International Journal of Surgery 12 (2014) 87–94

Contents lists available at ScienceDirect

International Journal of Surgery

journal homepage: www.journal-surgery.net

Review

Virtual reality simulation training in Otolaryngology

Asit Arora[a,*], Loretta Y.M. Lau[b], Zaid Awad[a], Ara Darzi[c], Arvind Singh[d], Neil Tolley[a]

[a] Department of ENT, St Mary's Hospital, Imperial College Healthcare NHS Trust, London W2 1NY, UK
[b] Imperial College School of Medicine, South Kensington Campus, London SW7 2AZ, UK
[c] Department of Biosurgery and Surgical Technology, St. Mary's Hospital, Imperial College London, W2 1NY, UK
[d] Department of Otolaryngology, Northwick Park Hospital, London HA1 3UJ, UK

ARTICLE INFO

Article history:
Received 5 October 2013
Accepted 14 November 2013
Available online 5 December 2013

Keywords:
Simulation training
Virtual reality
Otolaryngology
Temporal bone surgery
Myringotomy
Endoscopic sinus surgery
Validity
Surgical training

ABSTRACT

Objective: To conduct a systematic review of the validity data for the virtual reality surgical simulator platforms available in Otolaryngology.
Data sources: Ovid and Embase databases searched July 13, 2013.
Review methods: Four hundred and nine abstracts were independently reviewed by 2 authors. Thirty-six articles which fulfilled the search criteria were retrieved and viewed in full text. These articles were assessed for quantitative data on at least one aspect of face, content, construct or predictive validity. Papers were stratified by simulator, sub-specialty and further classified by the validation method used.
Results: There were 21 articles reporting applications for temporal bone surgery ($n = 12$), endoscopic sinus surgery ($n = 6$) and myringotomy ($n = 3$). Four different simulator platforms were validated for temporal bone surgery and two for each of the other surgical applications. Face/content validation represented the most frequent study type (9/21). Construct validation studies performed on temporal bone and endoscopic sinus surgery simulators showed that performance measures reliably discriminated between different experience levels. Simulation training improved cadaver temporal bone dissection skills and operating room performance in sinus surgery.
Conclusion: Several simulator platforms particularly in temporal bone surgery and endoscopic sinus surgery are worthy of incorporation into training programmes. Standardised metrics are necessary to guide curriculum development in Otolaryngology.

1. Introduction

Surgical trainees are required to achieve operative competency within a reduced period of clinical exposure compared to previous generations. Limited surgical exposure is compounded by the increasing workload of the trainer surgeon, the ethical and legal concerns over patient safety and the financial implications associated with accelerating the learning curve process[1–3]

In 2008, the Chief Medical Officer's annual report entitled 'Safer Medical Practice' advocated simulation-based surgical training in the UK.[4] The application of virtual reality (VR) simulation in surgical training was first proposed by Satava et al. in 1993 to deliver reproducible, consistent models which permit unlimited practice using standardised anatomy.[5] Training surgical tasks through repetitive, proctored sessions have been shown to improve the detection and analysis of surgical error.[6,7]

In the last decade, several VR simulators have been developed which produce a high fidelity representation of various operations in Otolaryngology. Three-dimensional projection, bimanual interaction and haptic (sensory) feedback are all features intended to enhance the user's experience. However, VR simulation is yet to be routinely incorporated into Otolaryngology training. In order for a simulator to be an effective training tool, it must include elements such as the ability for repetitive practice. Ideally it should be applicable for varying difficulty levels have established benchmarks and reliable outcome measures.[8] Robust validity data are essential to establish efficacy and guide application in surgical training.

The European Association for Endoscopic Surgery (EAES) guidelines outline the keystones of validation.[9] Face validity reflects the ability of a simulator to produce a realistic environment that resembles the actual surgical procedure. This is assessed using a trainer and trainee group using a structured questionnaire. Content validity is the assessment of the ability of the simulator to deliver what it is expected to achieve. This is demonstrated by satisfying

* Corresponding author. Department of ENT, St Mary's Hospital, Imperial College Healthcare NHS Trust, Praed St, London W2 1NY, UK. Tel.: +44 07976 897 446; fax: +44 207 886 1847.
E-mail address: asitarora@doctors.org.uk (A. Arora).

1743-9191/$ – see front matter
http://dx.doi.org/10.1016/j.ijsu.2013.11.007

pre-determined criteria that both groups agree upon. Construct validity confirms the simulator's ability to quantifiably differentiate between varying levels of expertise amongst participants or to test the ability of the model or tool to predict future performance (Predictive validity).

The objective of this review is to collect and critically analyse the evidence for VR simulation in Otolaryngology training and present a reference for program directors who are considering incorporating it into their training programs.

2. Methods

A systematic literature search was performed using Ovid Medline and Embase. Articles published until July 13, 2013 were included (Table 1). Fig. 1 summarises the search outcomes according to PRISMA guidelines.

2.1. Screening, eligibility and selection

Results from both databases produced a total of 432 citations. After removal of duplicates, 409 remained. Two independent reviewers (AA, LL) screened the citations based on title and abstract using the criteria outlined in Table 2 to determine relevance to Otolaryngology and postgraduate education and training. Thirty-six citations underwent full text review and references were hand searched for relevant studies. One additional paper (Fried et al., 2007) was included from reference searching. Each of these articles had quantitative data for at least one aspect of face, content, construct or predictive validity of the simulator.

2.2. Data extraction, analysis and outcomes

The author, date of publication, study design, and data from the eligible articles were tabulated in Microsoft Excel® (Microsoft Corporation, WA). Papers were stratified by simulator and subspecialty type and further classified by validation method; Face, content, construct and predictive validity.

Table 1
Search strategies for Ovid Medline and Embase databases (Search July 13, 2013).

Ovid (Medline)	
1	exp Otolaryngology/
2	head/or ear/or mouth/or nose/or parotid region/or exp skull base/or exp larynx/or exp nose/or exp pharynx/or exp trachea/
3	exp Neck/
4	virtual reality.mp.
5	simulator*.mp.
6	patient simulation/
7	1 or 2 or 3
8	4 or 5 or 6
9	7 and 8
10	limit 9 to (english language and humans)
Ovid (Embase)	
1	exp otorhinolaryngology/
2	exp "face, nose and sinuses"/
3	exp ear/or exp nose/or exp throat/
4	exp ethmoid bone/or exp facial bone/or exp hyoid bone/or exp mastoid/or exp temporal bone/or exp turbinate/
5	exp simulator/
6	exp virtual reality/
7	1 or 2 or 3 or 4
8	5 or 6
9	7 and 8
10	limit 9 to (english language and humans)

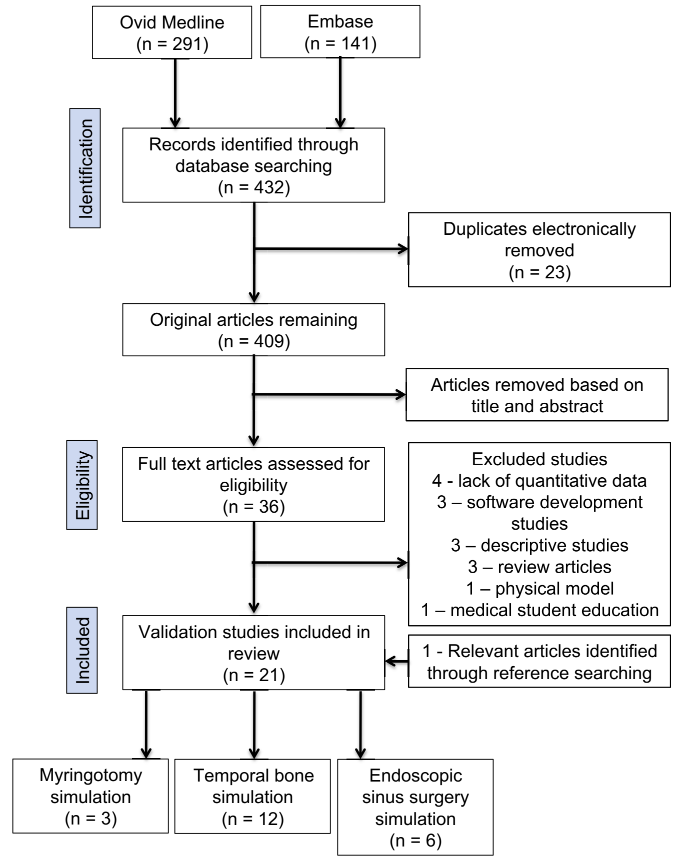

Fig. 1. Search outcomes of virtual reality surgical training simulation in Otolaryngology.

3. Results

There were 21 articles reporting on 3 main VR applications: temporal bone surgery, endoscopic sinus surgery and myringotomy (Table 3). A summary is shown in Table 4. There were 12 studies on temporal bone simulation using the Voxelman, Mediseus®, OSU (Ohio State University) or Stanford University platforms. Six studies were on endoscopic surgery simulation using the Endoscopic Sinus Surgery Simulator (ES3) or Dextroscope simulator. Three studies involved myringotomy simulation using the UWO (University of Western Ontario) haptics or optical tracker systems.

Other VR platforms were identified but articles describing these were excluded because they did not include quantitive or comparative data as outlined in the Method section (Table 5). The Mediseus is the only system with networking capability that allows a mentor to interactively guide the drilling process. The Voxelman is the only commercially available simulator at the present time.

3.1. Temporal bone surgery simulation

The Voxelman Temposurg, Mediseus®, OSU and Stanford platforms have all undergone validation studies with the aim of integration into postgraduate training programs in the UK, US and Australia (Table 4A).

3.1.1. Voxelman temporal bone simulator

Face validity was undecided although it was effective for training based on 20 respondents.[10] The largest evaluation of face and content validity was by Arora et al.[11] Eight-five participants were recruited comprising a trainer and trainee group. Although

Table 2
Criteria for review.

Inclusion
Quantitive data available
Face, content, construct or predictive validity addressed
Exclusion
Descriptive or development of simulator hardware/software
Posters
Conference proceedings
Reviews
Dental
Neurosurgery
Undergraduate education
Vestibular disorders and rehabilitation
Non virtual reality models

face validation was undecided it rated favourably for content validity including surgical anatomy and planning, drilling technique and hand eye coordination. Content validity was investigated in 2 further studies by examining the effect of iteration. Francis et al. assessed two component tasks of cortical mastoidectomy performed by 12 residents.[12] Technical performance improved with practice assessed by internal measures and a blinded expert evaluation. Nash et al. reported similar findings in 4 novices.[13]

Zirkle et al. evaluated construct validity by comparing the performance of novice and experienced groups in addition to cadaver temporal bone drilling.[14] The latter outperformed the novice group on both cadaver and simulation models. The simulator-derived metrics were able to determine between different levels of experience. Khemani et al. demonstrated construct validation on a subsequent iteration of this platform.[15] Sixty-five participants were recruited comprising 40 novices, 15 trainees and 10 experts. Experts and intermediates outperformed novices with respect to the total time taken to complete a standardised task, the total volume and efficiency of bone removal, time spent with the drill tip obscured and number of injuries to vital structures.

3.1.2. *Mediseus® temporal bone simulator*

O'Leary et al. assessed content validity recruiting 3 general surgery trainees without otolaryngology experience and 9 junior trainees. There was a positive consensus regarding usefulness for training.[16]

In a construct study, 27 participants were recruited comprising 12 experts, 6 residents and 9 novices.[17] Experts completed the task in significantly shorter time than others. Novices exerted higher forces, when close to vital structures, compared with experts ($P = 0.002$) and injured vital structures more frequently. Compared with residents, experts altered the force applied during deeper drilling near vital structures.

O'Leary et al. assessed predictive validity by comparing the performance of locating anatomical landmarks on the simulator and a cadaver temporal bone.[16] Trainees were able to identify anatomical structures with greater ability after simulation training. In a randomised controlled trial the mean performance score of cadaver temporal bone dissection was higher ($P = 0.04$) and fewer injuries were made ($P = 0.01$) in the simulation-trained group.[18] Another study used 20 participants to assess simulator effectiveness for self-directed learning.[19] The end product and technique scores were significantly higher in the simulation group (67% v 29%; $P < 0.001$).

3.1.3. *OSU temporal bone simulator*

Wiet et al. have published several studies reporting the development of this platform.[20–22] Predictive validation was assessed in a multi-institutional study in which 65 subjects were randomised to a two-week practice session using either the virtual or cadaver temporal bones. There was no difference between the two groups using a blinded rating tool to assess performance after training.

3.1.4. *Stanford temporal bone simulator*

In a construct study by Sewell et al., 8 experts and 7 novices performed a mastoidectomy on 2 occasions.[23] Each procedure was recorded and 2 experienced surgeons assigned global scores. The mean score of participants with prior surgical experience was higher ($P < 0.0001$) than the novices and the majority of performance metrics correlated strongly ($P < 0.05$) with the global scores.

Table 3
Available simulators in Otolaryngology.

Surgical application	Simulator	Capability			
		Hardware	Image rendering	Network	Image database
Temporal bone surgery	Mediseus surgical drilling simulator SDS, (CSIRO/University of Melbourne, Australia)	PHANTOM haptic device	Integrated commercial unit	Networking capable	CT data set
	VOXEL-MAN Temposurg (University Medical Center Hamburg-Eppendorf, Germany)	PHANTOM haptic device	Integrated commercial unit	None described	CT data set: cadaver, image libraries
	Ohio State University simulator (Ohio, USA)	PHANTOM haptic device	NVIDIA Quadro FX4600	None described	CT data set; cadaver
	Stanford surgical simulator (California, USA)	PHANTOM haptic device	Cyberscope	None described	CT data set
Sinus surgery	ES3 (Lockheed Martin, Ohio, USA)	PHANTOM haptic device, binocular display	SGI Octane workstation; Silicon Graphics Onyx with Infinite Reality graphics	None described	Visible human CT data set
	Dextroscope (Volume Interactions, Singapore)	Dextroscope workstation; stereoscopic glasses; stylus; control handle	Vizdexter 1.2, Volume Interactions, Singapore	None described	CT data set
Myringotomy	3D VR Myringotomy simulator with Haptics (University of Western Ontario, Canada)	PHANTOM haptic device; 3D binocular visor; styrofoam block for hand rest	Object-oriented Graphics Rendering Engine (OGRE 3D)	None described	Computer generated 3D meshes
	Myringotomy simulator with optical tracker (University of Western Ontario, Canada)	Optical tracker; 3D binocular visor, styrofoam block for hand rest	Open Dynamics Engine; Object-oriented Graphics Rendering Engine (OGRE); Object-oriented Input System (OIS)	None described	Computer generated 3D meshes

Table 4A
Validated temporal bone simulators.

Simulator	Author	Title	Validation objective	Data acquisition	Measured outcomes	
VOXEL-MAN Temposurg	Reddy-Kolanu et al., 2011	Evaluating the effectiveness of the Voxel-Man TempoSurg virtual reality simulator in facilitating learning mastoid surgery	Face	Likert scale rating	Feedback Repetitive practice Curriculum integration Range of difficulty Multiple learning strategies	Clinical variation Controlled environment Individualised learning Defined benchmarks
	Arora et al., 2012	Face and Content Validation of a Virtual Reality Temporal Bone Simulator	Face	Likert scale rating	Haptic feedback Ergonomics Appearance of structures Appearance of drill	Performance of drill Depth perception Quality of graphics
			Content	Likert scale rating	Teaching anatomy Teaching surgical planning Drilling technique	Instrument navigation Training hand-eye coordination Overall training tool
	Francis et al., 2012	Technical Skills Improve After Practice on Virtual-Reality Temporal Bone Simulator	Content	Objective assessment tool: OSATS	Time taken to complete task Injury to structures	
	Nash et al., 2012	Objective assessment of learning curves for the Voxel-Man TempoSurg temporal bone surgery computer simulator	Content	Quantitative metrics measurements	Overall task score Task completion time Target bone volume removed Number of mistakes	Drill path length
	Zirkle et al., 2007	Using a Virtual Reality Temporal Bone Simulator to Assess Otolaryngology Trainees	Construct	Quantitive metric measurements	Number of movements Distance travelled	Time taken to complete task
	Khemani et al., 2012	Objective Skills Assessment and Construct Validation of a Virtual Reality Temporal Bone Simulator	Construct	Quantitive metric measurements	Time taken to complete task Volume and efficiency of bone removal Forces applied Injury to structures	
Mediseus	O'Leary et al., 2008	Validation of a Networked Virtual Reality Simulation of Temporal Bone Surgery	Content	Likert scale rating	Teaching anatomy Teaching surgical planning Teaching drilling technique Usability of interface	Overall VR experience Transfer of learning Recommended VR for training use Support inclusion in training
			Predictive	Objective assessment tool	Identifying anatomy Surgical planning Surgical landmarks	
	Zhao et al., 2010	Differentiating levels of surgical experience on a virtual reality temporal bone simulator	Construct	Quantitive metrics measurements	Time taken to complete task Injuries to structures Forces applied	
	Zhao et al., 2011	Improving Temporal Bone Dissection Using Self-Directed Virtual Reality Simulation : Results of a Randomized Blinded Control Trial	Predictive	Objective assessment tool: OSATS	End product score Injury score Technique	
	Zhao et al., 2011	Can Virtual Reality Simulator Be Used as a Training Aid to Improve Cadaver Temporal Bone Dissection? Results of a Randomized Blinded Control Trial	Predictive	Objective assessment tool: OSATS, Wellings Scale	End product evaluation Injuries to structures Technique score	
Ohio State University simulator	Wiet et al., 2012	Virtual Temporal Bone Dissection System: OSU Virtual Temporal Bone System: Development and Testing	Predictive	Objective assessment tool: Wellings Scale	End product evaluation	
Stanford surgical simulator	Sewell et al., 2008	Providing Metrics and Performance Feedback in a Surgical Simulator	Construct	Quantitive metrics measurements	Drilling technique Suctioning technique Bone removal	Drill forces Drill velocities

3.2. Endoscopic sinus surgery simulation

3.2.1. Endoscopic sinus surgery simulator (ES3)

Four validation studies were retrieved (Table 4B). The ES3 simulator was assessed for performance characteristics by 14 trainees and 5 trainers.[24] The haptic device demonstrated a 77% success rate when used to identify structures without the aid of visual cues. Individuals who trained on the simulator demonstrated a significant improvement in psychomotor skills ($r = 0.63$, $P < 0.01$). Arora et al. found a significant correlation between hazards scores on the ES3 when compared to a device that measured depth perception ($r = 0.5$ $P < 0.001$).[25] Overall scores when compared to MIST-VR in executing important surgical tasks ($r = 0.57$, $P < 0.001$) were significant in addition to visual–spatial tests using cube and card comparisons ($r = 0.43$, $P < 0.01$, $r = 0.45$, $P < 0.01$).

In a construct validation study by Fried et al., 10 medical students 14 Otolaryngology residents and 10 experienced surgeons

Table 4B
Validated endoscopic sinus surgery simulators.

Simulator	Author	Title	Validation objective	Data acquisition	Measured outcomes
Endoscopic Sinus Surgery Simulator (ES3)	Rudman et al., 1998	Functional Endoscopic Sinus Surgery Training Simulator	Face	Likert scale rating	Visual interaction Manual interaction
	Arora et al., 2005	Assessment of Construct Validity of the Endoscopic Sinus Surgery Simulator	Construct	Quantitive assessment scores: PicSor, Mist-VR, cube comparison	Navigation Injection technique Dissection technique Hazardous maneouvres
	Fried et al., 2007	Construct Validity of the Endoscopic Sinus Surgery Simulator	Construct	Quantitive metric measurements	Mean score on task Dissection time
	Edmond et al., 2002	Impact of the Endoscopic Sinus Surgical Simulator on Operating Room Performance	Predictive	Objective assessment tool	Endoscope control Dissection skill
	Fried et al., 2010	From Virtual Reality to the Operating Room: The Endoscopic Sinus Surgery Simulator Experiment	Predictive	Quantitive metric measurements, objective assessment tool	Dissection task Case difficulty Tool manipulation Surgical confidence Number of errors
Dextroscope	Caversaccio et al., 2003	Virtual Simulator as a Training Tool for Endonasal Surgery	Predictive	Objective assessment tool	User-friendliness Identifying anatomical landmarks

were recruited.[26] The ES3 simulator was successfully used to distinguish between different ability levels.

Predictive validity was demonstrated for ESS and its successor, the ES3 Lockheed Martin simulator. Practice on the ES3 simulator illustrated an improvement in surgical confidence and instrument manipulation ($P = 0.01$, $P = 0.01$). In a study of 25 otolaryngology residents, simulation training reduced the overall operating time ($P < 0.001$) and the number of surgical errors made ($P = 0.05$).[27] Edmond et al. conducted an evaluation of the ES3 on operating room performance in 4 trainees.[28] The endoscopic video-recorded performance of two simulation-trained residents rated better than the other two residents across all measures.

3.2.2. Dextroscope endoscopic sinus simulator

In a face, content and predictive validation study, two PGY3 residents were trained on the Dextroscope simulator before performing the procedure in real and virtual settings.[29] The use of the simulator for learning manual skills was poorly rated although it was favourable for surgical anatomy. Simulation training did not improve operating room performance ($P = 0.19$).

3.3. Myringotomy simulation

3.3.1. Myringotomy with haptic device

A face validation study used 7 trainees and 4 experts[30] and showed good validity with high consistency between participants (Cronbach alpha: 0.92). Haptic feedback was poorly rated. Participants positively rated the system for improving hand eye coordination. The majority of experienced surgeons felt that the simulated task was easier to perform compared to the real procedure and therefore unrealistic. A subsequent version of this simulator was assessed by Ho et al.[31] Twelve Otolaryngologists were recruited but there were concerns regarding suboptimal face validity.

3.3.2. Myringotomy with optical tracker

In a face validation study, 4 residents and 2 experts were favourable for hand eye coordination and skills development.[32] Realism was undecided due to problems with image rendering and transition. The task was more difficult to perform compared to the operating room.

4. Discussion

This review identified over 400 potentially relevant articles, yet only a small proportion, approximately 5%, addressed the issue of validation using robust methodology and were eligible for inclusion. Many studies used small numbers and lacked power calculations. Data from experienced surgeons, trainees and novices with standardised forms and assessment tools are essential. Without this information it is not possible to draw conclusions regarding the usefulness of simulator platforms such as those listed in Table 5.

Twenty-one studies satisfied the selection criteria (Table 4). Due to heterogeneity in methodology, meta-analysis of outcomes on each simulator was not possible. Nevertheless, the findings demonstrate that VR simulation technology provides the platform for objective evaluation of technical skills in trainees.

4.1. Validation studies: synopsis of key findings

4.1.1. Face and content validity

Face and content validation was the most common study type representing nearly half the publications in this review (9/21). The high proportion reflects the fact that they are relatively straightforward to perform. For both temporal bone and endoscopic sinus surgery, reasonably realistic VR platforms exist. Developers strive to get their platform as close as possible to the real life experience and a simulator that provides the user with a truly immersive experience is more likely to achieve this. However, increasing realism is also associated with increased cost and does not necessarily improve skills transfer.[17,18] To a large extent the degree of realism required depends on the intended task and experience level of the intended user. Both the Voxelman and Mediseus temporal bone simulators did not achieve face validity although content validity was positively rated.[10,11,16] This suggests that VR simulation does not necessarily need to reach the highest level of fidelity to be an effective training tool, particularly at a junior level. The effect of repeated practice was more apparent in the experienced trainees on this simulator. They were able to perform the task in less time, with few errors and with better economy of motion. It suggests that one potential benefit is to enhance psychomotor and procedural learning when practice is based on an appropriate cognitive foundation.

Other models did not achieve content validity (myringotomy simulation). In this situation the simulator should only be used for performing validated tasks until further data becomes available.

Table 4C
Validated myringotomy simulators.

Simulator	Author	Title	Validation objective	Data acquisition	Measured outcomes
University of Western Ontario Myringotomy Simulator with Haptic Device	Sowerby et al., 2010	Development and Face Validity Testing of a Three Dimensional Myringotomy Simulator with Haptic Feedback	Face	Likert scale rating	Visual representation: anatomy, instruments Features of equipment Representation of movement
			Content	Likert scale rating	Development of skills Improving skills Hand eye coordination Usefulness as training tool
	Ho et al., 2012	Virtual Reality Myringotomy Simulation With Real-Time Deformation: Development and Validity Testing	Face	Likert scale rating, visual analogue scale	Efficacy of cutting algorithms Visual realism Response of instrumentation
University of Western Ontario Myringotomy Simulator with optical tracker	Wheeler et al., 2010	Interactive Computer-based Simulator for Training in Blade Navigation and Targeting in Myringotomy	Face	Likert scale rating	Visual representation: anatomy, instruments Features of equipment Representation of movement
			Content	Likert scale rating	Development of skills Improving skills Hand eye coordination Usefullness as training tool

4.1.2. Construct validity

This is the fundamental requirement for assessment and was evaluated in 6/21 studies. It was successfully demonstrated in 3 different simulator platforms in temporal bone surgery and in 1 sinus surgery simulator. Simulator-generated performance measures which reliably discriminate between different experience levels included time for task completion, number of injuries and economy of hand movement. The former does not necessarily reflect proficiency and should only be used in conjunction with other metrics to assess skills development. It would seem that novices approach a task with a degree of trepidation that results in slower completion times and more injuries to vital structures. Experienced surgeons perform surgical tasks in the operating room more efficiently, making fewer errors than novice surgeons.[33,34] The difference in performance between groups likely reflects the anatomic knowledge and ability to recognise a complication at an early stage.

Construct validity studies can also be used to guide which trainee level is likely to benefit the most from simulation training. This choice must be specific to each training programme due to the variation of expertise, syllabus, goals and pace at which training is being delivered. From the articles included in this review the definition of an experienced, intermediate and novice group varied as did the numbers within each group. The latter is important because too few numbers in a particular group preclude valid statistical analysis. Nevertheless, these studies demonstrate that the more junior the participants are, the more likely it is that training using virtual reality models will improve knowledge and skills.

4.1.3. Predictive validity

This was assessed in 7/21 studies. Four were performed with either the Mediseus ($n = 3$) or OSU temporal bone simulator. With the former, subjects were randomised to VR training or control groups. Two studies demonstrated that the VR simulator improves cadaveric temporal bone dissection performance compared with traditional teaching methods. The cadaver temporal bone is an excellent model for evaluating skills and represents the gold standard in this regard. The OSU study is the largest randomised validation paper of VR simulation in the literature. It demonstrated equivalence to the gold standard. However, skills acquired during VR simulation do not necessarily transfer to the clinical environment. VR-to-OR skills transfer is perhaps best viewed as a means of demonstrating a very deliberately designed VR training activity. Predictive validation studies of the ES3 used live operating as a comparison. Operative performance assessment included assessment applied to live observations or video recordings. Although subject numbers were small ($n = 4$) the results suggest that simulator training translates to improved performance in actual surgery. This includes shorter operating times, demonstration of higher confidence, better skills in instrument manipulation and fewer technical errors which is in keeping with skills transfer studies in other specialties.[35]

4.2. Limitations

A lack of standardisation meant that the results were not suitable for pooling which makes interpretation more difficult. Studies conducted by individuals who are affiliated with the manufacturer are liable to reporting bias. Results may also be affected by selection and performance bias. Voluntary enrolment and keeping the number of observers to a minimum minimise this risk but not all studies report whether this occurred.

When testing for face and content validity the subjective nature of evaluation is always a limitation. Most trainees are supportive, enthusiastic and appreciative of free training. Therefore caution must be exercised when interpreting face and content validity data by looking for high levels of agreement, large study numbers, anonymity and reproducibility. Most construct data do not differentiate between years of training. This information can be used as evidence for progress and to identify trainees who require remediation. Improvements in simulator realism and better objective measures are necessary before this can be addressed.

4.3. Integration of simulation training into an Otolaryngology curriculum

Virtual reality simulation is suited to the repeated practice, error correction and feedback required for a proficiency-based curriculum.[36] The development of standardised metrics to guide curriculum development is necessary. Appropriate selection of VR tasks, task difficulty levels and the definition of reasonable performance objectives are required. Both the duration and frequency of training sessions need to be established and study subjects need

Table 5
Non-validated Otolaryngology simulators.

Simulator type	Simulator	Capability				Key references	
		Hardware	Image rendering software	Networking	Image database	Author	Year
Temporal bone surgery	Interactive virtual dissection (Arizona)	None described	Stereoscopic images rendered based on determined surgical position of microscope	None described	Photographic SLR camera acquired images	Bernardo et al.	2003
	Virtual temporal bone	None described	Dextroscope Dextrobeam	None described	Visible Human CT data set	Kockro et al.	2009
	IERAPSI	PHANTOM haptic device (drill and suction)	Volume-rendering software developed at CRS4 (Center for Advanced Studies, Research and Development); No binocular display	None described	CT data set	Neri et al.	2006
	Visible Ear simulator (3D)	PHANTOM haptic device	Windows XP; Windows Vista; NVIDIA GeForce 8,800 GTX graphics card	None described	CT data set; Visible Ear digital library	Sorenson et al.	2009

appropriate time and support.[37] The role of experienced surgical trainers in providing guidance and curriculum development is essential.

4.4. Conclusion

Virtual reality simulation is emerging as a powerful training tool that can expedite skills acquisition. It offers a potential solution to the challenge faced by program directors in delivering effective surgical training. Several VR platforms in temporal bone surgery and endoscopic sinus surgery are worthy of incorporation into training programs particularly in early years. Virtual reality simulation will never replace the role of real life operative experience. It does, however, allow for unlimited repetition, a better understanding of the surgical anatomy and can also facilitate surgical planning.

New technology will continue to emerge and trainers need to be actively involved to improve quality and outcome. This review provides background information of the VR platforms currently available. This will help programme directors who are looking to implement an effective VR integrated curriculum in Otolaryngology.

Ethical approval

N/A.

Funding

N/A.

Author contribution

Mr Asit Arora: Conception and design of study, acquisition and interpretation of data; drafting the article and revising it critically for important intellectual content; final approval of the version to be published.

Ms Loretta Lau: Data acquisition and interpretation; drafting the article and revising it critically for important intellectual content; final approval of the version to be published.

Mr Zaid Awad: Data analysis, revising it critically for important intellectual content; final approval of the version to be published.

Professor Ara Darzi: Conception and design, revising article critically and final approval of the version to be published.

Mr Arvind Singh: Data analysis, revising it critically for important intellectual content; final approval of the version to be published.

Mr Neil Tolley: Conception of study, revising the article critically for important intellectual content; final approval of the version to be published.

Conflict of interest

The first author of this review (Asit Arora) is also the author of one of the included studies.

References

1. Tooke J. *Aspiring to excellence: findings and recommendations of the independent inquiry into modernising medical careers*. MMC Inquiry; 2007.
2. *Good doctors, safer patients: proposals to strengthen the system to assure and improve the performance of doctors and to protect the safety of patients*. Department of Health; 2006.
3. Calatayud D, Arora S, Aggarwal R, et al. Warm-up in a virtual reality environment improves performance in the operating room. *Ann Surg* 2010;**251**(6): 1181–5.
4. Donaldson L. *Safer medical practice: machines, manikins and polo mints*. Department of Health; 2009.
5. Satava RM. Virtual reality surgical simulator. The first steps. *Surg Endosc* 1993;**7**(3):203–5.
6. Zerey MHA, Kercher W, Heniford BT. Minimally invasive surgery training overview. In: Park AKR, editor. *Minimally invasive surgery training: theories, models, outcomes*. National Institute of Health, National Library of Medicine; 2010.
7. Aggarwal R, Mytton OT, Derbrew M, et al. Training and simulation for patient safety. *Qual Saf Health Care* 2010;(19 Suppl. 2):i34–43. England.
8. Cannon-Bowers JA, Bowers C, Procci K. Optimizing learning in surgical simulations: guidelines from the science of learning and human performance. *Surg Clin North Am* 2010;**90**:583–603. United States: 2010 Elsevier Inc.
9. Carter FJ, Schijven MP, Aggarwal R, et al. Consensus guidelines for validation of virtual reality surgical simulators. *Simul Healthc* 2006;**1**(3):171–9.
10. Reddy-Kolanu G, Alderson D. Evaluating the effectiveness of the Voxel-Man TempoSurg virtual reality simulator in facilitating learning mastoid surgery. *Ann R Coll Surg Engl* 2011;**93**(3):205–8.
11. Arora A, Khemani S, Tolley N, et al. Face and content validation of a virtual reality temporal bone simulator. *Otolaryngol Head Neck Surg* 2012;**146**(3):497–503.
12. Francis HW, Malik MU, Diaz Voss Varela DA, et al. Technical skills improve after practice on virtual-reality temporal bone simulator. *Laryngoscope* 2012;**122**(6): 1385–91.
13. Nash R, Sykes R, Majithia A, Arora A, Singh A, Khemani S. Objective assessment of learning curves for the Voxel-Man TempoSurg temporal bone surgery computer simulator. *J Laryngol Otol* 2012;**126**(7):663–9.
14. Zirkle M, Roberson DW, Leuwer R, Dubrowski A. Using a virtual reality temporal bone simulator to assess otolaryngology trainees. *Laryngoscope* 2007;**117**(2):258–63.
15. Khemani S, Arora A, Singh A, Tolley N, Darzi A. Objective skills assessment and construct validation of a virtual reality temporal bone simulator. *Otol Neurotol* 2012;**33**(7):1225–31.
16. O'Leary SJ, Hutchins MA, Stevenson DR, et al. Validation of a networked virtual reality simulation of temporal bone surgery. *Laryngoscope* 2008;**118**(6):1040–6.
17. Zhao YC, Kennedy G, Hall R, O'Leary S. Differentiating levels of surgical experience on a virtual reality temporal bone simulator. *Otolaryngol Head Neck Surg* 2010;**143**(5 Suppl. 3):S30–5.
18. Zhao YC, Kennedy G, Yukawa K, Pyman B, O'Leary S. Improving temporal bone dissection using self-directed virtual reality simulation: results of a randomized blinded control trial. *Otolaryngol Head Neck Surg* 2011;**144**(3):357–64.
19. Zhao YC, Kennedy G, Yukawa K, Pyman B, Stephen OL. Can virtual reality simulator be used as a training aid to improve cadaver temporal bone dissection? Results of a randomized blinded control trial. *Laryngoscope* 2011;**121**(4):831–7.

20. Wiet GJ, Yagel R, Stredney D, et al. A volumetric approach to virtual simulation of functional endoscopic sinus surgery. *Stud Health Technol Inform* 1997;**39**:167–79.
21. Wiet GJ, Stredney D, Sessanna D, Bryan JA, Welling DB, Schmalbrock P. Virtual temporal bone dissection: an interactive surgical simulator. *Otolaryngol Head Neck Surg* 2002;**127**(1):79–83.
22. Wiet GJ, Stredney D, Kerwin T, et al. Virtual temporal bone dissection system: OSU virtual temporal bone system: development and Testing. *Laryngoscope* 2012;**122**(Suppl. 1):S1–12.
23. Sewell C, Morris D, Blevins NH, et al. Providing metrics and performance feedback in a surgical simulator. *Comput Aided Surg* 2008;**13**(2):63–81.
24. Rudman DT, Stredney D, Sessanna D, et al. Functional endoscopic sinus surgery training simulator. *Laryngoscope* 1998;**108**(11 Pt 1):1643–7.
25. Arora H, Uribe J, Ralph W, et al. Assessment of construct validity of the endoscopic sinus surgery simulator. *Arch Otolaryngol Head Neck Surg* 2005;**131**(3):217–21.
26. Fried MP, Sadoughi B, Weghorst SJ, et al. Construct validity of the endoscopic sinus surgery simulator: II. Assessment of discriminant validity and expert benchmarking. *Arch Otolaryngol Head Neck Surg* 2007;**133**:350–7. United States.
27. Fried MP, Sadoughi B, Gibber MJ, et al. From virtual reality to the operating room: the endoscopic sinus surgery simulator experiment. *Otolaryngol Head Neck Surg* 2010;**142**(2):202–7.
28. Edmond Jr CV. Impact of the endoscopic sinus surgical simulator on operating room performance. *Laryngoscope* 2002;**112**(7 Pt 1):1148–58.
29. Caversaccio M, Eichenberger A, Hausler R. Virtual simulator as a training tool for endonasal surgery. *Am J Rhinol* 2003;**17**(5):283–90.
30. Sowerby LJ, Rehal G, Husein M, Doyle PC, Agrawal S, Ladak HM. Development and face validity testing of a three-dimensional myringotomy simulator with haptic feedback. *J Otolaryngol Head Neck Surg* 2010;**39**(2):122–9.
31. Ho AK, Alsaffar H, Doyle PC, Ladak HM, Agrawal SK. Virtual reality myringotomy simulation with real-time deformation: development and validity testing. *Laryngoscope* 2012;**122**(8):1844–51.
32. Wheeler B, Doyle PC, Chandarana S, Agrawal S, Husein M, Ladak HM. Interactive computer-based simulator for training in blade navigation and targeting in myringotomy. *Comput Methods Programs Biomed* 2010;**98**(2):130–9.
33. Darzi A, Smith S, Taffinder N. Assessing operative skill. Needs to become more objective. *BMJ* 1999;**318**(7188):887–8.
34. Dosis A, Aggarwal R, Bello F, et al. Synchronized video and motion analysis for the assessment of procedures in the operating theater. *Arch Surg* 2005;**140**:293–9. United States.
35. Haque S, Srinivasan S. A meta-analysis of the training effectiveness of virtual reality surgical simulators. *IEEE Trans Inf Technol Biomed* 2006;**10**(1):51–8.
36. Ericsson KA. Deliberate practice and acquisition of expert performance: a general overview. *Acad Emerg Med* 2008;**15**:988–94. United States.
37. Wiet GJ, Stredney D, Wan D. Training and simulation in otolaryngology. *Otolaryngol Clin North Am* 2011;**44**(6):1333–50.