Original Study

Virtual reality in spinal endoscopy: a paradigm shift in education to support spine surgeons

Ryan Lohre[1], Jeffrey C. Wang[2], Kai-Uwe Lewandrowski[3,4], Danny P. Goel[5]

[1]Department of Orthopaedics, University of British Columbia, Vancouver, BC, USA; [2]USC Spine Center, Keck School of Medicine at University of Southern California, Los Angeles, USA; [3]Center for Advanced Spine Care of Southern Arizona and Surgical Institute of Tucson, Tucson, AZ, USA; [4]Department of Neurosurgery, UNIRIO, Rio de Janeiro, Brazil; [5]Department of Orthopaedics, University of British Columbia, Vancouver, BC, Canada

Contributions: (I) Conception and design: R Lohre, DP Goel; (II) Administrative support: All authors; (III) Provision of study materials or patients: All authors; (IV) Collection and assembly of data: R Lohre, DP Goel; (V) Data analysis and interpretation: R Lohre, DP Goel; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Dr. Danny P. Goel, MD, MSc, FRCSC. 321 Water Street, Suite 500, Vancouver, BC V6B 1B8, Canada.
Email: danny@precisionostech.com.

Background: Minimally invasive spine surgery (MISS) and endoscopic spine surgery have continually evolving indications in the cervical, thoracic, and lumbar spine. Endoscopic spine surgery entails treatment of disc disease, stenosis, spondylolisthesis, radiculopathy, and deformity. MISS involves complex motor skills in regions of variable anatomy. Simulator use has been proposed to aid in training and skill retention, preoperative planning, and intraoperative use.

Methods: A systematic review of five databases was performed for publications pertaining to the use of virtual (VR), augmented (AR), and mixed (MR) reality in MISS and spinal endoscopic surgery. Qualitative data analysis was undertaken with focus of study design, quality, and reported outcomes. Study quality was assessed using the Medical Education Research Quality Instrument (MERSQI) score and level of evidence (LoE) by a modified Oxford Centre for Evidence-Based Medicine (OCEBM) level for simulation in medicine.

Results: Thirty-eight studies were retained for data collection. Studies were of intervention-control, clinical application, and pilot or cross-sectional design. Identified articles illustrated use of VR, AR, and MR in all study designs. Procedures included pedicle cannulation and screw insertion, vertebroplasty, kyphoplasty, percutaneous transforaminal endoscopic discectomy (PTED), lumbar puncture and facet injection, transvertebral anterior cervical foraminotomy (TVACF) and posterior cervical laminoforaminotomy. Overall MERSQI score was low-to-medium [M =9.71 (SD =2.60); range, 4.5–13.5], and LoE was predominantly low given the number of purely descriptive articles, or low-quality randomized studies.

Conclusions: The current scope of VR, AR, and MR surgical simulators in MISS and spinal endoscopic surgery was described. Studies demonstrate improvement in technical skill and patient outcomes in short term follow-up. Despite this, overall study quality and levels of evidence remain low. Cohesive study design and reporting with focus on transfer validity in training scenarios, and patient derived outcome measures in clinical studies are required to further advance the field.

Keywords: Virtual reality (VR); simulator; spinal endoscopic surgery; minimally invasive spine surgery (MISS)

Submitted Nov 09, 2019. Accepted for publication Nov 13, 2019.
doi: 10.21037/jss.2019.11.16
View this article at: http://dx.doi.org/10.21037/jss.2019.11.16

Introduction

Simulation by means of virtual reality (VR) in neurosurgery and orthopaedic surgery for educational, preoperative planning, and intra-operative utilization continues to improve with technological advances in computer processing. Simulation is endorsed by numerous organizations including the American College of Surgeons, American Academy of Orthopaedic Surgeons, Authorité Hauté Santé in France, Congress of Neurological Surgeons, and Accreditation Council for Graduate Medical Education (ACGME) (1). Simulators for educational purposes continue to develop along the six core competencies laid by the ACGME to replicate clinical training scenarios in the face of flexible duty hours, heterogeneity of experience, and changing resident education. The number of publications dedicated to simulation in the fields of neurosurgery and orthopaedics has continued to dramatically rise in the preceding decade, with transition from bench top and low fidelity models to contemporary VR (2). VR utilizes a computer processing unit with a head-mounted display (HMD) to provide visual and auditory cues coupled with controllers containing position trackers and force feedback, or haptics, to provide an immersive, multisensory experience.

Additional areas of simulation include augmented reality (AR) or mixed reality (MR). AR combines computer processing and a see-through display that projects a virtual construct onto real-world imagery. VR and AR exist on a continuous spectrum of MR, from the completely digitized and simulated environment present in VR to real and simulated environments of AR, and combinations between referred to as MR. VR in spine surgery has seen greatest use in educational simulation or pre-operative planning (3). Spinal anatomy is complex and variable, and the ability to visualize and interact with a virtual patient's spine prior to surgical intervention has benefit in rehearsal and planning. The focus of AR in spine surgery has been for intraoperative visualization and instrumentation (4). Most of these technologies rely on software capable of receiving patient CT scans of adequate resolution (approximately 1.25-mm slice width or smaller) and producing interactive volume renderings in a three-dimensional (3D) environment. VR for educational purposes or pre-operative planning in spine surgery and orthopaedic surgery has lagged behind other surgical subspecialties as well as commercial industry, with multiple reviews delineating the lack of standardized measures, low levels of evidence and levels of recommendation despite consensus statements from working directives and organizations (5-8). Furthermore, cost-effectiveness of these VR constructs has not been conclusively demonstrated in spine surgery, though has been alluded to in other areas of orthopaedic surgery.

Paralleling the rise in popularity of VR is minimally invasive spine surgery (MISS) and endoscopic spine surgery for treatment of spinal stenosis, degenerative disk disease, compression fractures, tumor or lesion ablation, and most recently for adult spinal deformity (ASD) correction (9). In a period of 1997 to 2017, the number of articles relating to endoscopic spine surgery increased 41 times (10). The rate of MISS procedures is approaching that of open procedures given increasing surgeon familiarity and potential benefits including faster recovery, reduced blood loss, shorter hospital stays, and outpatient settings (10,11). Endoscopic spine surgery requires correct localization, cannulation, and continuous irrigation while directly visualizing structures. Depending on the region of pathology, various portal techniques or endoscope docking and free-hand techniques may be employed (12). Endoscopic techniques require significant surgeon skill, and learning curves have been described as steep (13). Multi-centre, multi-year studies of percutaneous endoscopic lumbar discectomy (PELD) demonstrate continuous improvements in surgical time (14). Complication rates from dural tear, infection, epidural hematoma, and dysesthesias have been described as increased in early periods of endoscopic surgical adoption (15-17). Similarly, the pedicle screw placement learning curve has been estimated at 80 screws, or 25 cases to reach an asymptote of technique skill (18). Fluoroscopic time and radiation exposure are significantly higher with novice surgeons (10).

Further advances in technology with increasing evidence of MISS and endoscopic spine surgical benefit will promote incorporation of these procedures into regular practice. Despite the use of more technological instrumentation, surgeon recognition of the importance of clinical outcomes and patient safety is paramount. VR, AR, and MR simulators stand to provide uninhibited practice of surgical techniques for training purposes, and may benefit surgeons clinically for preoperative planning, and intraoperative use. Ultimately, VR, AR, and MR simulators could improve adoption of endoscopic spine surgery by traditionally trained spine surgeons, who often struggle with the adoption of spinal endoscopy due to its technical difficulties and expensive non-portable stand-alone

simulators. Therefore, we sought to determine the current evidence of the use of VR, AR, and MR simulators in MISS and spinal endoscopic surgery, including study quality, level of evidence (LoE), and outcomes. Furthermore, we wished to include recommendations for future studies, and considerations for forthcoming research and simulator development.

Methods

A systematic review of current available literature pertaining to VR and endoscopic/MISS was performed according to the guidelines for reporting systematic reviews (PRISMA).

Search strategy

The search strategy incorporated databases Medline OVID, Embase, PsychINFO, Cochrane reviews, and Google Scholar. Search strategy was completed in stages and combined using Boolean operators. Stage 1 incorporated terms relating to VR, MR or AR or computer-assisted surgery. Stage 2 incorporated educational terms or simulation terms. Stage 3 incorporated terms relating to neurosurgery or spine surgery. Stage 4 incorporated terms for minimally invasive, percutaneous, or endoscopic surgery. Lastly, Stage 5 incorporated all stage terms so that only articles containing these subjects were retained for initial review. Terms used included Medical Subject Headings (MeSH) for comprehensiveness. Secondary searches were conducted specifically for known commercially available VR suites including "PrecisionOS Technology", "PHANToM", "Dextroscope", "Surgical Rehearsal Platform", "Falcon", "Procedicus VIST Simulator", "ImmersiveTouch", and "Perk Tutor".

Inclusion criteria included English language and studies with primary data reporting of educational or simulation in minimally invasive/endoscopic/percutaneous spinal surgery with measurable outcomes. All relevant empirical study designs were included including randomized controlled trials, non-randomized controlled trials, prospective cohort studies, retrospective studies or cross-sectional studies. All percutaneous/endoscopic/minimally invasive manual surgical spine procedures were included including discectomy, fusion, foraminotomy and nerve decompression, and vertebroplasty. Pedicle screw insertion studies were included. Studies included required mention of preoperative planning or for educational purposes using VR/AR/MR simulators, or for intraoperative use. Incorporated participants in studies included all levels of training and all simulation devices pertaining to VR, MR, or AR. Exclusion criteria included neurosurgical studies incorporating brain or soft tissue surgery simulators, non-manual studies such as radiotherapy application, non-MISS studies, or studies not incorporating VR, AR, or MR.

Following initial search, title screening was performed by a single study member (R Lohre) for appropriateness of inclusion. Additional records identified through other sources were added to the initial search pool. Duplicates were screened and removed prior to abstract review. Abstract review was then performed by a single study member (R Lohre). Full articles retained were reviewed by study members (R Lohre and DP Goel) and any discrepancy on inclusion resolved by consensus. After full text review, the studies incorporated were searched for additional references manually. PubMed search term notifications were set up, as well as through the Mendeley platform for any new articles published during the manuscript draft period up to time of submission (November 2019).

Data extraction

Data extraction was performed by a single study member (R Goel). Study demographic data including authors, location of study, publication, simulated MISS procedure, simulator type and brand if available were extracted. Next, study parameters were extracted including type of study, aims or hypothesis, participants, validity assessments if any, outcome measures and main conclusions. Retained studies were examined for methodological quality using the validated Medical Education Research Study Quality Instrument (MERSQI). Scoring was provided via the MERSQI range of 5–18, with higher scores denoting higher quality study design, sampling, data type, assessment, validity measurements, data analysis and outcome measures and reporting. If assessments of each characteristic were not possible for a study, scores remained in keeping with standard MERSQI reporting with a highest value of 18. MERSQI scores reported are of the primary authors interpretation. LoE was provided for each study and was gathered either through publication reporting, or determined by study member (R Lohre) based on the modified Oxford Centre for Evidence-Based Medicine (OCEBM) guidelines proposed by Carter *et al.* for VR surgical simulators (19). A meta-analysis was not performed given the heterogeneity in study design, outcome measures and reporting variability.

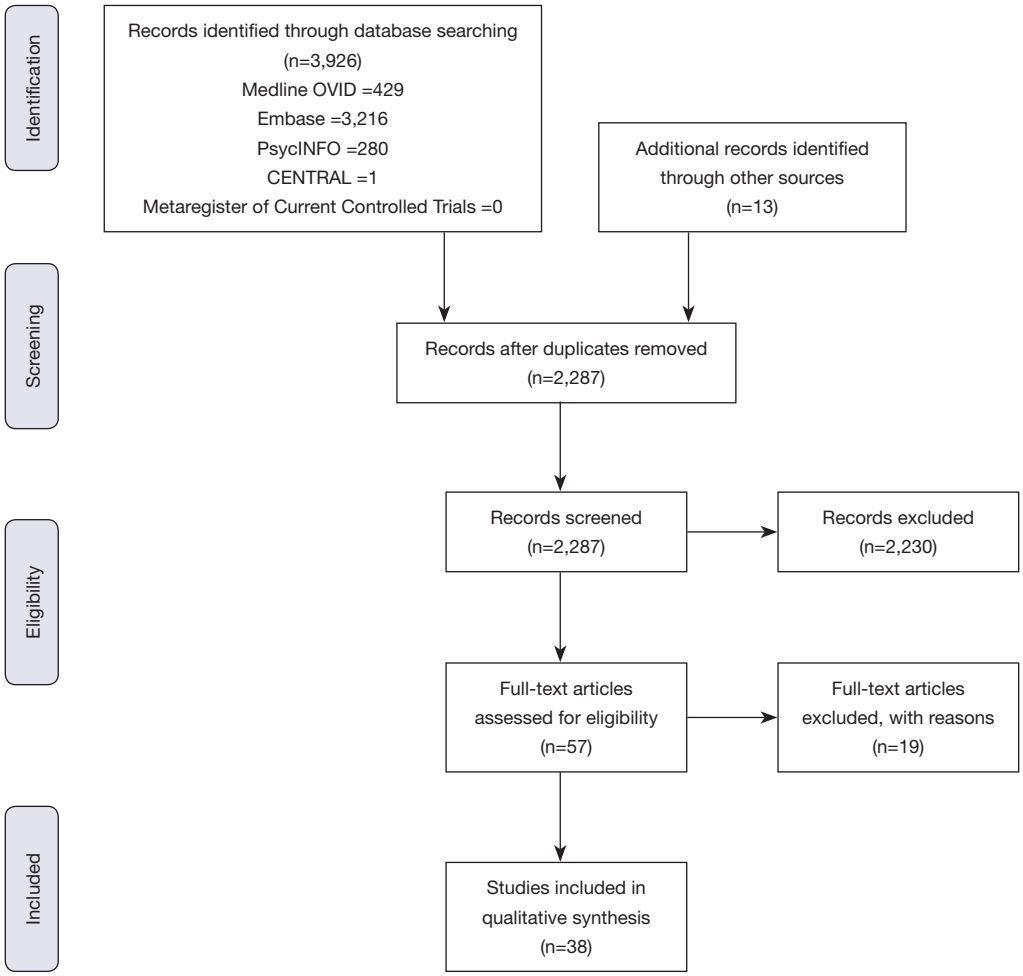

Figure 1 Search strategy and selection process of included studies.

Results

Following PRISMA guidelines, 2,287 unique publications were screened to produce 57 for full text review. Nineteen articles were further excluded prior to qualitative synthesis. Articles were excluded for not involving VR, AR, or MR (n=7), incorporating open procedures rather than MISS (n=5), incorporating other soft tissue structures such as brain (n=4), or not providing outcomes (n=3). Thirty-eight studies were included in the review, while a meta-analysis was not performed due to heterogeneity of study designs and reported outcomes. *Figure 1* depicts the screening and selection process for the systematic review.

Table 1 depicts characteristics of representative included studies including journal of publication, location of experiment, type of MISS or pedicle insertion procedure simulated, spectrum of simulator type, brand names of simulators and availability if described, and MERSQI and LoE scoring. For a full list of studies, please see *Tables S1,S2*. The earliest study was seen to occur in 2009 (33). Though no publication year cut-off was used during searching, the majority of articles (n=20; 52.6%) were published within 2 years prior to time of review writing. Eleven studies (28.9%) were conducted in China, 7 (18.4%) in the USA, 7 (18.4%) in Germany, 4 (10.5%) in Canada, 3 (7.9%) in Sweden, 2 (5.3%) in the UK, 2 (5.3%) in Japan, 1 (2.6%) in Finland, and 1 (2.6%) in Italy. Ten (90.9%) of the articles published in China were within 2 years of review writing. Fourteen (36.8%) of the studies specified VR simulators, 11 (28.9%) utilized AR simulators, and 10 (26.3%) utilized a combination of components designating MR. The majority (n=8; 57.1%) of VR simulators were unspecified, or of proprietary design and not commercially available. Other

Table 1 List of representative studies of available simulators and spine procedure

Study	Publication	Geographic location	Simulated procedure	Type of simulator	Simulator brand	MERSQI score	Level of evidence
Gibby *et al.* (20)	*Int J Comput Assist Radiol Surg*, 2019	District of Columbia, USA	Vertebral pedicle cannulation	AR—3D volume rendered images of spine model/ phantom	Microsoft HoloLens (Redmond, WA, USA), Novarad OpenSight (American Fork, UT, USA)	8.0	3
Luciano *et al.* (21)	*Neurosurgery*, 2013	Illinois, USA	Percutaneous needle insertion/pedicle cannulation	AR—volume rendering of single patient spine	ImmersiveTouch (San Francisco, CA, USA)	12.0	2c
Umebayashi *et al.* (22)	*J Am Acad Orthop Surg Glob Res Rev*, 2018	Aichi, Japan	Transvertebral anterior cervical foraminotomy (TVACF) and posterior cervical laminoforaminotomy	AR—digital overlay of intraoperative CT data to microscope	Medtronic StealthStation S7	7.0	3
Deib *et al.* (23)	*J Neurointerv Surg*, 2018	Maryland, USA	Vertebroplasty, kyphoplasty, percutaneous discectomy	AR—volume rendering of spine model/phantom	Microsoft HoloLens (Redmond, WA, USA)	4.5	3
Archavlis *et al.* (24)	*World Neurosurg*, 2016	Mainz, Germany	Microsurgical endoscopic assisted transpedicular corpectomy of the thoracic spine	VR—surface and volume renderings of patient spine CT data	Amira (FEI Visualization Sciences Group, Mérignac Cedex, France) and Dextroscope (Bracco Group, Kent Ridge Digital Labs, Singapore)	8.5	3
Gottschalk *et al.* (25)	*Spine J*, 2015	USA	Cervical lateral mass screw via Magerl technique	MR—simulated drill navigation and cadavers	Stealth 3D Navigation Unit, Medtronic, Minneapolis, MN, USA PixelStick, Plum Amazing LLC, Princeville, HI, USA	13.5	2a
Hu *et al.* (26)	*Int J Surg*, 2017	Shanghai, China	Percutaneous endoscopic lumbar discectomy	VR—volume renderings of lumbar spine	Boholo, Fengsuan Inc., Shanghai, China	11.0	2a
Yu *et al.* (27)	*World Neurosurg*, 2019	Shanghai, China	PTED	VR—volume rendering of lumbar spine	3D Slicer platform (http://www.slicer.org)	9.5	2a
Wei *et al.* (28)	*J Orthop Surg Res*, 2019	Nanjing, China	Percutaneous kyphoplasty	MR—volume rendering of spine and trajectory planning, with AR glasses and overlay	Baholo, Shanghai Front Computing Company, China; Medivi, Changzhou, China; Hololens, Microsoft, USA	11.5	2a
Chitale *et al.* (29)	*Neurosurgery*, 2013	MN, USA	Percutaneous lumbar pedicle screw placement	MR—volume rendered spine with phantom lumbar spine	Medtronic Surgical Technologies	10.0	3

Table 1 (*continued*)

Table 1 (*continued*)

Study	Publication	Geographic location	Simulated procedure	Type of simulator	Simulator brand	MERSQI score	Level of evidence
Weigl *et al.* (30)	*Surg Endosc*, 2016	Munich, Germany	Percutaneous vertebroplasty	MR—real OR with procedural mannequin and VR volume rendered lumbar spine	Gaumard HAL S2001 Mannequin Novint Falcon (Novint Technologies, Inc., USA)	13.5	2a
Wucherer *et al.* (31)	*Int J Comput Assist Radiol Surg*, 2014	Munich, Germany	Percutaneous vertebroplasty	MR—real OR with procedural mannequin and AR assisted CT images	Novint Falcon (Novint Technologis, Inc., USA)	8.0	3
Wucherer *et al.* (32)	*IEEE Trans Med Imaging*, 2015	Germany	Percutaneous vertebroplasty	MR—real OR with procedural mannequin and AR assisted fluoroscopic images	Novint Falcon (Novint Technologis, Inc., USA)	13.5	2b

AR, augmented reality; 3D, three-dimensional; CT, computed tomography; VR, virtual reality; MR, mixed reality; PTED, percutaneous transforaminal endoscopic discectomy; OR, operating room.

VR simulators used included Simulation and Visualization Research Group (n=1; 7.1%), NeurosimVR (Calgary, AB, Canada) (n=1; 7.1%), ImmersiveTouch (San Francisco, CA, USA) (n=2; 14.3%), and Boholo (Fengsuan Inc., Shanghai, China) (n=1; 7.1%). Three (27.3%) of the AR simulators used were unspecified. Two (18.2%) studies utilized Microsoft HoloLens (Redmond, WA, USA) for AR visualization and proprietary software, 2 (18.2%) utilized ImmersiveTouch (San Francisco, CA, USA), 1 utilized Virtual Protractor with Augmented Reality (VIPAR) though is not commercially available, one utilized Medtronic StealthStation, one utilized Perk Station (The Perk Lab, Queen's University, Canada), and one utilized a proprietary tracking system using Micron Tracker2 and a graphics user interface. Studies utilizing MR were varied in utilized simulators, with 2 (20.0%) not specifying simulators used. The most frequently used MR simulator was that of the Novint Falcon (Novint Technologis, Inc., USA), utilized in 3 (30.0%) of MR studies.

Pedicle screw insertion and cannulation for percutaneous approaches were both included in this review. Eighteen (47.4%) of studies involved localization of pedicles, and cannulation or insertion of pedicle screws, in general. Of these studies, the predominant location was that of lumbar vertebrae, seen in 10 (55.6%) studies. Two (11.1%) pedicle screw insertion studies involved the thoracic spine (24,34), 1 (5.6%) study involved cervical pedicle screw insertion (35), 1 (5.6%) involved cervical lateral mass screw insertion (36), 1 (5.6%) involved thoracolumbar pedicle screw insertion (27), and 1 (5.6%) was not specified and utilized a sheep cadaver (37). Seven studies (18.4%) pertained to vertebroplasty or kyphoplasty simulation (20,27,29,31,32,38,39). Two (5.3%) studies utilized simulators for transpedicular percutaneous endoscopic discectomy (TPED) (40). One (2.6%) study examined both vertebroplasty/kyphoplasty and discectomy using an AR system (27). Seven (18.4%) studies examined simulator use for needle localization of lumbar puncture (LP) or facet injections (25,33,41-45), with five studies utilizing AR or MR, and a single using a VR system. One (2.6%) study examined an AR simulator for use in transvertebral anterior cervical foraminotomy (TVACF) and posterior cervical laminoforaminotomy (30). Another single study (2.6%) examined the use of VR in microsurgical endoscopic assisted transpedicular corpectomy of the thoracic spine (26).

Table 2 illustrates design, hypotheses, aims, and outcomes of retained studies. Seven (18.4%) studies involved use

Table 2 List of studies by design, aims, validity assessments, and outcome measures

Study	Study design	Study aims or hypothesis	Study participants	Validity assessments	Outcome measures	Summary of results
Gibby *et al.*, 2019 (20)	One group post-test design with varying levels of training	Introduction of an AR guidance system for training	Two medical students, one neuroradiologist and one orthopaedic surgeon	Not completed	Time to needle placement, AR registration error, extrapolated needle position in pedicle	AR provided safe trajectories and intuitive insertion of needles
Luciano *et al.*, 2013 (21)	One group pre- and post-test	Evaluation of teaching effectiveness of AR in needle localization	63 fellows and residents attending AANS	Face validity—author developed	Failure rate of localization, performance accuracy of needle placement	Failure rate improved, performance accuracy improved (no significance testing)
Umebayashi *et al.*, 2018 (22)	Case series of TVACF and posterior cervical laminoforaminotomy	Pilot study for feasibility of AR in microscopic MISS	Two representative case examples	Not completed	Feasibility of intraoperative use and patient follow-up to 20 months	Intraoperative AR in microscopic MISS for TVACF and posterior cervical laminoforaminotomy was safe in two cases with no symptom recurrence
Deib *et al.*, 2018 (23)	One group/single user intervention	Pilot study for feasibility of AR use instead of angiography suite monitors	One interventional radiologist—three representative case examples using spine models/phantoms repeated four times	Not completed	Procedural times, beam time and dose time of HMD compared to traditional angiography suite monitors, and user preference	Similar procedural times, beam time and dose time between visualization methods (no statistics presented) User felt HMD was unobtrusive
Archavlis *et al.*, 2016 (24)	Cohort comparison of endoscopic assisted and mini open corpectomy with VR preoperative planning	Pilot feasibility study of VR preoperative planning for endoscopic and mini-open transpedicular corpectomy	Seven cases—two unstable burst fractures and five metastatic disease	Not completed	Comparison of degree of bone removal, distance from critical structures, and implant diameter of final surgery compared to VR preoperative planning	Preoperative parameters were met in all cases and surgeons identified VR preoperative planning with the use of endoscopic assistance as beneficial (no comparison of VR; used in all cases)
Gottschalk *et al.*, 2015 (25)	Intervention-control group	Effectiveness of MR surgical simulation training on novice trainees for lateral mass screw placement in cervical spine	15 orthopaedic surgery residents PGY1-6 randomized to three groups (group 1 no feedback, group 2 and 3 receiving 3D navigational feedback)	Face validity—author developed	Primary—aggregate mean difference from a "perfect" Magerl screw Secondary—adjacent structure injuries (facet violations, nerve or arterial)	3D navigation training in a MR setting significantly improved cervical lateral mass screw insertion

Table 2 (*continued*)

Table 2 (*continued*)

Study	Study design	Study aims or hypothesis	Study participants	Validity assessments	Outcome measures	Summary of results
Hu *et al.*, 2017 (26)	Intervention-control group	Effectiveness of VR planning in PELD for patient outcomes and surgical technique	40 patients receiving L4/5 or L5/S1 PELD (20 assigned to planning group and 20 assigned to conventional group)	Not completed	Technique outcomes: channel establishment time, operative time, fluoroscopic time Patient outcomes: 10-point visual analog pain scale, Oswestry Disability Index (ODI), modified Macnab's criteria for satisfaction, complications	Reduced channel establishment times, operative times, and fluoroscopic times with VR planning group compared to conventional. No difference in patient outcomes
Yu *et al.*, 2019 (27)	Intervention-control group	Efficacy of MR training of PTED for novice trainees	60 novice surgical residents and one experienced consultant. Residents were assigned to intervention (MR training, n=30) and control (2D-training, n=30).	Face validity—author developed	User satisfaction of MR training, puncture times, total operative time, fluoroscopy times	Face validity was demonstrated via questionnaire. Time to puncture, total operative time, and total fluoroscopy times were reduced
Wei *et al.*, 2019 (28)	Intervention-control group	Clinical outcomes of MR assisted percutaneous kyphoplasty	40 cases of osteoporotic vertebral compression fracture randomized into treatment with MR (n=20) or with traditional fluoroscopy (n=20)	Not completed	Operation time, fluoroscopy time, amount of PMMA injected, relative vertebral height (anterior/posterior height ratio), relative central vertebral height (center/posterior height ratio), change in vertebral kyphotic angle (KA), cement-both-endplates-contact, VAS and ODI scores at 1 year	The group receiving MR guidance had improvements of all technical parameters and improved patient VAS and ODI scores immediately and at all follow-up intervals to 1 year
Chitale *et al.*, 2013 (29)	Pre-test and post-test/intervention	Effectiveness of an MR training curriculum for neurosurgery residents in percutaneous pedicle screw placement	Eight residents completed a pre-intervention test, then a didactic learning session, then completed an MR pedicle insertion followed by a post-test.	Not completed	Fluoroscopy score (starting point and trajectory, fluoroscopy time, number of fluoroscopic shots), computed tomography score (time and starting point and trajectory), test score	A non-significant improvement in all domains was seen using the training model

Table 2 (*continued*)

Table 2 (*continued*)

Study	Study design	Study aims or hypothesis	Study participants	Validity assessments	Outcome measures	Summary of results
Weigl *et al.*, 2015 (30)	Comparison of two groups receiving different surgical disruptions	Determination of surgical disruption on surgeon performance and perceived workload	Nineteen junior surgeons were randomized to two disruption scenarios during vertebroplasty and following scenario, questionnaires were completed (SURG-TLX)	Not completed	SURG-TLX scores (mental workload), performance outcomes (trocar deviation, length of tooltip trajectory, fluoroscopy exposure time, overall duration, total number of fluoroscopic shots)	Surgical disruptions produced significantly higher SURG-TLX scores and were associated with poorer performance metrics, notably total number of fluoroscopic images
Wucherer *et al.*, 2014 (31)	Comparison of two groups receiving surgical disruption	Development of an MR training environment	Five surgeons performed an MR vertebroplasty with crisis scenarios/task disruptions	Face validity—author developed Construct validity—comparison to expert	None	Face validity of an MR training environment for surgical task disruptions in vertebroplasty was demonstrated and construct validity was attempted through single expert user
Wucherer *et al.*, 2015 (32)	Comparison of two groups performing MR vertebroplasty and receiving surgical disruption	Assessment of MR environment for usability in training vertebroplasty, and workload during crisis/surgical disruption	Nineteen junior surgeons performed MR vertebroplasty with crisis scenario/task disruptions	Face validity—author developed	SURG-TLX scores (mental workload), performance outcomes [root-mean-square deviation (RMSD), fluoroscopic exposure time, procedural time]	The users felt the task was realistic via face validity questionnaire. MR training enabled significantly faster procedure completion times

AR, augmented reality; AANS, American Association of Neurological Surgeons; TVACF, transvertebral anterior cervical foraminotomy; MISS, minimally invasive spine surgery; HMD, head-mounted display; VR, virtual reality; MR, mixed reality; 3D, three-dimensional; PELD, percutaneous endoscopic lumbar discectomy; PTED, percutaneous transforaminal endoscopic discectomy; VAS, Visual Analog Scale; PMMA, polymethylmethacrylate.

of simulators in patients (22,24,26,28,34,38,46). A total composite of patients involved in VR/AR/MR trials was n=123. The longest series of follow-up was to 20 months following a case series of two patients receiving TVACF and posterior cervical laminoforaminotomy (22). Follow-up was regarding symptom recurrence without validated patient outcome metrics. Two studies presented patient interval follow-up with recording of patient outcome metrics including Visual Analog Scale (VAS), Oswestry Disability Index (ODI), and modified McNab criteria (26,28). One of the aforementioned studies pertained to percutaneous lumbar discectomy procedure (PLED) of L4/5 and L5/S1 in 40 patients using a VR system for preoperative planning. The use of VR planning showed reduced technical times with equivalent patient outcomes at 6-month follow-up (26). The other aforementioned study pertained to MR use during percutaneous kyphoplasty. Forty patients similarly received randomization and treatment, with the MR group showing improvement in all technical parameters [operating time, fluoroscopy time, polymethylmethacrylate (PMMA) used, anterior/posterior (A/P) height ratios, kyphotic angle (KA) change, cement-both-endplates-contact, as well as VAS and ODI scores immediately post-op and at follow-up intervals to 1 year] (28).

Study aims or hypotheses varied considerably. Broad categories included determination of effectiveness of simulator platforms for training use, a pilot study for use in training or clinical scenarios, and clinical effectiveness and intervention studies. Twenty-four (63.2%) studies involved determination of simulator effectiveness in training, with 11 (45.8%) of these studies utilizing VR training systems, 8 (33.3%) utilizing MR training systems, and 5 (20.8%) studies utilizing AR training systems. Eleven (28.9%) studies specified pilot studies in introduction of VR/AR/MR systems for training or clinical use. Three (7.9%) studies specifically sought to compare clinical effectiveness of VR/AR/MR. Twenty-three (60.5%) studies were performed using an intervention-control design. These studies compared the use of VR (n=12, 52.2%), AR (n=3, 13.0%), and MR (n=8, 34.8%) use in preoperative planning and intra-operative use relative to established standards of training. Three intervention-control studies utilized disruption scenarios in an MR simulator operating room (OR) and did not specifically compare MR to other simulation methods (30-32). Four studies (10.5%) utilized a pre- and post-test design scheme (20,21,29,47), three of which utilized AR systems (20,21,47), and one MR system (29). The mean number of study participants in all studies was 16.4 (SD =17.5; range, 1–63). Study participants ranged from design technicians, undergraduate students, medical students, residents, orthopaedic surgeons, and neurosurgeons.

Twelve studies (31.6%) (21,25,27,31-33,36,39,41,42,48,49) attempted to establish validity in simulator use. Validity measures included demonstration of face, construct, and content validity. There were no studies that demonstrated transfer validity of educational or training simulators of VR, AR, or MR design to real OR scenarios. Ten of the 12 studies (83.3%) demonstrated face validity of the simulators via non-validated questionnaire responses of varying number of questions and Likert-responses (21,25,27,31-33,36,39,41,49). Two of the 10 studies also examined construct and content validity (31,33). One study examined the construct validity of a VR simulator for pedicle screw insertion, while another examined content validity of a VR simulator for pedicle screw insertion (42,48). All studies examining face validity demonstrated positive outcomes regarding realism of the system used compared to the simulated task. Three studies examined handling and functionality of simulator systems via non-validated questionnaires (32,33,49).

Quality of included studies was examined using the MERSQI score and LoE. There was a variation in MERSQI scores in included studies, with an average M =9.71 (SD =2.60; range, 4.5–13.5) out of 18. Individual MERSQI domains of "sampling" M =0.59 (SD =0.25; range, 0.5–1) and "validity" M=0.65 (SD =1.03; range, 0–3) had the lowest mean scores. Comparison of individual MERSQI domains is limited given variability in scale ranges and study characteristics. The LoE reported was based on Carter *et al.* modification of OCEBM guidelines, published by the Work Group for Evaluation and Implementation of Simulators and Skills Training Programmes, European Association of Endoscopic Surgeons (EAES). Eighteen (47.4%) studies achieved an LoE of 3 based on nonrandomized, noncomparative, descriptive trials. Fourteen (36.8%) studies were designated a LoE of 2a, for providing reasonable quality randomized trials, though that did not meet or provide sample size calculations. One (2.6%) article achieved an LoE of 1b, and was the highest LoE study included in the review for providing an adequate powered, randomized trial (48). A correlation of MERSQI score to LoE was not completed for included studies.

Discussion

The use of simulators in orthopaedic surgery training has continued to increase with support of regulatory bodies

and mounting evidence of knowledge retention and skill improvement. MISS and spinal endoscopic surgery have particular reliance on complicated trajectories and uncoupling of hand-eye cues similar to that of arthroscopic procedures. Reports of high complication rates in early adoption may adversely steer established surgeons from performing these procedures. As evidence grows for simulation in endoscopic surgical training, production of simulators that are portable, cost-effective, and enjoyable to use cannot be underscored. Our systematic review highlighted 38 studies involving VR, AR, and MR simulation in MISS including pedicle screw insertion. These studies were performed globally and were relatively equivalent in terms of volume of VR, AR, or MR system used. There was a clear lack of commercially available simulators, with the majority being developed using proprietary instrumentation or software.

Intervention-control studies were predominantly utilized for VR, AR and MR simulator studies. In these studies, the digital simulator trained groups routinely outperformed the control groups in measured parameters except in one (39). Though a meta-analysis could not be performed to aggregate data, descriptive analysis demonstrates that simulator training improved both knowledge and technical skill of learners, including novice and expert (orthopaedic and neurosurgeons) populations. This was gathered through objective outcomes analysis of user error rates, improvement of technical skills, time to completion of tasks, or fluoroscopic usage.

Additional identified research presented pilot studies, cross-sectional studies, or clinical outcome studies for the use of VR, AR, and MR on patients. Two studies, incorporating follow-up of a total of 80 patients at 6 months and 1 year following VR and MR use in planning and implementation of PLED and kyphoplasty demonstrated superiority in the VR/MR groups over controls (26,28). There were no intraoperative or immediate complications of any patient highlighted in the clinical use publications. Currently there are few publications pertaining to the use of VR, AR, and MR clinically for MISS and spinal endoscopy and evidence remains limited, however as more centres become familiar with available technology, the number of clinical outcome studies is likely to rise. Future studies should attempt standardization of clinical outcomes, supporting well-conducted randomized trials of VR, AR, and MR use in spinal endoscopy and MISS. These outcomes should combine radiographic parameters with patient-reported outcome measures at regular intervals with minimum follow-up of 2 years. Standardized reporting of these trials is also encouraged to allow for direct comparison to other studies in systematic reviews and meta-analyses as research continues.

Overall, study quality was demonstrated as low-to-medium based on MERSQI scores. The mean score was lower than other reviews attempting the same scoring system, and may be attributed to scope of review and inclusion criteria. Our study highlights a greater number of papers than other reviews, and is novel in presenting simulator application for MISS surgery. Furthermore, LoE was determined for all included papers based on a modified OCEBM framework. Overall LoE was low based on descriptive studies. Most intervention-control study designs were of moderate quality based on this modified framework. Studies presented had a number of objective limitations including small population sizes, lack of clearly defined outcomes, absent statistical analyses, limited validity breadth including demonstration of transfer validity, and lack of clinical application. As is the case in other areas of orthopaedics, VR, AR, and MR simulators in MISS education have not demonstrated clear transfer validity to real OR environments. Studies in other surgical disciplines have employed robust study design to accomplish this, and it is likely that this will be attempted in MISS given the surge in recent publications and simulator availability. Moving forward, consensus documents should be produced and adhered to for development and validation of MISS immersive simulators. A standardized framework would be effective in providing clear research protocols to multiple centres, precipitating larger and better-quality studies. For intervention-control, and randomized trials comparing simulator modalities, we advocate for attempting to produce higher quality trials with power and sample size calculations and well-defined hypotheses. The outcomes should be supported by appropriate descriptive statistics. Study designs should allow for prospective, long-term follow-up of skill retention, in multiple levels of trainee or established expert, to establish efficacy of training. As validity of simulators in MISS is lacking, we also recommend adherence to guidelines of establishing face, content, construct, and transfer validity. Transfer validity to real operative scenarios would definitively establish simulator training effectiveness in MISS and spinal endoscopy (50). Finally, reporting requires standardization to allow for aggregate data collection to provide higher levels of evidence and support meta-analyses and we recommend the recently updated reporting guidelines for health care simulation research (51).

There are a number of identifiable limitations of the presented review. First, only English articles were collected and presented. There is an inherent bias to published outcomes of simulators presented in this article, with no identified negative outcomes of simulator use, and only one equivalency study presented (39). Further sources of unaccounted for bias include financial disclosures and vested interest in simulator products by authors, and reporting of multiple experiments with the same group of learners. These articles were included in the review as they presented different data or aspects of the experiment, however impart bias with experimental repetition and thus affected MERSQI scoring (30,31,52-54). Following search strategy, data aggregation and scoring via MERSQI and LoE was performed by a single reviewer, thus potentially incorporating bias and preventing inter-rater reliability calculations. Evaluation of study quality with MERSQI has previously been used in spine simulation studies, however has inconsistencies based on reviewer and interpretation of criteria, thus our use of a single reviewer potentially expounds these inconsistencies in reporting (8). Additional use of the Newcastle-Ottawa Scale-Education (NOS-E) may be considered in future analyses as this has been validated previously (35).

Beyond novice surgeons, tracking of user data, individualization and updating cases, and ability to perform surgeries virtually prior to a real environment stand to benefit practicing and expert surgeons. Regarding development of simulators, efforts should be made to produce multi-use, adaptable designs to replicate real world OR scenarios in spinal endoscopy and MISS. Adaptability in software modules, hardware, and control schemes will allow for the growing indication of spinal endoscopic procedures and allow for growth within the specialty. New avenues of research with VR, AR, and MR systems include collection of large data with artificial intelligence applications to develop expert sourced recommendations for trajectory of implants or implant types. This may be adaptable to endoscopic approach, pedicle trajectories, or osteotomy for deformity correction.

Sample surgical training workflow

We provide a sample iterative workflow for established spine surgeons wishing to train in contemporary endoscopic spine surgery. PrecisionOS Technology™ as a representative example, provides an entirely immersive, multi-sensory OR environment for training. *Figure 2* provides images of the virtual suite once the surgeon places the heads up display (HUD) on themselves. Complementing the sensory experience is ability to practice both clinical decision-making and real, technical skill on representative stock cases, or import additional cases (available in 2020). Computing software allows for tracking of progress through stock cases, both in error and success rates, and technical outcomes utilizing a well-known concept in education referred to as deliberate practice (55).

Figure 2 Immersive virtual reality simulator for spinal endoscopic surgical training. Image courtesy of PrecisionOS Technology.

A surgeon using the HUD and haptic controllers would select one of several representative cases appropriately indicated for endoscopic surgery and subsequently perform this case in immersive VR (iVR). The equipment is free of wires and can be used in any environment. Uninhibited repetition with variation, also referred to as interleaving is one of several paramount learning nuances created in this technology (56). The degrees of freedom of the techniques, ability to fail and a guide-mode with direct feedback is an available option to aid in early learning, which may be turned on, or off as experience grows. Akin to other avenues of training in orthopaedics or neurosurgery, or high-risk industries such as aviation, the surgeon can practice numerous cases with realistic anatomy, localization of tool or implant trajectories, and receive immediate feedback. *Figure 3* demonstrates needle localization in a virtual patient. The surgeon can then perform the real operation knowing a clear plan, having performed a similar operation in a realistic, multi-sensory manner as many times as desired preoperatively. As the surgeon increases his/her endoscopic caseload, the iVR system may be tailored to increasing complexity of cases, and aggregate user outcome data in virtual cases can be compared to real-life patient outcomes. Our review has demonstrated evidence in the use of VR systems for training of both novices and established surgeons in endoscopic procedures as well as clinical cases.

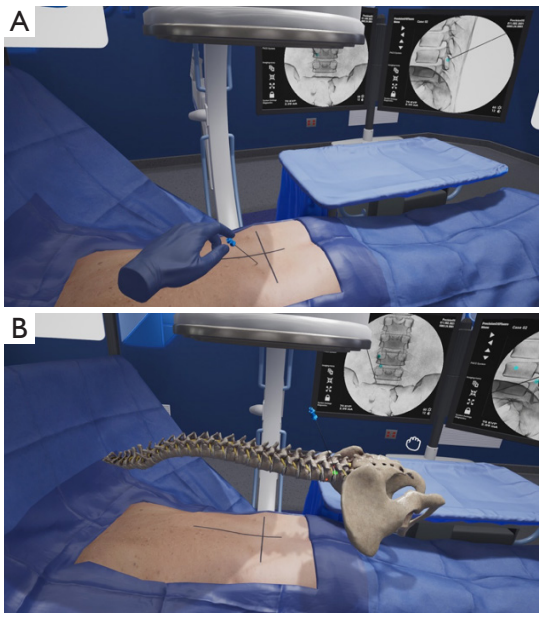

Figure 3 Needle localization in immersive virtual reality (A) with ability to visualize anatomy and plan trajectories not possible in other simulator modalities (B). Image courtesy of PrecisionOS Technology.

Following the above outline for endoscopic procedural adoption may provide a cost-effective, enjoyable, and safe avenue of changing practice. We have previously performed a blinded randomized controlled trial demonstrating 570% efficiency in training time for those using iVR when compared to didactic or pre-surgical planning using written surgical techniques (R Lohre and DP Goel, submitted) and superior technical skills with equivalent verbal and written knowledge scoring. Furthermore, a unique option available in iVR, is that of "multiplayer". This feature permits an expert surgeon to educate novice surgeons (regardless of geographical location) within the same virtual OR and within the convenience of their home city. The time and cost efficiency imparted through its use is further augmented by the ability to research, collaborate, and educate through this modality.

Conclusions

The presented review highlights 38 studies incorporating VR, AR, and MR in MISS and spinal endoscopic surgery. Published studies have utilized this technology for surgical training and clinical application for less than 10 years. Virtual simulators routinely outperformed traditional methods of training for MISS procedures including pedicle screw insertion in both novice and expert cohorts. Surgical simulators have been used to improve clinical results in early follow-up for both vertebroplasty and PLED with reported patient outcome measures. Overall included study quality was low to medium, with limited LoE and no direct transfer validity to real scenarios demonstrated for training studies. Higher quality, randomized studies with clear objectives, longer term results, and standardized reporting are required to more clearly demonstrate the benefit of virtual simulators in MISS and spinal endoscopic surgery. VR, AR, and MR simulators in spinal endoscopic surgery stand to complement surgical training programs, preoperative planning and intra-operative use. Strong consideration to evaluate and use iVR for a comprehensive evaluation including patient positioning, variation in cases and patient specific planning is currently available and may provide a cost efficient and scalable method to introduce simulation internationally with cross collaboration and case sharing.

Acknowledgments

None.

Footnote

Conflicts of Interest: Dr. DP Goel is the CEO of PrecisionOS Technology, a virtual reality company. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Disclaimer: The views expressed in this editorial represent those of the authors and no other entity or organization. This manuscript is not meant for or intended to endorse any products or push any other agenda other than to report the associated clinical outcomes with use of endoscopy versus laser. The motive for compiling this clinically relevant information is by no means created and/or correlated to directly enrich anyone due to its publication. This publication was intended to substantiate contemporary endoscopic spinal surgery concepts to facilitate technology advancements.

References

1. Aïm F, Lonjon G, Hannouche D, et al. Effectiveness of Virtual Reality Training in Orthopaedic Surgery. Arthroscopy 2016;32:224-32.
2. Vaughan N, Dubey VN, Wainwright TW, et al. A review of virtual reality based training simulators for orthopaedic surgery. Med Eng Phys 2016;38:59-71.
3. Guha D, Alotaibi NM, Nguyen N, et al. Augmented Reality in Neurosurgery: A Review of Current Concepts and Emerging Applications. Can J Neurol Sci 2017;44:235-45.
4. Burström G, Nachabe R, Persson O, et al. Augmented and Virtual Reality Instrument Tracking for Minimally Invasive Spine Surgery: A Feasibility and Accuracy Study. Spine 2019;44:1097-104.
5. Bernardo A. Virtual Reality and Simulation in Neurosurgical Training. World Neurosurg 2017;106:1015-29.
6. Konakondla S, Fong R, Schirmer CM. Simulation training in neurosurgery: advances in education and practice. Adv Med Educ Pract 2017;8:465-73.
7. Oliveira LM, Figueiredo EG. Simulation Training Methods in Neurological Surgery. Asian J Neurosurg 2019;14:364-70.
8. Pfandler M, Lazarovici M, Stefan P, et al. Virtual reality-based simulators for spine surgery: a systematic review. Spine J 2017;17:1352-63.
9. Wewel JT, Godzik J, Uribe JS. The utilization of minimally invasive surgery techniques for the treatment of spinal deformity. J Spine Surg 2019;5:S84-S90.
10. Iprenburg M, Wagner R, Godschalx A, et al. Patient radiation exposure during transforaminal lumbar endoscopic spine surgery: a prospective study. Neurosurg Focus 2016;40:E7.
11. Butler AJ, Alam M, Wiley K, et al. Endoscopic Lumbar Surgery: The State of the Art in 2019. Neurospine 2019;16:15-23.
12. Hwa Eum J, Hwa Heo D, Son SK, et al. Percutaneous biportal endoscopic decompression for lumbar spinal stenosis: a technical note and preliminary clinical results. J Neurosurg Spine 2016;24:602-7.
13. Jacquot F, Gastambide D. Percutaneous endoscopic transforaminal lumbar interbody fusion: is it worth it?. Int Orthop 2013;37:1507-10.
14. Shi R, Wang F, Hong X, et al. Comparison of percutaneous endoscopic lumbar discectomy versus microendoscopic discectomy for the treatment of lumbar disc herniation: a meta-analysis. Int Orthop 2019;43:923-37.
15. Marappan K, Jothi R, Paul Raj S. Microendoscopic discectomy (MED) for lumbar disc herniation: comparison of learning curve of the surgery and outcome with other established case studies. J spine Surg 2018;4:630-7.
16. Nowitzke AM. Assessment of the learning curve for lumbar microendoscopic discectomy. Neurosurgery 2005;56:755-62.
17. Soliman HM. Irrigation endoscopic discectomy: a novel percutaneous approach for lumbar disc prolapse. Eur Spine J 2013;22:1037-44.
18. Gonzalvo A, Fitt G, Liew S, et al. The learning curve of pedicle screw placement: how many screws are enough? Spine (Phila Pa 1976) 2009;34:E761-5.
19. Carter FJ, Schijven MP, Aggarwal R, et al. Consensus guidelines for validation of virtual reality surgical simulators. Simul Healthc 2006;1:171-9.
20. Gibby JT, Swenson SA, Cvetko S, et al. Head-mounted display augmented reality to guide pedicle screw placement utilizing computed tomography. Int J Comput Assist Radiol Surg 2019;14:525-35.
21. Luciano CJ, Banerjee PP, Sorenson JM, et al. Percutaneous spinal fixation simulation with virtual reality and haptics. Neurosurgery 2013;72 Suppl 1:89-96.
22. Umebayashi D, Yamamoto Y, Nakajima Y, et al. Augmented Reality Visualization-guided Microscopic Spine Surgery: Transvertebral Anterior Cervical Foraminotomy and Posterior Foraminotomy. J Am Acad Orthop Surg Glob Res Rev 2018;2:e008.
23. Deib G, Johnson A, Unberath M, et al. Image guided percutaneous spine procedures using an optical see-through head mounted display: proof of concept and rationale. J Neurointerv Surg 2018;10:1187-91.
24. Archavlis E, Schwandt E, Kosterhon M, et al. A Modified Microsurgical Endoscopic Assisted Transpedicular Corpectomy of the Thoracic Spine Based on Virtual 3D Planning, Technical Note. World Neurosurg 2016;91:424-33.
25. Gottschalk MB, Yoon ST, Park DK, et al. Surgical training using three-dimensional simulation in placement of cervical lateral mass screws: a blinded randomized control trial. Spine J 2015;15:168-75.
26. Hu Z, Li X, Cui J, et al. Significance of preoperative planning software for puncture and channel establishment in percutaneous endoscopic lumbar DISCECTOMY: A study of 40 cases. Int J Surg 2017;41:97-103.
27. Yu H, Zhou Z, Lei X, et al. Mixed Reality-Based Preoperative Planning for Training of Percutaneous Transforaminal Endoscopic Discectomy: A Feasibility Study. World Neurosurg 2019;129:e767-75.

28. Wei P, Yao Q, Xu Y, et al. Percutaneous kyphoplasty assisted with/without mixed reality technology in treatment of OVCF with IVC: a prospective study. J Orthop Surg Res 2019;14:255.
29. Chitale R, Ghobrial GM, Lobel D, et al. Simulated Lumbar Minimally Invasive Surgery Educational Model With Didactic and Technical Components. Neurosurgery 2013;73:107-10.
30. Weigl M, Stefan P, Abhari K, et al. Intra-operative disruptions, surgeon's mental workload, and technical performance in a full-scale simulated procedure. Surg Endosc 2016;30:559-66.
31. Wucherer P, Stefan P, Weidert S, et al. Task and crisis analysis during surgical training. Int J Comput Assist Radiol Surg 2014;9:785-94.
32. Wucherer P, Stefan P, Abhari K, et al. Vertebroplasty Performance on Simulator for 19 Surgeons Using Hierarchical Task Analysis. IEEE Trans Med Imaging 2015;34:1730-7.
33. Färber M, Hummel F, Gerloff C, et al. Virtual reality simulator for the training of lumbar punctures. Methods Inf Med 2009;48:493-501.
34. Abe Y, Sato S, Kato K, et al. A novel 3D guidance system using augmented reality for percutaneous vertebroplasty: technical note. J Neurosurg Spine 2013;19:492-501.
35. Cook DA, Reed DA. Appraising the quality of medical education research methods: the Medical Education Research Study Quality Instrument and the Newcastle-Ottawa Scale-Education. Acad Med 2015;90:1067-76.
36. Koch A, Pfandler M, Stefan P, et al. Say, What Is on Your Mind? Surgeons' Evaluations of Realism and Usability of a Virtual Reality Vertebroplasty Simulator. Surg Innov 2019;26:234-43.
37. Ma L, Zhao Z, Chen F, et al. Augmented reality surgical navigation with ultrasound-assisted registration for pedicle screw placement: a pilot study. Int J Comput Assist Radiol Surg 2017;12:2205-15.
38. Elmi-Terander A, Burström G, Nachabe R, et al. Pedicle Screw Placement Using Augmented Reality Surgical Navigation With Intraoperative 3D Imaging: A First In-Human Prospective Cohort Study. Spine (Phila Pa 1976) 2019;44:517-25.
39. Kulcsár Z, O'Mahony E, Lövquist E, et al. Preliminary evaluation of a virtual reality-based simulator for learning spinal anesthesia. J Clin Anesth 2013;25:98-105.
40. Zhou Z, Hu S, Zhao YZ, et al. Feasibility of Virtual Reality Combined with Isocentric Navigation in Transforaminal Percutaneous Endoscopic Discectomy: A Cadaver Study. Orthop Surg 2019;11:493-9.
41. Mostafa AE, Ryu WHA, Chan S, et al. Designing NeuroSimVR: A Stereoscopic Virtual Reality Spine Surgery Simulator. 2017. doi: http://dx.doi.org/10.11575/PRISM/31003.
42. Xiang L, Zhou Y, Wang H, et al. Significance of preoperative planning simulator for junior surgeons' training of pedicle screw insertion. J Spinal Disord Tech 2015;28:E25-9.
43. Keri Z, Sydor D, Ungi T, et al. Computerized training system for ultrasound-guided lumbar puncture on abnormal spine models: a randomized controlled trial. Can J Anaesth 2015;62:777-84.
44. Moult E, Ungi T, Welch M, et al. Ultrasound-guided facet joint injection training using Perk Tutor. Int J Comput Assist Radiol Surg 2013;8:831-6.
45. Yeo CT, Ungi T, U-Thainual P, et al. The effect of augmented reality training on percutaneous needle placement in spinal facet joint injections. IEEE Trans Biomed Eng 2011;58:2031-7.
46. Naddeo F, Cataldo E, Naddeo A, et al. An automatic and patient - specific algorithm to design the optimal insertion direction of pedicle screws for spine surgery templates. Med Biol Eng Comput 2017;55:1549-62.
47. Luciano CJ, Banerjee PP, Bellotte B, et al. Learning retention of thoracic pedicle screw placement using a high-resolution augmented reality simulator with haptic feedback. Neurosurgery 2011;69:ons14-9; discussion ons19.
48. Rambani R, Ward J, Viant W. Desktop-based computer-assisted orthopedic training system for spinal surgery. J Surg Educ 2014;71:805-9.
49. Sutherland C, Hashtrudi-Zaad K, Sellens R, et al. An augmented reality haptic training simulator for spinal needle procedures. IEEE Trans Biomed Eng 2013;60:3009-18.
50. Van Nortwick SS, Lendvay TS, Jensen AR, et al. Methodologies for establishing validity in surgical simulation studies. Surgery 2010;147:622-30.
51. Cheng A, Kessler D, Mackinnon R, et al. Reporting Guidelines for Health Care Simulation Research: Extensions to the CONSORT and STROBE Statements. Simul Healthc 2016;11:238-48.
52. Hou Y, Lin Y, Shi J, et al. Effectiveness of the Thoracic Pedicle Screw Placement Using the Virtual Surgical Training System: A Cadaver Study. Oper Neurosurg (Hagerstown) 2018;15:677-85.
53. Hou Y, Shi J, Lin Y, et al. Virtual surgery simulation versus traditional approaches in training of residents in cervical

pedicle screw placement. Arch Orthop Trauma Surg 2018;138:777-82.
54. Shi J, Hou Y, Lin Y, et al. Role of Visuohaptic Surgical Training Simulator in Resident Education of Orthopedic Surgery. World Neurosurg 2018;111:e98-104.
55. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Acad Med 2004;79:S70-81.
56. Henz D, Schöllhorn WI. Differential Training Facilitates Early Consolidation in Motor Learning. Front Behav Neurosci 2016;10:199.

Cite this article as: Lohre R, Wang JC, Lewandrowski KU, Goel DP. Virtual reality in spinal endoscopy: a paradigm shift in education to support spine surgeons. J Spine Surg 2020;6(Suppl 1):S208-S223. doi: 10.21037/jss.2019.11.16

Supplementary

Table S1 List of studies by simulated procedure and type of simulator

Study	Publication	Geographic location	Simulated procedure	Type of simulator	Simulator brand	MERSQI score	Level of evidence
Xiang *et al.* (42)	*J Spinal Disord Tech*, 2015	Chongqing, China	Thoracolumbar pedicle screw insertion	VR—3D volume rendering with projection fluoroscopy simulator of patient spine	Proprietary cross-platform simulator written in C++	9.0	2b
Gibby *et al.* (20)	*Int J Comput Assist Radiol Surg*, 2019	District of Columbia, USA	Vertebral pedicle cannulation (MISS)	AR—3D volume rendered images of spine model/phantom	Microsoft HoloLens (Redmond, WA, USA), Novarad OpenSight (American Fork, UT, USA)	8.0	3
Luciano *et al.* (47)	*Neurosurgery*, 2011	Illinois, USA	Thoracic pedicle screw placement	AR—3D volume rendering of single patient spine	ImmersiveTouch (San Francisco, CA, USA)	11.0	2c
Luciano *et al.* (21)	*Neurosurgery*, 2013	Illinois, USA	Percutaneous needle insertion/ pedicle cannulation	AR—volume rendering of single patient spine	ImmersiveTouch (San Francisco, CA, USA)	12.0	2c
Burström *et al.* (4)	*Spine*, 2019	Stockholm, Sweden	Percutaneous vertebral pedicle cannulation (MISS)	AR—*in vivo* cannulation of porcine vertebral pedicles in hybrid OR using CBCT	Not described	10.0	3
Umebayashi *et al.* (22)	*J Am Acad Orthop Surg Glob Res Rev*, 2018	Aichi, Japan	Transvertebral anterior cervical foraminotomy (TVACF) and posterior cervical laminoforaminotomy (MISS)	AR—digital overlay of intraoperative CT data to microscope	Medtronic StealthStation S7	7.0	3
Deib *et al.* (23)	*J Neurointerv Surg*, 2018	Maryland, USA	Vertebroplasty, kyphoplasty, percutaneous discectomy (MISS)	AR—volume rendering of spine model/ phantom	Microsoft HoloLens (Redmond, WA, USA)	4.5	3
Archavlis *et al.* (24)	*World Neurosurg*, 2016	Mainz, Germany	Microsurgical endoscopic assisted transpedicular corpectomy of the thoracic spine (MISS)	VR—surface and volume renderings of patient spine CT data	Amira (FEI Visualization Sciences Group, Mérignac Cedex, France) and Dextroscope (Bracco Group, Kent Ridge Digital Labs, Singapore)	8.5	3
Gasco *et al.* (57)	*Neurol Res*, 2014	Texas, USA	Placement of lumbar spine pedicle screws	VR—volume renderings of lumbar spine	ImmersiveTouch (San Francisco, CA, USA)	13.5	2a
Kulcsár *et al.* (39)	*J Clin Anesth*, 2013	Limerick, Ireland	Lumbar puncture	MR—volume rendering of lumbar spine	Sensable, Wilmington, DE, USA H3D and Volume Haptics Toolkit (VHTK), SenseGraphics, Krista, Sweden	12.0	2a
Gottschalk *et al.* (25)	*Spine J*, 2015	USA	Cervical lateral mass screw via Magerl technique	MR—simulated drill navigation and cadavers	Stealth 3D Navigation Unit, Medtronic, Minneapolis, MN, USA PixelStick, Plum Amazing LLC, Princeville, HI, USA	13.5	2a
Hu *et al.* (26)	*Int J Surg*, 2017	Shanghai, China	Percutaneous endoscopic lumbar discectomy (MISS)	VR—volume renderings of lumbar spine	Boholo, Fengsuan Inc., Shanghai, China	11.0	2a
Shi *et al.* (54)	*World Neurosurg*, 2018	Shanghai, China	Lumbar pedicle screw placement	VR—volume renderings of lumbar spine	Unspecified. "Virtual Surgery Training System" (VSTS)	9.0	2a
Choque-Velasquez *et al.* (58)	*World Neurosurg*, 2018	Finland	Microsurgical skills (i.e., microsuturing)	MR—VR HMD with camera showing real time micro suturing via an "eye hands blind" (EHB) technique	Unspecified VR glasses, Mac (Apple Inc., Cupertino, CA, USA), iPhone (Apple Inc., Cupertino, CA, USA), TeamViewer software, Reality Augmented Software	5.5	3
Xin *et al.* (59)	*World Neurosurg*, 2019	Shanghai, China	Thoracolumbar pedicle screw placement	VR—volume rendering of thoracolumbar spine	Unspecified VR system, UG NX8.0, Seimens, Munich, Germany	12.0	2a
Yu *et al.* (27)	*World Neurosurg*, 2019	Shanghai, China	PTED (MISS)	VR—volume rendering of lumbar spine	3D Slicer platform (http://www.slicer.org)	9.5	2a
Wei *et al.* (28)	*J Orthop Surg Res*, 2019	Nanjing, China	Percutaneous kyphoplasty (MISS)	MR—volume rendering of spine and trajectory planning, with AR glasses and overlay	Baholo, Shanghai Front Computing Company, China; Medivi, Changzhou, China; Hololens, Microsoft, USA	11.5	2a
Elmi-Terander *et al.* (38)	*Spine (Phila Pa 1976)*, 2019	Stockholm, Sweden	Lumbosacral pedicle screw placement	MR—volume rendered spine with VR preoperative planning and AR intraoperative workflow	Not specified—not commercially available	12.5	3
Hou *et al.* (53)	*Arch Orthop Trauma Surg*, 2018	Shanghai, China	Cervical pedicle screw placement	VR—volume rendered cervical spine	Unspecified. "Virtual Surgery Training System" (VSTS)	10.0	2a
Hou *et al.* (52)	*Oper Neurosurg*, 2018	Shanghai, China	Thoracic pedicle screw placement	VR—volume rendered thoracic spine	Unspecified. "Virtual Surgery Training System" (VSTS)	10.0	2a
Liu *et al.* (60)	*IEEE*, 2017	Shenyang, China	Lumbar pedicle screw placement	MR—volume rendered virtual model, 3D-printed model and infrared (IR) camera instrument tracking	Unspecified "Minimally invasive spine system training" (MISST)	5.5	3
Mostafa *et al.* (41)	Technical report—University of Calgary Publication, 2017	Calgary, AB, Canada	Lumbar pedicle screw placement	VR—volume rendered thoracolumbar spine	NeurosimVR, ImmersiveTouch (San Francisco, CA, USA)	4.5	3
Naddeo *et al.* (46)	*Med Biol Eng Comput*, 2017	Italy	Lumbar pedicle screw placement	MR—volume rendered lumbar spine with production of patient specific drilling templates	In-house software designed using Rhinoceros 3D (Seattle, WA, USA)	7.0	3
Ma *et al.* (37)	*Int J Comput Assist Radiol Surg*, 2017	Beijing, China	Unspecified pedicle screw placement	AR—digital overlay of CT data with US data	Unspecified developed surgical navigation system	6.5	3
Elmi-Terander *et al.* (61)	*Spine (Phila Pa 1976)*, 2019	Sweden	Thoracolumbar pedicle localization and screw insertion	AR—volume rendered spine with hybrid OR navigation	Not specified—not commercially available	8.0	3
Zhou *et al.* (40)	*Orthop Surg*, 2019	Shanghai, China	Lumbosacral TPED (MISS)	VR—volume rendered spine for preoperative planning and isocentric navigation	Not specified	9.0	3
Keri *et al.* (43)	*Can J Anaesth*, 2015	Kingston, ON, Canada	Lumbar puncture	MR—Volume rendered spine with US localization	Lumbar Puncture Simulator II (Kyoto Kagaku), Perk Tutor	12.5	2a
Abe *et al.* (34)	*J Neurosurg Spine*, 2013	Hokkaido, Japan	Percutaneous vertebroplasty (MISS)	AR—volume rendered spine with AR assisted needle path in spine phantom	Virtual Protractor with Augmented Reality (VIPAR)—not commercially available	11.5	3
Chitale *et al.* (29)	*Neurosurgery*, 2013	MN, USA	Percutaneous lumbar pedicle screw placement (MISS)	MR—volume rendered spine with phantom lumbar spine	Medtronic Surgical Technologies	10.0	3
Färber *et al.* (33)	*Methods Inf Med*, 2009	Hamburg, Germany	Lumbar puncture	VR—volume rendered spine	Sensable Phantom Premium 1.5	10.5	2b
Moult *et al.* (44)	*Int J Comput Assist Radiol Surg*, 2013	Kingston, ON, Canada	Percutaneous facet joint injection	MR—volume rendered lumbar spine with US localization	Perk Tutor, SonixTouch US system with SonixGPS (Ascension™)	12.5	2a
Rambani *et al.* (48)	*J Surg Educ*, 2014	United Kingdom	Lumbar pedicle screw insertion	VR—volume rendered lumbar spine	Simulation and Visualization Research Group modified to VR	12.5	1b
Sutherland *et al.* (49)	*IEEE Trans Biomed Eng*, 2013	Kingston, Ontario, Canada	Percutaneous LP	AR—volume rendered spine overlay on phantom	Torso Mannequin, Micron Tracker2 optical tracking system, PHANToM haptic device graphical user interface	7.0	3
Weigl *et al.* (30)	*Surg Endosc*, 2016	Munich, Germany	Percutaneous vertebroplasty (MISS)	MR—real OR with procedural mannequin and VR volume rendered lumbar spine	Gaumard HAL S2001 Mannequin Novint Falcon (Novint Technologies, Inc., USA)	13.5	2a
Wucherer *et al.* (31)	*Int J Comput Assist Radiol Surg*, 2014	Munich, Germany	Percutaneous vertebroplasty (MISS)	MR—real OR with procedural mannequin and AR assisted CT images	Novint Falcon (Novint Technologis, Inc., USA)	8.0	3
Wucherer *et al.* (32)	*IEEE Trans Med Imaging*, 2015	Germany	Percutaneous vertebroplasty (MISS)	MR—real OR with procedural mannequin and AR assisted fluoroscopic images	Novint Falcon (Novint Technologis, Inc., USA)	13.5	2b
Koch *et al.* (36)	*Surg Innov*, 2019	Munich, Germany	Percutaneous vertebroplasty (MISS)	VR—volume rendered spine with haptic controllers with simulated fluoroscopy	Unspecified novel VR vertebroplasty simulator	11.0	3
Yeo *et al.* (45)	*IEEE Trans Biomed Eng*, 2011	Munich, Germany	Percutaneous facet injection	AR—volume rendered image of patient spine overlaid on phantom model	Perk Station (The Perk Lab, Queen's University, Canada)	12.5	2a

MERSQI, Medical Education Research Quality Instrument; VR, virtual reality; 3D, three-dimensional; MISS, minimally invasive spine surgery; AR, augmented reality; CT, computed tomography; MR, mixed reality; HMD, head-mounted display; PTED, percutaneous transforaminal endoscopic discectomy; US, ultrasound; OR, operating room; TPED, transpedicular percutaneous endoscopic discectomy; LP, lumbar puncture.

Table S2 List of studies by design, aims, validity assessments, and outcome measures

Study	Study design	Study aims or hypothesis	Study participants	Validity assessments	Outcome measures	Summary of results
Xiang *et al.*, 2015 (42)	Intervention-control group	Determination of benefit of training platform	Group 1—two junior surgeons Group 2—two senior spine surgeons	Content validity	Time to pedicle screw insertion, pedicle breach grading	Training on the VR simulator improved time to pedicle screw insertion and reduced breach
Gibby *et al.*, 2019 (20)	One group post-test design with varying levels of training	Introduction of an AR guidance system for training	Two medical students, one neuroradiologist and one orthopaedic surgeon	Not completed	Time to needle placement, AR registration error, extrapolated needle position in pedicle	AR provided safe trajectories and intuitive insertion of needles
Luciano *et al.*, 2011 (47)	One group pre- and post-test	Evaluation of learning retention of AR in pedicle screw insertion	51 fellows and residents attending American Association of Neurological Surgeons (AANS)	Not completed	Failure rate of localization, performance accuracy of screw trajectory	No change in localization failure rate, improvement of performance accuracy after training
Luciano *et al.*, 2013 (21)	One group pre- and post-test	Evaluation of teaching effectiveness of AR in needle localization	63 fellows and residents attending AANS	Face validity—author developed	Failure rate of localization, performance accuracy of needle placement	Failure rate improved, performance accuracy improved (no significance testing)
Burström *et al.*, 2019 (4)	One group intervention	Pilot and feasibility study of AR in percutaneous pedicle cannulation	2 spine surgeons	Not completed	Navigation time, instrumentation accuracy compared to planned VR path (entry point, device tip, axial and sagittal angular deviation)	Hammering and drilling using hybrid AR, VR planning and pedicle cannulation is accurate and feasible
Umebayashi *et al.*, 2018 (22)	Case series of TVACF and posterior cervical laminoforaminotomy	Pilot study for feasibility of AR in microscopic MISS	Two representative case examples	Not completed	Feasibility of intraoperative use and patient follow-up to 20 months	Intraoperative AR in microscopic MISS for TVACF and posterior cervical laminoforaminotomy was safe in two cases with no symptom recurrence
Deib *et al.*, 2018 (23)	One group/single user intervention	Pilot study for feasibility of AR use instead of angiography suite monitors	One interventional radiologist—three representative case examples using spine models/phantoms repeated four times	Not completed	Procedural times, beam time and dose time of HMD compared to traditional angiography suite monitors, and user preference	Similar procedural times, beam time and dose time between visualization methods (no statistics presented) User felt HMD was unobtrusive
Archavlis *et al.*, 2016 (24)	Cohort comparison of endoscopic assisted and mini open corpectomy with VR preoperative planning	Pilot feasibility study of VR preoperative planning for endoscopic and mini-open transpedicular corpectomy	Seven cases—two unstable burst fractures and five metastatic disease	Not completed	Comparison of degree of bone removal, distance from critical structures, and implant diameter of final surgery compared to VR preoperative planning	Preoperative parameters were met in all cases and surgeons identified VR preoperative planning with the use of endoscopic assistance as beneficial (no comparison of VR; used in all cases)
Gasco *et al.*, 2014 (57)	Intervention-control group	Effectiveness of VR training to place lumbar pedicle screws	26 medical students interested in neurosurgical residency programs	Not completed	Number of errors (length, coronal error, breach)	Simulation trained group demonstrated significant total error reduction
Kulcsár *et al.*, 2013 (39)	Prospective interventional-control group	Effectiveness of VR training in lumbar puncture for novice trainees	27 medical students within 12 months of graduation—randomly assigned to intervention (14 students) or conventional (13 students)	Face validity—author developed	Multiple choice written examination, global rating scale and task-specific checklist for both control and intervention groups, repeated clinical global rating scale and task-specific checklist via video review by two reviewers	No significant differences in global rating scale or task-specific scoring or knowledge testing between simulator trained and conventional trained novice medical students
Gottschalk *et al.*, 2015 (25)	Intervention-control group	Effectiveness of MR surgical simulation training on novice trainees for lateral mass screw placement in cervical spine	15 orthopaedic surgery residents PGY1-6 randomized to three groups (group 1 no feedback, group 2 and 3 receiving 3D navigational feedback)	Face validity—author developed	Primary—aggregate mean difference from a "perfect" Magerl screw Secondary—adjacent structure injuries (facet violations, nerve or arterial)	3D navigation training in a MR setting significantly improved cervical lateral mass screw insertion
Hu *et al.*, 2017 (26)	Intervention-control group	Effectiveness of VR planning in PELD for patient outcomes and surgical technique	40 patients receiving L4/5 or L5/S1 PELD (20 assigned to planning group and 20 assigned to conventional group)	Not completed	Technique outcomes: channel establishment time, operative time, fluoroscopic time Patient outcomes: 10-point visual analog pain scale, Oswestry Disability Index (ODI), modified Macnab's criteria for satisfaction, complications	Reduced channel establishment times, operative times, and fluoroscopic times with VR planning group compared to conventional. No difference in patient outcomes
Shi *et al.*, 2018 (54)	Intervention-control group	Effectiveness of VR training platform for teaching lumbar pedicle screw insertion	10 inexperienced residents assigned to intervention (n=5) or control (n=5) for pedicle screw insertion training prior to insertion in cadaver	Not completed	Screw penetration rate, acceptable screw placement by three raters (completely in pedicle or non-medial wall penetration <2 mm), screw penetration distance	Reduced number of screw penetrations, increased acceptable screw placement by author criteria, and decreased screw penetration distance by group trained using VR system
Choque-Velasquez *et al.*, 2018 (58)	Single group cross-sectional	Ease of use of novel VR microsurgical training system	Single user (neurosurgeon with 3 years of practice experience) performing 5 consecutive micro-sutures at baseline, 3 and 5 months	Not completed	Time to perform 5 consecutive sutures in seconds	The single user showed improved time to completion of task with repeated use but did not control for any variables
Xin *et al.*, 2018 (59)	Intervention-control group	Efficacy of VR training for thoracolumbar pedicle screw placement	16 novice surgical trainees randomized to intervention group (VR trained) or control (spine model demonstration) and repeated three times	Not completed	Screw accuracy (number of screws without breach), screw acceptance (<25% of screw diameter breach and no anterior cortex perforation), mean time of screw insertion	Improved positional accuracy of screws, acceptable screws based on author criteria, and faster insertion time in the VR trained group compared to control
Yu *et al.*, 2019 (27)	Intervention-control group	Efficacy of MR training of PTED for novice trainees	60 novice surgical residents and one experienced consultant. Residents were assigned to intervention (MR training, n=30) and control (2D-training, n=30).	Face validity—author developed	User satisfaction of MR training, puncture times, total operative time, fluoroscopy times	Face validity was demonstrated via questionnaire. Time to puncture, total operative time, and total fluoroscopy times were reduced
Wei *et al.*, 2019 (28)	Intervention-control group	Clinical outcomes of MR assisted percutaneous kyphoplasty	40 cases of osteoporotic vertebral compression fracture randomized into treatment with MR (n=20) or with traditional fluoroscopy (n=20)	Not completed	Operation time, fluoroscopy time, amount of PMMA injected, relative vertebral height (anterior/posterior height ratio), relative central vertebral height (center/posterior height ratio), change in vertebral kyphotic angle (KA), cement-both-endplates-contact, VAS and ODI scores at 1 year	The group receiving MR guidance had improvements of all technical parameters and improved patient VAS and ODI scores immediately and at all follow-up intervals to 1 year
Elmi-Terander *et al.*, 2019 (38)	One group/single intervention	Efficacy of hybrid MR operating room with VR preoperative planning of lumbosacral pedicle screw placement	Twenty consecutive cases with 253 lumbosacral pedicle screws inserted by single surgeon	Not completed	Screw placement accuracy based on Gertzbein criteria, number of severely misplaced screws (Gerzbein grade 3), average screw placement time, intra-operative or immediate post-operative adverse events	94.1% screw placement accuracy with no severely misplaced screws. The authors deemed an acceptable screw navigation time though offered no comparison or standard. There were no reported adverse events
Hou *et al.* (*Arch Orthop Trauma Surg*, 2018) (53)	Intervention-control group	Efficacy of VR training platform for teaching cervical pedicle screw placement to novice residents using cadavers	Ten novice residents assigned to VR intervention (n=5) and control (n=5) groups	Not completed	Screw penetration rates, screw acceptance rates (no pedicle penetration or penetration <50% diameter), quantified screw penetration distance	Significant improvement in screw penetration rates between VR trained (10%) and control (62.5%) (P<0.05), screw acceptance (100% in VR *vs.* 50% in control, P<0.05), and penetration distance (1.12±0.47 mm for VR *vs.* 2.08±0.39 mm, P<0.05)
Hou *et al.* (*Oper Neurosurg*, 2018) (52)	Intervention-control group	Efficacy of VR training platform for teaching thoracic pedicle screw placement to novice residents using cadavers	Ten novice residents assigned to VR intervention (n=5) and control (n=5) groups	Not completed	Screw penetration rates, screw acceptance rates (no pedicle penetration or penetration <2 mm), quantified screw penetration distance	Significant improvement in screw penetration rates between VR trained (7.14%) and control (30%) (P<0.05), screw acceptance (100% in VR *vs.* 92.86% in control, P<0.05), and penetration distance (1.23±0.56 mm for VR *vs.* 2.37±0.23 mm, P<0.05)
Liu *et al.*, 2017 (60)	One group single case	Pilot feasibility study of MISS MR pedicle screw trainer using single learner and expert spine surgeon	One novice and one expert surgeon using MISS MR system	Not completed	Computer calculated user score based on screw trajectory	Single case example demonstrates proof of concept and more appropriate screw trajectory from expert surgeon. User scores provided were not described
Mostafa *et al.*, 2017 (41)	Single group intervention-control	Pilot study to assess usability of novel VR simulator for pedicle screw placement	One group of surgeons (residents and consultants, n=6) and another of design experts (n=6) performing pedicle screw placement in NeurosimVR and ImmersiveTouch VR for comparison	Face validity—author developed	Five-point Likert-scale questionnaire pertaining to: difficulty level, skill/feedback, realistic haptics, repetitive practice potential, visual quality, hints/guidance, individualized learning potential, objective performance measures	Design experts preferred NeurosimVR but was equivalent for medical professionals. No statistics were provided for comparison of features between groups or within groups
Naddeo *et al.*, 2017 (46)	Case series	A pilot study to determine the applicability and use of VR-derived, patient specific implants for pedicle screw placement	Nine cases performed by single spine surgeon, with two elaborated on for results	Not completed	Time per screw insertion, total number of radiographs per screw	Reduced time to screw insertion (36.25 *vs.* 9.5 min with template) and reduced number of radiographs per screw (12.5 with no template *vs.* 2 with template). No statistical analysis was undertaken and no patient follow-up reported
Ma *et al.*, 2017 (37)	Case series	A pilot study to determine applicability of MR CT and US-guided pedicle cannulation	One surgeon performing eight K-wire pedicle insertions in phantom and then four K-wire pedicle insertions in a sheep cadaver	Not completed	Mean targeting error, mean angle error	MR US-guided K-wires had less targeting errors (2.41 *vs.* 5.18 mm) and angle errors (3.13 *vs.* 5.89 deg.). No comparative statistics were performed
Elmi-Terander *et al.*, 2019 (61)	Technical series	A pilot study to determine the applicability of a hybrid OR using AR navigation for pedicle cannulation	Two neurosurgeons performing 66 Jamshidi needle placements and 18 cannulated pedicle screw insertions in thoracolumbar spine of cadavers	Not completed	Navigation time per insertion, screw placement accuracy within pedicle, error between planned path and Jamshidi needle placement	Navigation time was 90±53 seconds. Two pedicle screws breached (89% accuracy). Error angle of Jamshidi needle was 0.9±0.8 deg. No radiation was used. Determined to be feasible to use in patients
Zhou *et al.*, 2019 (40)	Intervention-control group	A pilot study for feasibility of VR preoperative planning for lumbosacral TPED combined with isocentric navigation	Four surgeons performed TPED for L3/4, L4/5 and L5/S1 on cadavers without navigation and then with VR preoperative planning and isocentric navigation	Not completed	Puncture channel time, radiation exposure time	Results varied between levels. VR preoperative planning and isocentric navigation reduced puncture time at L4/5 and L5/S1, and reduced radiation exposure time at all levels
Keri *et al.*, 2015 (43)	Intervention-control group	Efficacy of MR platform in teaching lumbar puncture localization to novice trainees	Twenty-four residents randomly assigned to MR or control groups with three virtual patients	Not completed	Needle path, tissue damage, total time of procedure, needle insertion time, success rate	MR trained residents demonstrated statistically significant improvement in needle path, tissue damage, and needle insertional time compared to control
Abe *et al.*, 2013 (34)	Intervention-control group + cross-sectional	Efficacy of novel AR guidance system for percutaneous vertebroplasty	Two surgeons performed 40 AR assisted vertebroplasty trials, then performed in five patients with osteoporotic vertebral compression fractures	Not completed	Insertion angle error (EIA) in coronal and sagittal planes, technical outcomes [pedicle breach, PMMA leakage]	AR assisted vertebroplasty had significantly improved EIA in coronal and sagittal planes. In clinical use, there was no pedicle breach or PMMA leakage
Chitale *et al.*, 2013 (29)	Pre-test and post-test/intervention	Effectiveness of an MR training curriculum for neurosurgery residents in percutaneous pedicle screw placement	Eight residents completed a pre-intervention test, then a didactic learning session, then completed an MR pedicle insertion followed by a post-test	Not completed	Fluoroscopy score (starting point and trajectory, fluoroscopy time, number of fluoroscopic shots), computed tomography score (time and starting point and trajectory), test score	A non-significant improvement in all domains was seen using the training model
Färber *et al.*, 2009 (33)	Intervention-control group	Pilot study for effectiveness of a VR simulator for teaching LP	Forty-two medical students completed three virtual LP and questionnaire	Face and content validity—author developed	Computer generated score, number of successful virtual LPs	VR LP training improved performance on the utilized trainer (no significance testing performed)
Moult *et al.*, 2013 (44)	Intervention-control group	Evaluation of MR training platform compared to traditional training for percutaneous lumbar facet joint injections	Twenty-six pre-medical undergraduate students completed L3/4 and L4/5 MR facet injections	Not completed	Total time, total needle path, time inside phantom, needle path inside phantom, percent success rate	MR trained novices significantly completed more successful facet injections than control
Rambani *et al.*, 2014 (48)	Intervention-control group	To develop a training system for pedicle screw insertion and validate its effectiveness	Twelve junior orthopaedic trainees completed VR lumbar pedicle screw insertions	Construct validity	Scoring system based on total time, pedicle screw insertion accuracy, number of exposures required to complete insertion	Significant improvement in scores for VR trained group over control in using the simulator
Sutherland *et al.*, 2013 (49)	Comparison of three cohorts using single intervention	To demonstrate the creation of, and perceived benefit of use of AR system for teaching LP	Ten participants (four radiology residents, three medical students, three technicians) performed trials and a final LP followed by questionnaire	Face validity—author developed	None	Qualitative description of face validity based on questionnaire with no statistics performed between participant groups. Overall positive feedback
Weigl *et al.*, 2016 (30)	Comparison of two groups receiving different surgical disruptions	Determination of surgical disruption on surgeon performance and perceived workload	Nineteen junior surgeons were randomized to two disruption scenarios during vertebroplasty and following scenario, questionnaires were completed (SURG-TLX)	Not completed	SURG-TLX scores (mental workload), performance outcomes (trocar deviation, length of tooltip trajectory, fluoroscopy exposure time, overall duration, total number of fluoroscopic shots)	Surgical disruptions produced significantly higher SURG-TLX scores and were associated with poorer performance metrics, notably total number of fluoroscopic images
Wucherer *et al.*, 2014 (31)	Comparison of two groups receiving surgical disruption	Development of an MR training environment	Five surgeons performed an MR vertebroplasty with crisis scenarios/task disruptions	Face validity—author developed Construct validity—comparison to expert	None	Face validity of an MR training environment for surgical task disruptions in vertebroplasty was demonstrated and construct validity was attempted through single expert user
Wucherer *et al.*, 2015 (32)	Comparison of two groups performing MR vertebroplasty and receiving surgical disruption	Assessment of MR environment for usability in training vertebroplasty, and workload during crisis/surgical disruption	Nineteen junior surgeons performed MR vertebroplasty with crisis scenario/task disruptions	Face validity—author developed	SURG-TLX scores (mental workload), performance outcomes [root-mean-square deviation (RMSD), fluoroscopic exposure time, procedural time]	The users felt the task was realistic via face validity questionnaire. MR training enabled significantly faster procedure completion times
Koch *et al.*, 2019 (36)	Single group outcomes evaluation	Assessment of VR vertebroplasty simulator and development of a surgeon-reported experiential classification system	Thirteen orthopaedic trauma surgeons and neurosurgeons performed a percutaneous VR vertebroplasty	Face validity—author developed	Intraoperative performance metrics (procedure time, path length of tool-tip, motion smoothness, X-ray source length, number of X-ray shots, expert observations [Objective Structured Assessment of Technical Skills (OSATS) score, pass/fail recommendation]	Over half (53.8%) passed based on expert criteria. Face validity was demonstrated. Verbal feedback was collected and classified based on author defined task analysis. Simulation performance was not correlated with negative verbal feedback
Yeo *et al.*, 2011 (45)	Intervention-control group	Effectiveness of AR simulation in teaching percutaneous facet joint injections	Forty students (medical, engineering, first year residents) randomized to AR group (n=20) and control groups (n=20) and performed needle insertions	Not completed	Rate of success, overall procedural time, needle time inside phantom, needle path inside phantom, tissue damage, out of plane and in-plane deviation	AR group demonstrated significantly improved success rate and less potential for tissue damage than control

VR, virtual reality; AR, augmented reality; TVACF, transvertebral anterior cervical foraminotomy; MISS, minimally invasive spine surgery; HMD, head-mounted display; MR, mixed reality; PELD, percutaneous endoscopic lumbar discectomy; PTED, percutaneous transforaminal endoscopic discectomy; PMMA, polymethylmethacrylate; VAS, Visual Analog Scale; US, ultrasound; OR, operating room; TPED, transpedicular percutaneous endoscopic discectomy; LP, lumbar puncture.

References

57. Gasco J, Patel A, Ortega-Barnett J, et al. Virtual reality spine surgery simulation: an empirical study of its usefulness. Neurol Res 2014;36:968-73.
58. Choque-Velasquez J, Colasanti R, Collan J, et al. Virtual Reality Glasses and "Eye-Hands Blind Technique" for Microsurgical Training in Neurosurgery. World Neurosurg 2018;112:126-30.
59. Xin B, Chen G, Wang Y, et al. The Efficacy of Immersive Virtual Reality Surgical Simulator Training for Pedicle Screw Placement: A Randomized Double-Blind Controlled Trial. World Neurosurg 2018;124;e324-30.
60. Liu X, Bai H, Song G, et al. Augmented reality system training for minimally invasive spine surgery. 2017 IEEE International Conference on Robotics and Biomimetics (ROBIO); 5-8 Dec. 2017; Macau, China. IEEE, 2017.
61. Elmi-Terander A, Nachabe R, Skulason H, et al. Feasibility and Accuracy of Thoracolumbar Minimally Invasive Pedicle Screw Placement With Augmented Reality Navigation Technology. Spine (Phila Pa 1976) 2018;43:1018-23.