Questionnaire	Detailed Information
Nosek et al (2007)[64]	Cancer Genetics Tower	None	Detailed	—	X	X	—	—	X	Described the pilot testing used for the game: an anonymous online survey that included rating of the game elements and capability for text-based responses.
O'Neill et al (2012)[65]	GRAPHIC	None	None	—	—	X	—	—	—	—
Rondon et al (2013)[66]	Anatesse 2.0 Quiz	None	None	—	—	—	—	—	—	Used a previously developed quiz game.
Roy et al (2006)[67]	Virtual Standardized Patients	None	None	—	—	X	—	—	—	Cases developed in a previously created software program.
Schwaab et al (2011)[68]	Second Life Case-Based Mock Emergency Medicine Board Examination	None	None	—	—	X	—	—	—	Cases developed in a previously created software program.
Schwarz et al (2013)[69]	MEFANET	None	None	—	X	X	—	—	—	—
Schwid et al (2001)[70]	Anesthesia Simulator 3.0	None	None	—	—	—	—	—	—	Used a previously developed simulator.
Shewaga et al (2013)[71]	Z-DOC	None	None	X	—	—	—	—	—	—
Sullivan et al (2016)[72]	Virtual Surgical Patient Cases	None	Brief	—	—	X	—	X	—	Cases developed in a previously created software program. Focus groups with groups of twenty to thirty students and also faculty instructors gave feedback on their experience with the software and case scenarios. Did not report whether this feedback was used to modify game.
Sward et al (2008)[73]	Pediatric Rotation Game	None	None	—	—	—	—	—	—	Used a previously developed game.
Tolsgaard et al (2016)[74]	Virtual Patient Cases	None	None	—	—	X	—	—	—	Cases developed in a previously created software program.
Weiner et al (2016)[75]	Web-SP Seminar for oral Surgery	None	None	—	—	X	—	—	—	Cases developed in a previously created software program.
Wilkening et al (2016)[76]	Branched-Narrative Virtual Patients for Interprofessional Education	None	Brief	—	—	X	—	X	—	Cases developed in a previously created software program. Psychiatrists and psychiatric pharmacists, reviewed each case before resident administration. One-hour debriefing period was conducted and included summative satisfaction evaluation. Did not report whether this feedback was used to modify game.

Youngblood et al (2008)[77]	Virtual ED	None	None	—	—	X	—	—	—	—
Zaveri et al (2016)[78]	Virtual Reality Sedation Environment	None	Brief	—	X	X	—	—	X	Cases developed in a previously created software program. Thirty-two residents gave feedback on games via usability survey. Did not report whether this feedback was used to modify game.
Total no. studies	61	—	—	11	19	33	7	9	10	—

None indicates no development process described, only the name of software program or development team reported, and/or the developer used already existing software to create new cases; brief, development process described broadly, would be unable to replicate the development process with the reported information; detailed, development process described in detail, would be likely to completely or partially replicate these processes, authors did not need publish the exact instruments or questionnaires.
ED indicates emergency department; eNLG, online Neurological Localisation Game; GaMeTT, Game Medical Team Training; GRAPHIC, Games Research Applied to Public Health with Innovative Collaboration; MEFANET, Medical Faculties Network; SICKO, Surgical Improvement of Clinical Knowledge Ops; SL, Second Life; TKA, Total Knee Arthroscopy; VICTEAMS, Virtual Characters for Team Training: Emotional, Adaptive, Motivated and Social; V-PIN, Virtual Pathology Instructor.

Preparation and Design

Team Assembly

Identify and fund the appropriate members needed to develop the technical components (game developers) and medical content [subject matter experts (SMEs) and end users]. A development schedule including frequent meetings should be agreed on. We created a team of game developers, medical animators, and SMEs including three pediatric nephrologists (from broad range of practice backgrounds including academic, private practice, and a low-resource environment), a pediatric intensivist, a pediatric resident, and a medical student to develop the PD simulator. Scheduled biweekly meetings promoted accountability and required regular feedback between teams.

Medical Concepts Transfer

Because one of the key challenges described in medical game development is difficulty conveying medical concepts to nonmedical developers, medical concepts transfer helps orient the game developers to the medical information to be shared or processes to be used in the game. The SMEs demonstrated how to perform PD on a mannequin simulating the following: PD machine setup, initial prescription selection, dialysate adjustments, troubleshooting common problems, and clinical assessment of the patient. The game developers asked questions and took notes and photographs for reference.

Content Production

Medical content, including any necessary physiological modeling, is developed and approved by SMEs with expertise in the field, ideally on the basis of an appropriate instructional design for the type of game and applying validated or expert-derived guidelines or recommendations. The text, short problems (tactics), and cases were written incorporating existing evidence, guidelines, and expert-derived algorithms. The content was felt to reflect current standards of care and was reviewed and approved by each of the SMEs. Patient animations were created. Content was delivered to the game developers.

Learner Experience Mapping

The game developer and medical team discuss and collaboratively determine the game functionality, flow, feedback, and scoring on the basis of the game theory and proposed game modeling. Using the developed medical content, storyboards can be created to describe the flow and required functionality of the game. Over several design meetings, the team discussed and collectively decided on the hospital room layout, tabs necessary for interacting with the clinical components (such as patient, monitor, and chart) in a clinically realistic manner, device interface, flow, and functionality.

Development

Wireframes

Wireframes are illustrations of proposed game components and assist in visual communication and design of the structure, functionality, learner interface, and positioning of an application. The game developer created a wireframe to convey the proposed composition and functionality for the simulator, incorporating learner experience from the previous step. After review by the entire team, revisions were made, and a clickable wireframe was drafted, allowing for testing of functionality and learner experience (Fig. 3).

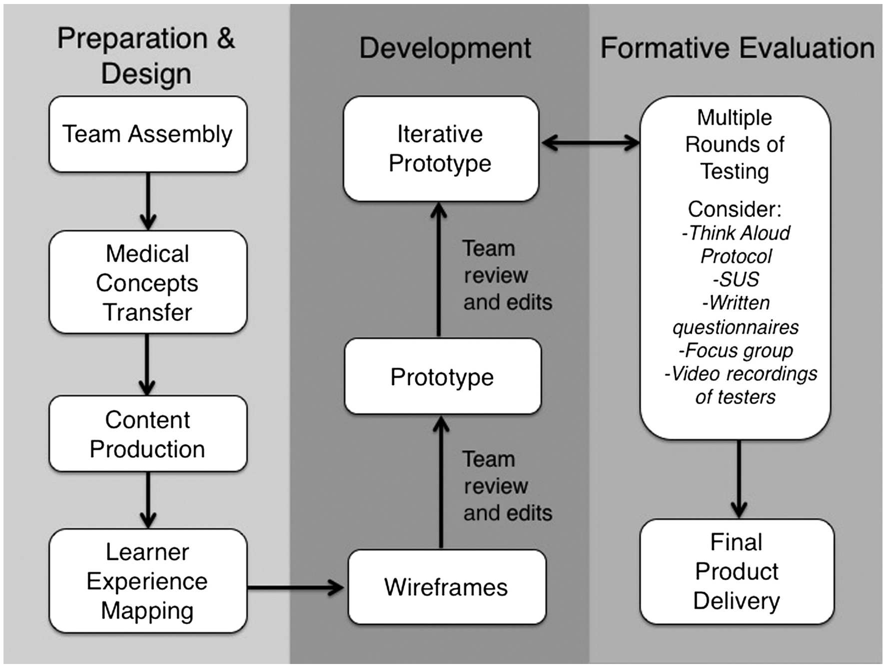

FIGURE 2. The development process, including design, development, and iterative formative evaluation steps.

Prototype

Software prototyping involves producing a series of partial systems early in the development cycle to facilitate team discussions about design and functionality allowing changes to be made easily to the prototype. A fully interactive and styled prototype of the PD simulator was built using Flash and the Adobe Air software program (Fig. 3). During biweekly team meetings, changes were proposed to existing functionality, and next steps were discussed and agreed on. Once most functionality was established, modifications to optimize the feedback and scoring systems were proposed and implemented. The prototype underwent several rounds of review and feedback, and an iterative prototype for testing was generated. During the SME review, all SMEs independently tested the prototype for both content accuracy and usability. When developing the prototype, a spreadsheet was drafted to communicate the input and expected outputs for every possible user action on the simulator with the developer. One SME (AO) validated the content by formally and systematically testing each possible response to all sections of the tactics and cases to ensure that the appropriate and expected outcomes and feedback were given for each step. Errors were identified and corrected, and each subsequent version was retested in this systematic manner until no errors were identified. The other SMEs informally tested content by entering both correct and incorrect inputs to assess for content accuracy.

Iterative Prototype

Feedback from usability testing can be used to address bugs, usability problems, and content modifications in the iterative prototype. The revised prototype is then ready for the next round of usability testers. Critical problems including bugs, content and spelling errors, and scoring mistakes were sent to the game developer and fixed immediately, and other usability issues were discussed individually by the team at meetings. The team came to consensus about whether those changes were necessary to improve the game and would be desired by most learners and whether they were within the scope and budget

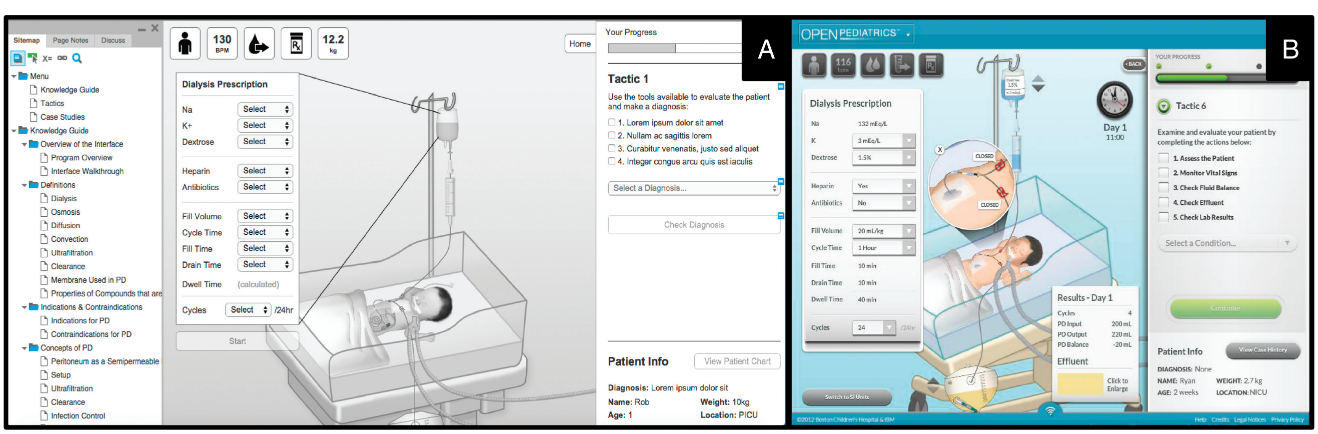

FIGURE 3. A, Clickable wireframe. B, Iterative prototype.

of the project. Approved changes were made to the prototype and usability testing continued with the new prototype.

Formative Evaluation

Usability Testing

Formal usability testing is conducted to identify content, design, functionality, and usability problems with the game. Many strategies for testing have been reported including focus groups, Think Aloud Protocol testing, SUS surveys, Likert scale questionnaires, written or oral open-ended feedback, and video recording of users. Only small numbers of tester are needed to identify most usability problems.[80] *Testers should include SMEs and end users.* Testing was conducted in four rounds over 4 weeks, with thirteen participants undergoing Think Aloud Protocol testing, the SUS survey, Likert scale survey, and open-ended questionnaire. During Think Aloud Protocol testing, each participant used the simulator, and an observer gave prompts to assess key aspects of the design, content, and functionality. Each session lasted roughly 1 hour, during which the observer recorded and categorized participant comments as content edits, user interface (UI) edits, and bugs. Next, an SUS survey and a quantitative questionnaire (Likert scale–rated statements assessing usefulness, enjoyment, interest clarity, and utility), and open-ended questionnaire (asking for overall improvement suggestions) were verbally administered with responses recorded. After each cycle of testing (3–4 participants per cycle), the prototype was modified on the basis of feedback as stated previously. Edits decreased, and user satisfaction and usability increased during the course of testing. Between cycles, the total bugs identified, content edits, and UI edits decreased (Fig. 4).

Final Product Delivery

After usability testing is complete and all edits have been made, the final product is delivered for beta testing and/or release. Validation, acceptance, and assessment of educational gains are often conducted after release. An agreement detailing how long the game developer will make additional edits to the game and at what cost should be created. After usability testing with resultant modifications made, the PD simulator was released for beta testing and validation.

DISCUSSION

To our knowledge, this is the first report to review and describe the existing literature on serious medical game development. Our results demonstrate that few authors report the development process used to create their game. Building on the development practices described in the existing literature and considering some of the reported challenges with game

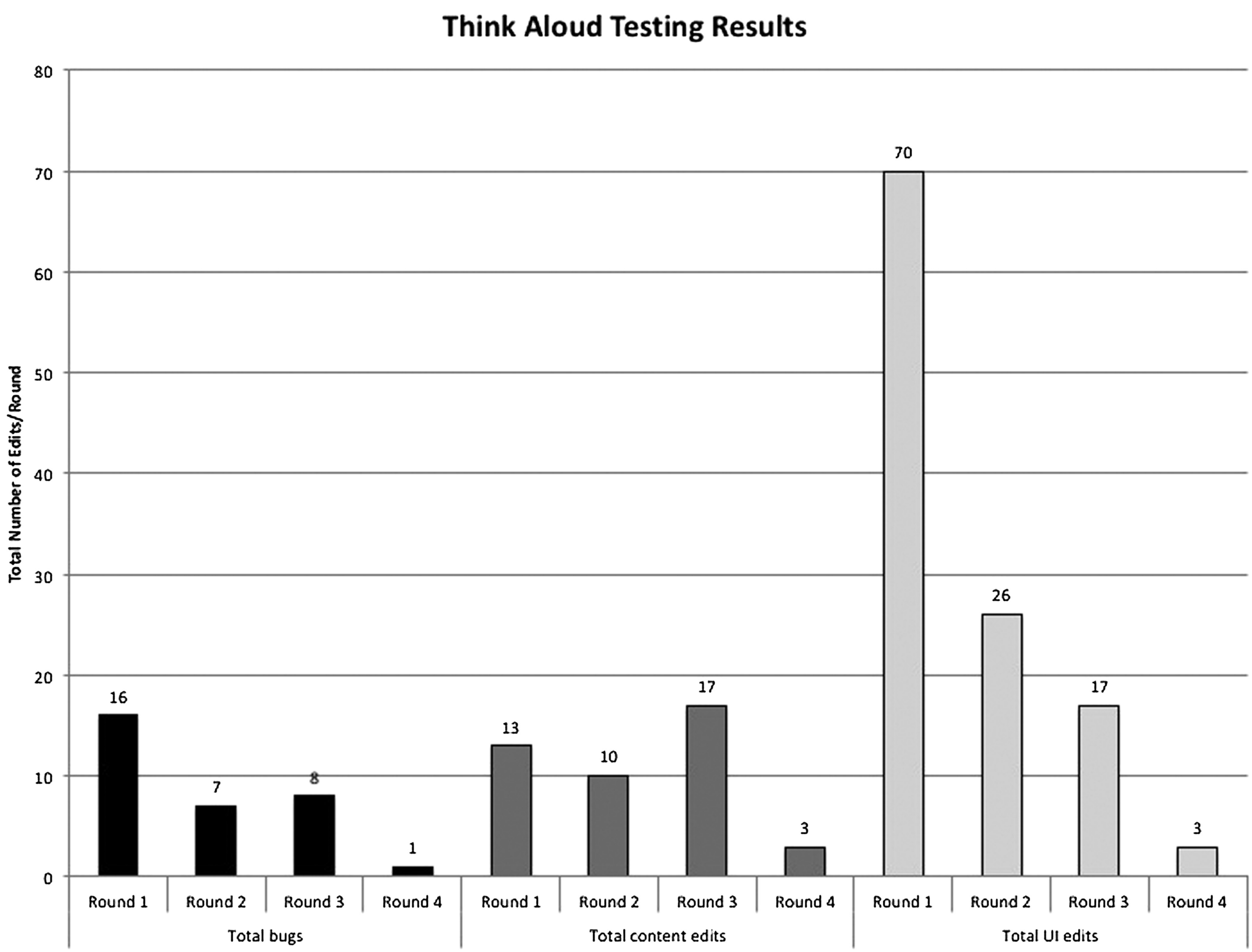

FIGURE 4. Think Aloud Protocol testing results for each round of testing, organized according to total bugs, total content edits, and total UI edits.

development, we proposed a structured, three-phase iterative framework that can be used to guide future serious game development in medical education, especially for those new at game development.

To develop our PD simulator, we sought to develop a structured development framework incorporating the best practices from medical serious game development. We aimed to incorporate the practices that would yield the most efficient processes to optimize communication, development time, and cost. Although many articles describe serious game medical content development in great detail, we were unable to identify any articles detailing strategies for medical content transfer to the game development team. To improve transfer of medical concepts, we added a Medical Content Transfer step to our development process based on the advice from our game developer, because this strategy had been successful in streamlining process and aligning expectations in their work developing applications for advertising.

Our choice of strategies for usability testing may be argued. Many strategies for testing have been described and validated,[8–15] but we chose to include Think Aloud Protocol and questionnaires including the SUS because they were low cost, easy to administer, time efficient, and provided a great deal of information in qualitative and quantitative formats. Think Aloud Protocol methodology allowed for identification of usability issues, and the SUS scores monitored how iterations impacted game usability.[14,15] The SUS is a widely-used scale that assesses the ease of use of a product.[15] The SUS scores can be trended over time to detect changes in product usability.[15] We hoped these strategies would be sufficient to identify most usability problems. Since our release of the simulator 8 months ago, we have only uncovered one additional usability problem requiring correction, suggesting that the testing we conducted was sufficient. Alternatively, we could have chosen a focus groups strategy. However, we felt that it would be easier to schedule busy SMEs individually. Finally, we could have chosen to video record the participants, potentially in a usability studio, as was performed by some game developers. We acknowledge that we may have missed some feedback by trying to record it in real time, and having recordings would allow us to go back and gather additional data. However, our testing was relatively time- and cost-efficient, because all data were completely recorded during the session itself, requiring no additional evaluation time, and we did not require access to a costly usability studio to perform testing.

One limitation of this study is that the literature review was limited to medical education game development, and thus, evaluation strategies employed by serious games in other fields were not included, which may have limited our ability to find useful frameworks. In addition, many serious games in medical education have been created and released, likely internally, without publication, further limiting the strategies we could identify.

Another limitation of this study is that usability testing was conducted with participant responses given verbally to the direct observer, who was involved in the game development. Participants may have been less likely to voice negative feedback verbally, potentially elevating the SUS. However, the Think Aloud Protocol outlines a verbal and observed testing process, which encourages reflection and increases the likelihood of open-ended feedback, potentially enriching the data. An additional limitation is the small participant number for testing. However, the sample size used is standard practice for Think Aloud Protocol testing.[13–15,80] Higher numbers have been reported to offer diminishing value in informing a development process.[80] By conducting several rounds of testing with a small number of participants, we were able to efficiently identify bugs and usability issues without exhausting a large pool of testers or imposing a significant burden of testing time for the SMEs. Through small numbers of diverse testers, not only did we show decreased edits with each round, indicating improvement, but also we also found that proportionally fewer edits were related to technical issues and proportionally more comprised sophisticated content edits, which we believe allowed us to test our simulator more effectively.

In conclusion, very few game developers report details of the design framework used to develop their game. Our three-phase development process incorporates the best practices from medical education and software design, may help improve communication and encourage more efficient and effective product development, and may be generalizable for future serious game development in the education of health professionals. The next steps should include applying the framework to other serious games types with differing education technology principles to assess its generalizability and efficacy. We would also like to compare other usability testing strategies to identify the most useful and efficient method for testing.

ACKNOWLEDGMENTS

The authors thank the Department of Anesthesiology, Perioperative and Pain Medicine, and the Technology and Innovation Development Office at Boston Children's Hospital for their generous financial support for the development of the PD simulator. We thank the PD simulator development team including our SMEs, Daniel Hames, Mignon McColluch, Debbie Stein, and Sharon Su, and our game developers at Genuine for their creativity and guidance during development, especially Rob Brecher and Paul Devlin. We are thankful to Brittanie Marques and Sarah Kim for their beautiful animations and illustrations, Nabila Mirza for her statistical analysis, the numerous reviewers and testers who have helped improve the learner experience, and the Boston Children's Hospital Simulator Program for allowing us to demonstrate PD in action in the simulation suite.

REFERENCES

1. Wang R, DeMaria S Jr, Goldberg A, Katz D. A systematic review of serious games in training health care professionals. *Simul Healthc* 2016;11(1):41–51.
2. Graafland M, Schraagen JM, Schijven MP. Systematic review of serious games for medical education and surgical skills training. *Br J Surg* 2012;99(10):1322–1330.
3. de Wit-Zuurendonk LD, Oei SG. Serious gaming in women's health care. *BJOG* 2011;118(Suppl 3):17–21.
4. Cook DA, Triola MM. Virtual patients: a critical literature review and proposed next steps. *Med Educ* 2009;43(4):303–311.
5. Cook DA, Hatala R, Brydges R, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. *JAMA* 2011;306(9):978–988.
6. Roberts DH, Newman LR, Schwartzstein RM. Twelve tips for facilitating Millennials' learning. *Med Teach* 2012;34(4):274–278.
7. Morris BJ, Croker S, Zimmerman C, Gill D, Romig C. Gaming science: the "Gamification" of scientific thinking. *Front Psychol* 2013;4:607.

8. Davids MR, Chikte UM, Halperin ML. Development and evaluation of a multimedia e-learning resource for electrolyte and acid-base disorders. *Adv Physiol Educ* 2011;35(3):295–306.

9. Gleason AW. RELM: developing a serious game to teach evidence-based medicine in an academic health sciences setting. *Med Ref Serv Q* 2015;34(1):17–28.

10. Vankipuram A, Khanal P, Ashby A, et al. Design and development of a virtual reality simulator for advanced cardiac life support training. *IEEE J Biomed Health Inform* 2014;18(4):1478–1484.

11. Davies AM. *Software Prototyping Advances in Computers* 1995;40:39–63.

12. Tripp SD, Bichelmeyer B. Rapid prototyping: an alternative instructional design strategy. *Educ Technol Res Dev* 1990;38(1):31–44.

13. Still B, Morris MJ. The blank-page technique: reinvigorating paper prototyping in usability testing. *IEEE Transactions on Personal Communication* 2010;53(2):144–157.

14. Jaaskelainen R. Think-aloud protocol. In: Gambier Y, van Doorslaer Luc, eds. *Handbook of Translation Studies*. 1st ed. Philadelphia, PA: John Benjamins Publishing Company; 2010:371–373.

15. Brooke J. SUS-A quick and dirty usability scale. In: Jordan ALMPW, Weerdmeester BA, eds. *Usability Evaluation in Industry*. London: Taylor & Francis; 1996:4–7.

16. Wolbrink TA, Kissoon N, Burns JP. The development of an internet-based knowledge exchange platform for pediatric critical care clinicians worldwide*. *Pediatr Crit Care Med* 2014;15(3):197–205.

17. Akl EA, Mustafa R, Slomka T, Alawneh A, Vedavalli A, Schunemann HJ. An educational game for teaching clinical practice guidelines to internal medicine residents: development, feasibility and acceptability. *BMC Med Educ* 2008;8:50.

18. Amer RS, Denehy GE, Cobb DS, Dawson DV, Cunningham-Ford MA, Bergeron C. Development and evaluation of an interactive dental video game to teach dentin bonding. *J Dent Educ* 2011;75(6):823–831.

19. Andreatta PB, Maslowski E, Petty S, et al. Virtual reality triage training provides a viable solution for disaster-preparedness. *Acad Emerg Med* 2010;17(8):870–876.

20. Antoniou PE, Athanasopoulou CA, Dafli E, Bamidis PD. Exploring design requirements for repurposing dental virtual patients from the web to second life: a focus group study. *J Med Internet Res* 2014;16(6):e151.

21. Atack L, Parker K, Rocchi M, Maher J, Dryden T. The impact of an online interprofessional course in disaster management competency and attitude towards interprofessional learning. *J Interprof Care* 2009;23(6):586–598.

22. Bergeron B. Augmented assessment as a means to augmented reality. *Stud Health Technol Inform* 2006;119:49–51.

23. Boeker M, Andel P, Vach W, Frankenschmidt A. Game-based e-learning is more effective than a conventional instructional method: a randomized controlled trial with third-year medical students. *PLoS One* 2013;8(12):e82328.

24. Buttussi F, Pellis T, Cabas Vidani A, Pausler D, Carchietti E, Chittaro L. Evaluation of a 3D serious game for advanced life support retraining. *Int J Med Inform* 2013;82(9):798–809.

25. Cavazza M, Charles F. Towards interactive narrative medicine. *Stud Health Technol Inform* 2013;184:59–65.

26. Cowan B, Sabri H, Kapralos B, Moussa F, Cristancho S, Dubrowski A. A serious game for off-pump coronary artery bypass surgery procedure training. *Stud Health Technol Inform* 2011;163:147–149.

27. Craig FE, McGee JB, Mahoney JF, Roth CG. The Virtual Pathology Instructor: a medical student teaching tool developed using patient simulator software. *Hum Pathol* 2014;45(10):1985–1994.

28. Creutzfeldt J, Hedman L, Fellander-Tsai L. Effects of pre-training using serious game technology on CPR performance—an exploratory quasi-experimental transfer study. *Scand J Trauma Resusc Emerg Med* 2012;20:79.

29. Creutzfeldt J, Hedman L, Fellander-Tsai L. Using virtual world training to increase situation awareness during cardiopulmonary resuscitation. *Stud Health Technol Inform* 2014;196:83–85.

30. Creutzfeldt J, Hedman L, Medin C, Heinrichs WL, Fellander-Tsai L. Exploring virtual worlds for scenario-based repeated team training of cardiopulmonary resuscitation in medical students. *J Med Internet Res* 2010;12(3):e38.

31. Creutzfeldt J, Hedman L, Medin C, Wallin CJ, Fellander-Tsai L. Effects of repeated CPR training in virtual worlds on medical students' performance. *Stud Health Technol Inform* 2008;132:89–94.

32. Del Cura-Gonzalez I, Lopez-Rodriguez JA, Sanz-Cuesta T, et al. Effectiveness of a strategy that uses educational games to implement clinical practice guidelines among Spanish residents of family and community medicine (e-EDUCAGUIA project): a clinical trial by clusters. *Implement Sci* 2016;11:71.

33. Delasobera BE, Goodwin TL, Strehlow M, et al. Evaluating the efficacy of simulators and multimedia for refreshing ACLS skills in India. *Resuscitation* 2010;81(2):217–223.

34. Dev P, Heinrichs WL, Youngblood P. CliniSpace: a multiperson 3D online immersive training environment accessible through a browser. *Stud Health Technol Inform* 2011;163:173–179.

35. Diehl LA, de Souza RM, Gordan PA, Esteves RZ, Coelho IC. User assessment of "InsuOnLine," a game to fight clinical inertia in diabetes: a pilot study. *Games Health J* 2015;4(5):335–343.

36. Diehl LA, Souza RM, Alves JB, et al. InsuOnline, a serious game to teach insulin therapy to primary care physicians: design of the game and a randomized controlled trial for educational validation. *JMIR Res Protoc* 2013;2(1):e5.

37. Duque G, Fung S, Mallet L, Posel N, Fleiszer D. Learning while having fun: the use of video gaming to teach geriatric house calls to medical students. *J Am Geriatr Soc* 2008;56(7):1328–1332.

38. Evans KH, Daines W, Tsui J, Strehlow M, Maggio P, Shieh L. Septris: a novel, mobile, online, simulation game that improves sepsis recognition and management. *Acad Med* 2015;90(2):180–184.

39. Fonseca LM, Aredes ND, Dias DM, Scochi CG, Martins JC, Rodrigues MA. Serious game e-Baby: nursing students' perception on learning about preterm newborn clinical assessment. *Rev Bras Enferm* 2015;68(1):9–14.

40. Fonseca LM, Dias DM, Goes Fdos S, et al. Development of the e-Baby serious game with regard to the evaluation of oxygenation in preterm babies: contributions of the emotional design. *Comput Inform Nurs* 2014;32(9):428–436.

41. Foster J, Dallemagne C. Teaching undergraduate nursing students renal care in a 3D Gaming Environment. *Stud Health Technol Inform* 2009;146:837–840.

42. Graafland M, Bemelman WA, Schijven MP. Prospective cohort study on surgeons' response to equipment failure in the laparoscopic environment. *Surg Endosc* 2014;28(9):2695–2701.

43. Graafland M, Vollebergh MF, Lagarde SM, van Haperen M, Bemelman WA, Schijven MP. A serious game can be a valid method to train clinical decision-making in surgery. *World J Surg* 2014;38(12):3056–3062.

44. Graafland M SM. A serious game to improve situation awareness in laparoscopic surgery. *Games for Health* 2013:173–182.

45. Hannig A, Kuth N, Ozman M, Jonas S, Spreckelsen C. eMedOffice: a web-based collaborative serious game for teaching optimal design of a medical practice. *BMC Med Educ* 2012;12:104.

46. Huguet L, Lourdeaux D, Sabouret N, Ferrer MH. Perturbed communication in a virtual environment to train medical team leaders. *Stud Health Technol Inform* 2016;220:146–149.

47. Johnsen HM, Fossum M, Vivekananda-Schmidt P, Fruhling A, Slettebo A. A serious game for teaching nursing students clinical reasoning and decision-making skills. *Stud Health Technol Inform* 2016;225:905–906.

48. Juanes JA, Alonso P, Hernandez F, Ruisoto P, Muriel C. Anatomical-ultrasound visor for regional anaesthesia. *J Med Syst* 2016;40(7):158.

49. Kanthan R, Senger JL. The impact of specially designed digital games-based learning in undergraduate pathology and medical education. *Arch Pathol Lab Med* 2011;135(1):135–142.

50. Katz DDS, Khanal P, Kahol K. Serious gaming to improve the safety of central venous catheter placement. *J Educ Informatics Cybern* 2013;11(3):78–81.

51. Kerfoot BP, Baker H. An online spaced-education game for global continuing medical education: a randomized trial. *Ann Surg* 2012;256 (1):33–38.

52. Kerfoot BP, Baker H, Pangaro L, et al. An online spaced-education game to teach and assess medical students: a multi-institutional prospective trial. *Acad Med* 2012;87(10):1443–1449.

53. Kerfoot BP, Turchin A, Breydo E, Gagnon D, Conlin PR. An online spaced-education game among clinicians improves their patients' time to blood pressure control: a randomized controlled trial. *Circ Cardiovasc Qual Outcomes* 2014;7(3):468–474.

54. Kizakevich PN, Culwell A, Furberg R, et al. Virtual simulation-enhanced triage training for Iraqi medical personnel. *Stud Health Technol Inform* 2007;125:223–228.

55. Knight JF, Carley S, Tregunna B, et al. Serious gaming technology in major incident triage training: a pragmatic controlled trial. *Resuscitation* 2010;81(9):1175–1179.

56. Kurenov SN, Cance WW, Noel B, Mozingo DW. Game-based mass casualty burn training. *Stud Health Technol Inform* 2009;142:142–144.

57. Lagro J, van de Pol MH, Laan A, Huijbregts-Verheyden FJ, Fluit LC, Olde Rikkert MG. A randomized controlled trial on teaching geriatric medical decision making and cost consciousness with the serious game GeriatriX. *J Am Med Dir Assoc* 2014;15(12):957.e1–957.e6.

58. LeRoy Heinrichs W, Youngblood P, Harter PM, Dev P. Simulation for team training and assessment: case studies of online training with virtual worlds. *World J Surg* 2008;32(2):161–170.

59. Lim EC, Seet RC. Using an online neurological localisation game. *Med Educ* 2008;42(11):1117.

60. Lin DT, Park J, Liebert CA, Lau JN. Validity evidence for Surgical Improvement of Clinical Knowledge Ops: a novel gaming platform to assess surgical decision making. *Am J Surg* 2015;209(1):79–85.

61. Makransky G, Bonde MT, Wulff JS, et al. Simulation based virtual learning environment in medical genetics counseling: an example of bridging the gap between theory and practice in medical education. *BMC Med Educ* 2016;16:98.

62. McKenzie K. A comparison of the effectiveness of a game informed online learning activity and face to face teaching in increasing knowledge about managing aggression in health settings. *Adv Health Sci Educ Theory Pract* 2013;18(5):917–927.

63. Mohan D, Angus DC, Ricketts D, et al. Assessing the validity of using serious game technology to analyze physician decision making. *PLoS One* 2014;9(8):e105445.

64. Nosek TM, Cohen M, Matthews A, et al. A serious gaming/immersion environment to teach clinical cancer genetics. *Stud Health Technol Inform* 2007;125:355–360.

65. O'Neill E, Reynolds PA, Hatzipanagos S, Gallagher JE. Graphic (games research applied to public health with innovative collaboration)—designing a serious game pilot for dental public health. *Bull Group Int Rech Sci Stomatol Odontol* 2013;51(3):e30–e31.

66. Rondon S, Sassi FC, Furquim de Andrade CR. Computer game-based and traditional learning method: a comparison regarding students' knowledge retention. *BMC Med Educ* 2013;13:30.

67. Roy MJ, Sticha DL, Kraus PL, Olsen DE. Simulation and virtual reality in medical education and therapy: a protocol. *Cyberpsychol Behav* 2006;9(2):245–247.

68. Schwaab J, Kman N, Nagel R, et al. Using second life virtual simulation environment for mock oral emergency medicine examination. *Acad Emerg Med* 2011;18(5):559–562.

69. Schwarz D, Stourac P, Komenda M, et al. Interactive algorithms for teaching and learning acute medicine in the network of medical faculties MEFANET. *J Med Internet Res* 2013;15(7):e135.

70. Schwid HA, Rooke GA, Michalowski P, Ross BK. Screen-based anesthesia simulation with debriefing improves performance in a mannequin-based anesthesia simulator. *Teach Learn Med* 2001;13(2):92–96.

71. Shewaga R, Knox A, Ng G, Kapralos B, Dubrowski A. Z-DOC: a serious game for Z-plasty procedure training. *Stud Health Technol Inform* 2013;184:404–406.

72. Sullivan SA, Bingman E, O'Rourke A, Pugh CM. Piloting virtual surgical patient cases with 3rd-year medical students during the surgery rotation. *Am J Surg* 2016;211(4):689–696.e1.

73. Sward KA, Richardson S, Kendrick J, Maloney C. Use of a Web-based game to teach pediatric content to medical students. *Ambul Pediatr* 2008;8 (6):354–359.

74. Tolsgaard MG, Jepsen RM, Rasmussen MB, et al. The effect of constructing versus solving virtual patient cases on transfer of learning: a randomized trial. *Perspect Med Educ* 2016;5(1):33–38.

75. Weiner CK, Skalen M, Harju-Jeanty D, et al. Implementation of a Web-based patient simulation program to teach dental students in oral surgery. *J Dent Educ* 2016;80(2):133–140.

76. Wilkening GL, Gannon JM, Ross C, et al Evaluation of branched-narrative virtual patients for interprofessional education of psychiatry residents. *Acad Psychiatry* 2016. In press.

77. Youngblood P, Harter PM, Srivastava S, Moffett S, Heinrichs WL, Dev P. Design, development, and evaluation of an online virtual emergency department for training trauma teams. *Simul Healthc* 2008;3(3):146–153.

78. Zaveri PP, Davis AB, O'Connell KJ, Willner E, Aronson Schinasi DA, Ottolini M. Virtual reality for pediatric sedation: a randomized controlled trial using simulation. *Cureus* 2016;8(2):e486.

79. De Leo G, Diggs LA, Radici E, Mastaglio TW. Measuring sense of presence and user characteristics to predict effective training in an online simulated virtual environment. *Simul Healthc* 2014;9(1):1–6.

80. Nielsen J. Estimating the number of subjects needed for a thinking aloud test. *Int J Hum Comput Stud* 1994;41(3):385–397.