Implementation of an Interactive Virtual-World Simulation for Structured Surgeon Assessment of Clinical Scenarios

Vishal Patel, MRCS, Rajesh Aggarwal, MA, PhD, FRCS, Daniel Cohen, MRCS, Dave Taylor, MSc, MBCS, Ara Darzi, KBE, Hon FREng, FMedSci, MD, FACS, FRCS

BACKGROUND: A novel simulation technology has emerged through the use of online 3-dimensional virtual worlds in which it is feasible to create virtual patients. This study establishes the face, content and construct validity of online 3-dimensional virtual patients in Second Life (a 3-dimensional virtual world accessible via the Internet).

STUDY DESIGN: Sixty-three surgeons of the following grades participated in this study: intern (n = 20); junior resident (n = 15); senior resident (n = 18), and attending (n = 10). All subjects assessed a series of 3 virtual patients (level 1) with different surgical presentations, such as lower gastrointestinal bleeding, acute pancreatitis, and small bowel obstruction. The junior resident group managed an additional 3 cases (level 2) with the same presentation but of increasing complexity. The senior resident and attending groups completed a total of 9 cases (level 1 to 3). The primary outcomes measures were the face and content validity rated on a 7-point Likert scale and a performance score based on a performance rating.

RESULTS: The simulation demonstrated high face and content validity ratings. Eight of 9 cases, with the exception of the level 3 small bowel obstruction, demonstrated significant differences in performance among the user groups (p < 0.01). Additional subset analysis demonstrated that the attending group performed best for performance ratings.

CONCLUSIONS: This novel form of simulation demonstrated high face and content validity. Performance assessed in managing a series of virtual patients varies with different levels of surgical training. This simulation can be used to differentiate among these levels and can be implemented as a unique form of assessment. (J Am Coll Surg 2013;217:270-279. © 2013 by the American College of Surgeons)

Patient assessment and subsequent management is the cornerstone of the art of practicing medicine. Rather than support this approach, previous formal assessment in medical training has concentrated on obtaining relevant postgraduate qualification as opposed to training assessment exercises. However, the evolution of the Worldwide Web has led to the progression of online 3-dimensional (3D) virtual worlds, which has granted an opportunity for medical educators to enhance training paradigms through the use of online virtual patients.1,2

This article introduces the first online virtual patients within a virtual-world setting that can be used for surgical training or assessment.3 This depicts the process of how a trainee surgeon accesses the virtual world of Second Life (a 3D virtual world accessible via the Internet) to assess and manage a series of virtual patients and is subsequently evaluated on their performance. The aim of this study is to establish the face and content validity of the simulation as well as use this as an assessment tool by defining the construct validity for trainee surgeons of different grades.

METHODS
Subjects
The study recruited 63 participants. All participants were UK postgraduate surgeons. Subjects were recruited at
with the OS group. However, despite the statistical significance achieved, the clinical importance of this finding is questionable because the model projected a GFR change of -15 mL/min/1.73 m² in 36 months for the ER group, which we do not consider clinically significant, especially because it represents a “worst case scenario” according to the model. Also despite the model being adjusted for preoperative comorbidities, there is still a chance of selection bias (unaccounted-for differences between the groups) as in any retrospective analysis.

Limitation of these analyses is the small sample size as well as the lack of information on split renal function and other factors that may have influenced renal function during follow-up. The heterogeneity of aneurysm morphology and chosen treatment in each group could also represent relevant confounders even if not significant on statistical analysis. Despite these acknowledged limitations, this study presents the first comparative evidence of renal function after OS or ER for RAA and demonstrated no clinically relevant difference in GFR between the groups during available follow-up. Unlike previous reports,\(^4\)\(^6\) we did not observe a reduction in the number of antihypertensive medication between baseline and last follow-up in either the OS or the ER group.

CONCLUSIONS
Both OS and ER were safe and effective in treatments of RAA with similar early postoperative outcomes. Endovascular repair was associated with a shorter length of hospital stay. Postprocedure decline in renal function did not relevantly differ between the 2 groups during follow-up. The 2 approaches, when used appropriately, would complement rather than compete with each other.

Author Contributions
Study conception and design: Tsilimparis, Reeves, Ricotta
Acquisition of data: Tsilimparis, Dayama
Analysis and interpretation of data: Tsilimparis, Reeves, Dayama, Perez, Debus, Ricotta
Drafting of manuscript: Tsilimparis, Debus, Ricotta
Critical revision: Tsilimparis, Reeves, Dayama, Perez, Debus, Ricotta

REFERENCES
Table 1. Essential Themes for Each Case

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lower gastrointestinal bleed</td>
<td>Not life threatening; requires monitoring</td>
<td>Plus: substantial bleed, requires transfusion</td>
<td>Plus: persistent bleed + normal endoscopic investigations, requires definitive management</td>
</tr>
<tr>
<td>Acute pancreatitis</td>
<td>Simple pancreatitis; no systemic upset</td>
<td>Plus: Low urine output + exacerbation of COPD, requires IV rehydration + nebulizers</td>
<td>Plus: type II respiratory failure + acute renal failure, requires BiPAP and hemofiltration</td>
</tr>
<tr>
<td>Small bowel obstruction</td>
<td>Small bowel obstruction; requires conservative management</td>
<td>Plus: onset of strangulation, requires diagnosis and proceed to laparotomy</td>
<td>Plus: postoperative anastomotic leak, requires diagnosis and appropriate intervention</td>
</tr>
</tbody>
</table>

BiPAP, bilevel positive airway pressure.

different training levels, primarily intern (n = 20), junior resident (year 2 resident, n = 15), senior resident (year 4 resident, n = 18), and attending (n = 10). All nonattending grades were recruited from within the pan-London UK training scheme through email invitations. A total of 203 email invitations were distributed to the different training grades: intern (n = 64), junior resident (n = 81), and senior resident (n = 58). Attending grades were recruited from both London teaching and community hospitals via a similar process, with 46 email invitations distributed.

Setting
The study predominantly took place within the Research Laboratories of St Mary’s Hospital, Imperial College, National Health Service Trust, London. All online simulations occurred on a single laptop (Mac Book 15-inch, Pro-Intel Core 2 Duo, 2.5 GHz; Apple Inc.).

Online patient simulation
The virtual online patient simulation consists of a series of 9 cases. These have been designed and developed at 3 increasing levels of complexity for each surgical condition encountered, eg, rectal bleeding, acute pancreatitis, and small bowel obstruction (Table 1), with the acute pancreatitis case detailed more extensively (Fig. 1). These conditions were selected as the focus of the clinical encounters as they reflect common pathology encountered by surgical trainees.

The case content for each complexity level was designed after consultation with the United Kingdom Intercollegiate Surgical Curriculum Project. This curriculum provides the framework for systematic training from completion of the intern grade through to attending level in the United Kingdom. The main features of the cases were storyboarded by 2 surgeons (a senior resident and an attending). After the outline of the case progression, other recommendations were sought by 2 additional attendings, both of whom are nationally recognized surgery trainers of all training grades, according to the Intercollegiate Surgical Curriculum Programme framework. Consensus was obtained that all cases were deemed appropriate for intern (level 1), junior surgical trainee (level 2), and senior surgical trainee (level 3) for the current Intercollegiate Surgical Curriculum Programme framework.

The virtual-patient development process consisted of a team of developers with varying backgrounds, including general surgeons, an information technology project manager, and both Web-based and virtual-world developers with expertise in different scripting languages. The main features of all cases were planned using an open source Java program, VUE (Visual Understanding Environment). The details of such cases were depicted in Microsoft Word format (Microsoft). These features were transferred through the use of allied Web technologies into a virtual patient (Fig. 2) in Second Life.

Three different clinical environments were developed within Second Life—an emergency room, a clinical ward, and a high-dependency unit—in which the virtual patient was placed and these scripted objects. Additional details of the design and development have been described more extensively in a short video available at http://www.screencast.com/t/FAuxPADvNENi. The
Mr Hart is a 65 year old, man who presents with epigastric pain and 2 episodes of vomiting. His pain is burning in nature and radiates to his back. His past medical history includes COPD and high blood pressure. His observations are normal but his abdominal examination reveals epigastric tenderness. His Amylase is 1024 but his Glasgow score is 1, due to his age.

His breathing deteriorates on the ward where his Oxygen saturations reduce to 90% on air. His respiratory examination reveals a high-pitched polyphonic wheeze. He also has experienced a reduced urine output, which is less than 30 ml over the last four hours. His Arterial Blood Gas reveals a pH of 7.33 with retention of carbon dioxide (pCO₂ 6.5). His serum creatinine level has risen to 142.

He is subsequently transferred to the high dependency unit. His breathlessness deteriorates despite bronchodilator therapy. His oxygen saturations are now 87% on fIO₂ of 0.24. He has stopped producing urine over the last four hours and his serum creatinine has further risen to 328 with a serum potassium level of 6.4. His arterial blood gas shows a mixed respiratory and metabolic acidosis with a pH of 7.24.

To assess and manage a virtual patient, the subject enters their login details, which enables them to select a specific case. This login consists of a unique identifier for the user. The login process has the following purposes: primarily to generate the specific information for that case onto the virtual patient and to generate a comma-separated value file to the Web player. This comma-separated value file can then be retrieved from the Web player using Microsoft Excel (Microsoft). This Excel file then depicts each decision a user has made when managing a case, which can subsequently be used for assessment of cases.

The purpose of each case is for the user to assess the virtual patient through history taking, examination, and investigations, including blood profiles and x-rays. The subject is able to do this through selection of various options on a menu board (eg, an examination finding.
or a blood result). This then enables the user to view this result (Fig. 3) and interpret the findings. The subject can then select an appropriate management plan based on the information retrieved.

**Study design**

All participants completed an orientation process before commencing the surgical scenarios. This involved reading a manual detailing how to use the software to proceed in the simulation. This process was followed by a 10-minute period of familiarization with Second Life and the use of the avatar. Subsequently, subjects assessed and managed a single virtual patient, which was used as an orientation case. The orientation case consisted of a virtual patient with urinary retention. Post orientation, a demographic questionnaire was completed. The subjects then proceeded through the simulation.

The number of cases completed was related to the training level of the participant. Therefore, the most junior trainee (ie, intern) would complete 3 cases (level 1 cases only). The junior resident completed the level 1 and 2 cases, and the senior resident and attending completed all 3 cases at levels 1, 2, and 3. The cases were stratified according to the appropriate level of training, as the aim is to ultimately develop proficiency-based training in which the junior surgeons are deemed competent at their level before being able to proceed to the next level. For this to be performed, it was first necessary to identify what constituted proficiency at each level.

**Outcomes measures**

**Face/content validity**

As the 3D surgical virtual patients were a novel concept, we sought to determine participants' reactions to both the cases and the environment. This was determined through both a face and content validity scale. These were both developed by expert consensus within our group to explore views on the prominent features of the simulation. Subjects were asked to complete a 7-item face validity. This explored general items, such as the way in which to retrieve patient information through history, examination, etc, as well as the perceived realism of the environment. In addition, a 9-item content validity questionnaire was given to depict the realism of the content of each clinical case encountered.

**Objective assessment of performance**

The primary outcomes measure assessed was a performance score for each case. Each participant was scored according to a checklist criteria scale (Table 2). The performance score represented a total score of this scale. This checklist was designed from expert surgeon consensus within our group. A unique scale was created due to the limited number of assessment tools available for the management of patients; specifically surgical patients. The scale constituted a 7-point Likert rating scale for different categories, which encompassed the assessment of a patient. These included a category for the quality of history, examination, diagnostic investigation selection, diagnosis choice, and appropriateness of management decisions. The minimum obtainable score
on the performance scale was 7 and maximum score was 56. Consensus within our group was established to identify what constituted acceptable components for each rating. Two blinded raters (VP and DC) assessed 40 cases (10% of the total cases) performed with the inter-rater reliability determined for these cases using the Cronbach α coefficient. These cases were selected at random through the use of a random-number generator. A single rater (VP) rated the remaining cases. Ratings were conducted post hoc using the Excel spreadsheet depicting each participant’s decision.

**Statistical analysis**

Data were analyzed using SPSS version 17 for Mac (SPSS Inc). Pearson chi-square test was used to establish if there was any association among participants’ sex and opinions about e-learning and previous virtual-world use. The internal consistency of the face and content validity scale was assessed using the Cronbach α coefficient. In view of the nonparametric nature of the data, the Kruskal-Wallis test was used to assess for any differences among each user group for each case with regard to both performance score. Additional group comparative analysis was performed using the Mann-Whitney U test to detect where the outcomes differences lay among the specific groups.

**Table 3. Participant’s Views on E-Learning and Previous Use of Virtual Worlds**

<table>
<thead>
<tr>
<th>Groups</th>
<th>Intern</th>
<th>Senior resident</th>
<th>Attending</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find e-learning useful</td>
<td>18</td>
<td>14</td>
<td>17</td>
<td>7</td>
</tr>
<tr>
<td>No</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Previous virtual-world use</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Yes</td>
<td>19</td>
<td>13</td>
<td>12</td>
<td>8</td>
</tr>
</tbody>
</table>
Table 4. Results of the Face Validity Questionnaire

<table>
<thead>
<tr>
<th>Likert scale questions to assess face validity</th>
<th>Intern</th>
<th>Junior resident</th>
<th>Senior resident</th>
<th>Attending</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The process of taking a history from the virtual patient is realistic</td>
<td>6 (4-7)</td>
<td>6 (4-7)</td>
<td>6 (5-7)</td>
<td>5 (4-7)</td>
</tr>
<tr>
<td>2. The process of examining the virtual patient is realistic</td>
<td>6 (4-7)</td>
<td>6 (4-7)</td>
<td>6 (5-7)</td>
<td>5 (4-7)</td>
</tr>
<tr>
<td>3. The process of obtaining different investigations and their results for the virtual patient is realistic</td>
<td>6 (5-7)</td>
<td>6 (4-7)</td>
<td>6 (4-7)</td>
<td>5 (4-7)</td>
</tr>
<tr>
<td>4. The process of managing the virtual patient is realistic</td>
<td>6 (5-7)</td>
<td>6 (4-7)</td>
<td>6 (5-7)</td>
<td>6 (4-7)</td>
</tr>
<tr>
<td>5. The simulated A&amp;E was a realistic representation of a real A&amp;E</td>
<td>5 (2-7)</td>
<td>5 (2-7)</td>
<td>4 (2-7)</td>
<td>4 (4-6)</td>
</tr>
<tr>
<td>6. The simulated ward was a realistic representation of a real clinical ward</td>
<td>5 (2-7)</td>
<td>4.5 (2-7)</td>
<td>5.5 (4-7)</td>
<td></td>
</tr>
<tr>
<td>7. The simulated HDU was a realistic representation of a real HDU</td>
<td>4.5 (2-6)</td>
<td>5 (4-6)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Values are median (range).

A&E, accident and emergency; HDU, high dependency unit.

Performance analysis

Participant’s performance was analyzed through a performance score on completion of the cases. The inter-rater reliability for the assessment of performance score between the 2 raters (VP and DC) was high, with a Cronbach’s coefficient ≥0.811 for 40 cases assessed from the 3 levels. A single rater assessed the remaining cases (VP). The performance score for the 4 groups are reported in Table 6.

Level 1 cases

Analysis of performance demonstrated widespread significant differences among the different user groups (ie, intern, junior resident, senior resident, and attending) for the 3 cases at level 1: rectal bleeding (median 48, 50, 54, and 56, respectively; p < 0.001), acute pancreatitis (median 47, 51, 55, and 55.5, respectively, p < 0.001) (Fig. 4), and small bowel obstruction (median 49, 53, 55, and 56; p < 0.001).

Additional analysis revealed that the attending group and the senior resident group performed significantly better than the intern group for all cases (rectal bleeding: median 56, 54 vs 48; p < 0.001; acute pancreatitis: median 55.5, 55 vs 47; p < 0.001; and small bowel obstruction: median 56, 55 vs 49; p < 0.001). Both the attending and the senior resident group also performed significantly better than the junior resident group for the rectal bleeding case (median 56 vs 50; p < 0.001 and median 54 vs 50; p = 0.003) and the acute pancreatitis case (median 55.5 vs 51; p < 0.010 and median 55.5 vs 51; p < 0.001)"

Table 5. Results of the Content Validity Questionnaire

<table>
<thead>
<tr>
<th>Likert scale questions to assess content validity</th>
<th>Intern</th>
<th>Junior Resident</th>
<th>Senior Resident</th>
<th>Attending</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The case rectal bleeding level 1 was a realistic representation of a real-life clinical scenario</td>
<td>5.5 (2-7)</td>
<td>5 (4-6)</td>
<td>6 (4-7)</td>
<td>6 (5-7)</td>
</tr>
<tr>
<td>2. The case acute pancreatitis level 1 was a realistic representation of a real-life clinical scenario</td>
<td>6 (3-7)</td>
<td>5 (4-7)</td>
<td>6 (3-7)</td>
<td>6 (5-7)</td>
</tr>
<tr>
<td>3. The case small bowel obstruction level 1 was a realistic representation of a real-life clinical scenario</td>
<td>6 (3-7)</td>
<td>5 (5-7)</td>
<td>6 (3-7)</td>
<td>6 (5-7)</td>
</tr>
<tr>
<td>4. The case rectal bleeding level 2 was a realistic representation of a real-life clinical scenario</td>
<td>X</td>
<td>5 (5-7)</td>
<td>6 (3-6)</td>
<td>6 (5-7)</td>
</tr>
<tr>
<td>5. The case acute pancreatitis level 2 was a realistic representation of a real-life clinical scenario</td>
<td>X</td>
<td>5 (5-7)</td>
<td>5.5 (3-7)</td>
<td>6 (4-7)</td>
</tr>
<tr>
<td>6. The case small bowel obstruction level 2 was a realistic representation of a real-life clinical scenario</td>
<td>X</td>
<td>5 (5-7)</td>
<td>6 (3-6)</td>
<td>6 (4-7)</td>
</tr>
<tr>
<td>7. The case rectal bleeding level 3 was a realistic representation of a real-life clinical scenario</td>
<td>X</td>
<td>X</td>
<td>6 (3-7)</td>
<td>6 (4-7)</td>
</tr>
<tr>
<td>8. The case acute pancreatitis level 3 was a realistic representation of a real-life clinical scenario</td>
<td>X</td>
<td>X</td>
<td>5.5 (3-6)</td>
<td>6 (5-6)</td>
</tr>
<tr>
<td>9. The case small bowel obstruction level 3 was a realistic representation of a real-life clinical scenario</td>
<td>X</td>
<td>X</td>
<td>6 (3-7)</td>
<td>6 (5-7)</td>
</tr>
</tbody>
</table>

Values are median (range).
**Table 6. All Groups Performance Score for Each Case**

<table>
<thead>
<tr>
<th>Level</th>
<th>Case</th>
<th>Group comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rectal bleeding</td>
<td>Intern vs junior resident p = 0.006*</td>
</tr>
<tr>
<td></td>
<td>Acute pancreatitis</td>
<td>Intern vs junior resident p = 0.001*</td>
</tr>
<tr>
<td></td>
<td>Small bowel obstruction</td>
<td>Intern vs junior resident p = 0.001*</td>
</tr>
<tr>
<td>2</td>
<td>Rectal bleeding</td>
<td>Intern vs senior resident p = 0.001*</td>
</tr>
<tr>
<td></td>
<td>Acute pancreatitis</td>
<td>Intern vs senior resident p = 0.001*</td>
</tr>
<tr>
<td></td>
<td>Small bowel obstruction</td>
<td>Intern vs senior resident p = 0.001*</td>
</tr>
<tr>
<td>3</td>
<td>Rectal bleeding</td>
<td>Intern vs attending p = 0.003*</td>
</tr>
<tr>
<td></td>
<td>Acute pancreatitis</td>
<td>Intern vs attending p = 0.003*</td>
</tr>
<tr>
<td></td>
<td>Small bowel obstruction</td>
<td>Intern vs attending p = 0.003*</td>
</tr>
</tbody>
</table>

*Statistical significance with p < 0.05.

55 vs 51; p = 0.004). The junior resident group performed better than the intern group for the acute pancreatitis case (median 51 vs 47; p = 0.006) and the small bowel obstruction case (median 53 vs 49; p = 0.001).

**Level 2 and 3 cases**

For all 3 cases, there were significant differences among the user groups at level 2. The attending and the senior resident group performed significantly better than the junior resident group for all 3 cases (rectal bleeding: median 49, 47 vs 40; p < 0.001; acute pancreatitis: median 49 vs 41; p = 0.001; small bowel obstruction: median 49 vs 42; p = 0.001 and median 45 vs 42; p = 0.002) (Fig. 5).

Analysis of performance between the attending and the senior resident groups for the level 3 cases showed a significant difference in the rectal bleeding case (median 47 vs 42; p < 0.001) and the pancreatitis case (median 49 vs 43; p = 0.012), but not the small bowel obstruction case (median 54 vs 49; p = 0.001 and median 45 vs 42; p = 0.002) (Fig. 5). The completion of the level 3 cases by the attending group resulted in the defining of benchmark criteria for performance score for each individual case.

**Self report**

All user groups performed significantly better than their perceived performance in the respective final small bowel obstruction case they undertook (Fig. 6) (p < 0.001) for all groups (attending: median 56 vs 44.5; p < 0.001; senior resident: median 54 vs 47.5; p < 0.001; junior resident: median 42 vs 36; p = 0.013; and intern: median: 49 vs 40; p < 0.001).
Medical education is undergoing a paradigm shift in which the time-honored specified training period is being replaced by a competency-based system. The Accreditation Council for Graduate Medical Education has listed a series of clinical competencies that are deemed integral to maintenance of the high standards of clinical practice. These stipulate 6 core competencies, including patient care, medical knowledge, practice-based learning, and improvement. The recent implementation of curtailed working hours, though, has had a profound impact on training opportunities and potentially the achievement of competence. Simulation-based education has been cited as a potential solution to this shortfall. However, in tandem, accurate and reliable methods through objective performance assessment are required to determine surgical competence. This can then make a substantial contribution to the quality of patient care in several important ways through surgeon training, selection, certification, and additional revalidation.

It is imperative that the inception of novel assessment methods display feasibility, validity, and reliability. This study has displayed inter-rater reliability, feasibility for implementation and the principal steps toward validity, ie, high face, content, and construct validity. This provides valuable and meaningful information about the simulation’s potential use as an assessment tool because it displays the extent of realism toward actual clinical situations—face validity, appropriate contextual content within the scenarios-content validity—as well as being able to discriminate among different levels of performance-construct validity. It was beyond the remits of this study to demonstrate concurrent validity comparing this assessment method with other related tools.

Computer-based simulation used for assessment already exists and has been used within the United States Medical Licensing Examination step 3 examination. Although this has represented a substantial advance in the use of simulation in assessment with its introduction having undergone the culmination of >20 years of research, 3D immersive and realistic environments create a potentially richer learning environment than standard Web-based methods. This is through the process of providing the user with a level of immersion in which they are responsive to the events and surroundings within the virtual place. A lack of such immersion can alter a user’s normal behavior and affect their performance within a simulated task.

An alternative method for structured assessment in medical training that is highly immersive is the objective structured clinical examination, which has become integral to both undergraduate and postgraduate assessment of performance. This method is reliant on the use of simulated patients, which is fraught with disadvantages, such as the time consumption of recruiting, training, and organization of the simulated patients, as well as the cost, which can be significantly high. Satava has questioned numerous institutions about the cost of using patient actors for objective structured clinical examination and found the range for training/assessment from $200,000 to $800,000 per year. Virtual patients, as described in this article, provide a much more cost-effective solution, as this simulation would cost $5,000 to host per annum. However, a critical
next step with such a simulation is to compare it with the objective structured clinical examination as a form of performance assessment.

This study has focused on the performance of the participants of different training grades. However, this study also highlighted an interesting finding in that all groups perceived their performance to be worse than their actual objective performance in the management of their final case, small bowel obstruction level 1, 2, or 3. This might be a generic perception of clinicians that resembles findings from previous studies in which clinicians perceived their performance to be different than their actual performance. An alternative explanation is that the surgeons’ unfamiliarity and first exposure to the software might have possibly resulted in them having less confidence in managing these virtual patients and self-rating themselves less than their actual objective measure of performance.

Confidence in using this technology was not established in this study; however, a high proportion of the participants had stipulated their interest in the use of e-learning. This raises the question of whether selection bias has occurred within the study, ie, the majority who volunteered to participate found e-learning useful. The important question remains as to whether a large group of participants who did not find e-learning useful would generate similar performance metrics. This would need to be established before mass dissemination of the simulation into educational programs.

This is the first single-player, 3D virtual-patient scenario in which subjects have been formally assessed on their individual performance. This study has demonstrated that it is feasible for 3D virtual patients to be used by trainees and subsequently for their performance to be assessed and benchmarked against more senior training groups. However, additional work needs to be done, including the development of a more rigorous scoring method and the dissemination of such cases to a larger cohort of trainees within the regional training scheme. This simulation was developed with the assistance of the London Deanery and without any financial assistance from Second Life. The London Deanery is a regional UK training group that consists of >900 postgraduate surgical trainees. The next aim is to deploy these cases into a more formal method of assessment, which will form part of the Annual Review of Competence Progression for General Surgical Trainees.

This unique form of 3D simulation has much potential that could be exploited. For example, such simulations could form a component of formal surgical examinations, informal regular assessments, or as part of the selection process for progressive surgical training, such as residency. However, it can be argued whether cognitive decision-making abilities obtained through such simulations are transferrable to the real-life context of patient assessment, ie, demonstrating predictive validity. This was not determined in the remit of this study; however, additional work is being done to establish whether the use of such scenarios could have an effect on patient care. Within our unit, a combined virtual-reality training simulation has been constructed involving preoperative virtual patients, the use of virtual-reality simulators to mimic the intraoperative phase, and postoperative virtual patients. Our aim was to establish whether training a group of surgeon trainees (intern to attending grade) within a department on this simulated virtual-patient pathway could result in an improvement of various patient outcomes. Although this study is in its infancy and focuses on training, it highlights the potential of the use of these 3D virtual patients for the future. Another potential use is the development of 3D virtual patients with rare or infrequently encountered pathology, such as a patient with enterocutaneous fistula. This will offer the trainee exposure and rehearsal for such cases before encountering them in real life.

CONCLUSIONS
This study has focused on emergency surgical conditions, but the reproducible nature of virtual-patient design renders it feasible to extend the virtual surgical patient across different settings, such as the pre- and postoperative elective cases or outpatient cases. More importantly, these cases can be extended to every form of medical subspecialty that encompasses patient contact. This study has demonstrated the inherent benefits of the use of such 3D virtual patients. However, it is acknowledged that additional work is required to develop such simulations. With such obvious benefits, it is suggested that such a unique and feasible technology be embraced to herald a new avenue in medical training and assessment.

Author Contributions
Study conception and design: Patel, Aggarwal, Taylor, Darzi
Acquisition of data: Patel, Aggarwal
Analysis and interpretation of data: Patel, Aggarwal, Cohen
Drafting of manuscript: Patel, Aggarwal
Critical revision: Cohen, Taylor, Darzi

Acknowledgment: The authors would like to acknowledge the contribution of James Kinross, PhD, FRCS and Karen Kerr, PhD for their assistance.

REFERENCES
Outcomes of Hepatectomy for Hepatolithiasis Based on 3-Dimensional Reconstruction Technique

Chi-hua Fang, MD, Jun Liu, MD, Ying-fang Fan, MD, Jian Yang, MD, Nan Xiang, MD, Ning Zeng, MD

BACKGROUND: The aim of our study was to evaluate the perioperative and long-term outcomes of hepatectomy based on 3-dimensional reconstruction technique for hepatolithiasis by comparing it with traditional hepatectomy.

STUDY DESIGN: From December 2005 to September 2012, 56 consecutive patients underwent hepatectomy based on 3-dimensional reconstruction technique for hepatolithiasis in our hospital (group A). During the same period, 42 patients with hepatolithiasis who met the inclusion criteria for hepatectomy were selected for traditional hepatectomy (group B). All operations were performed by the authors. There was no significant difference in preoperative data between the 2 groups statistically.

RESULTS: Compared with patients in group B, those in group A had a significantly lower stone residual rate (intermediate rate, 3.6% vs 19.0%; final rate, 0% vs 9.5%) and stone recurrence rate (3.6% vs 23.8%), a lower intrahepatic duct stricture residual rate (1.8% vs 14.3%), and a faster operating time (218.8 ± 55.5 minutes vs 254.7 ± 65.6 minutes). Intraoperative blood transfusion, intraoperative blood loss, postoperative hospital stay, and recurrent cholangitis rate were similar. No significant dominance was found in group A with respect to serum aminotransferase level, serum bilirubin level, serum albumin level, and prothrombin time. There was a significant dominance in group A for serum hemoglobin level (116.3 ± 16.0 g/L vs 108.0 ± 13.9 g/L; p < 0.05). Twenty-two complications occurred, 10 in group A and 12 in group B. Neither group had any perioperative mortality.

CONCLUSIONS: Hepatectomy for hepatolithiasis based on 3-dimensional reconstruction technique is feasible and safe in selected patients. Compared with traditional hepatectomy, it is more effective for diagnosis and treatment of hepatolithiasis. (J Am Coll Surg 2013;217:280–288. © 2013 by the American College of Surgeons)