The Canadian Journal of Neurological Sciences (2024), **51**, 255–264
doi:10.1017/cjn.2023.46

Canadian Journal of Neurological Sciences
Journal Canadien des Sciences Neurologiques

Original Article

The Development And Applications Of Augmented And Virtual Reality Technology In Spine Surgery Training: A Systematic Review

Youngkyung Jung[1], Varun Muddaluru[2], Pranjan Gandhi[1], Markian Pahuta[3] and Daipayan Guha[4]

[1]Department of Medicine, Faculty of Health Sciences, McMaster University, Hamilton, Ontario, Canada, [2]Royal College of Surgeons in Ireland, Dublin, Ireland, [3]Division of Orthopedic Surgery, Hamilton General Hospital, McMaster University, Hamilton, ON, Canada and [4]Division of Neurosurgery, Hamilton General Hospital, McMaster University, Hamilton, ON, Canada

ABSTRACT: ***Background:*** The COVID-19 pandemic has accelerated the growing global interest in the role of augmented and virtual reality in surgical training. While this technology grows at a rapid rate, its efficacy remains unclear. To that end, we offer a systematic review of the literature summarizing the role of virtual and augmented reality on spine surgery training. ***Methods:*** A systematic review of the literature was conducted on May 13th, 2022. PubMed, Web of Science, Medline, and Embase were reviewed for relevant studies. Studies from both orthopedic and neurosurgical spine programs were considered. There were no restrictions placed on the type of study, virtual/augmented reality modality, nor type of procedure. Qualitative data analysis was performed, and all studies were assigned a Medical Education Research Study Quality Instrument (MERSQI) score. ***Results:*** The initial review identified 6752 studies, of which 16 were deemed relevant and included in the final review, examining a total of nine unique augmented/virtual reality systems. These studies had a moderate methodological quality with a MERSQI score of 12.1 + 1.8; most studies were conducted at single-center institutions, and unclear response rates. Statistical pooling of the data was limited by the heterogeneity of the study designs. ***Conclusion:*** This review examined the applications of augmented and virtual reality systems for training residents in various spine procedures. As this technology continues to advance, higher-quality, multi-center, and long-term studies are required to further the adaptation of VR/AR technologies in spine surgery training programs.

RÉSUMÉ : Le développement et les applications des technologies de la réalité augmentée et de la réalité virtuelle dans la formation destinée à la chirurgie de la colonne vertébrale : une revue systématique. ***Objectif :*** La pandémie de COVID-19 a accéléré l'intérêt mondial croissant pour le rôle des technologies de la réalité augmentée et de la réalité virtuelle dans la formation chirurgicale. Bien que ces technologies se développent rapidement, leur efficacité reste incertaine. À cet égard, nous voulons proposer ici une revue systématique de littérature résumant le rôle de ces technologies dans la formation dédiée à la chirurgie de la colonne vertébrale. ***Méthodes :*** Une analyse systématique de la littérature a été réalisée le 13 mai 2022. Les bases de données PubMed, Web of Science, Medline et Embase ont ainsi été examinées à la recherche d'études pertinentes. Seules des études provenant de programmes orthopédiques et neurochirurgicaux de la colonne vertébrale ont été prises en compte. Précisons qu'aucune restriction n'a été imposée quant aux types d'étude, aux modalités de réalité virtuelle ou augmentée ou aux types de procédure adoptée. Enfin, une analyse qualitative de données a été effectuée et toutes les études retenues ont reçu un score « MERSQI » (M*edical Education Research Study Quality Instrument*). ***Résultats :*** Un examen initial a permis d'identifier 6752 études, dont 16 ont été jugées pertinentes et incluses dans notre analyse finale. Au total, cela a représenté neuf systèmes uniques de réalité augmentée ou virtuelle. La qualité méthodologique de ces études s'est révélée moyenne, leur score MERSQI étant de 12,1 + 1,8. À noter que la plupart de ces études ont été menées dans des établissements monocentrique dont les taux de réponse ne sont pas clairs. Le regroupement statistique (*statistical pooling*) des données a été limité par l'hétérogénéité des modèles d'étude. ***Conclusion :*** Cette analyse s'est penchée sur les applications des technologies de la réalité augmentée ou virtuelle destinées à former des résidents quant aux diverses procédures chirurgicales de la colonne vertébrale. Comme ces technologies continuent de progresser, des études de meilleure qualité, multicentriques et à long terme sont nécessaires pour approfondir leur adaptation aux programmes de formation à la chirurgie de la colonne vertébrale.

Keywords: virtual reality; Augmented reality; Spine surgery; Surgical education; Systematic review

(Received 10 January 2023; final revisions submitted 18 March 2023; date of acceptance 12 April 2023; First Published online 28 April 2023)

Introduction

Augmented and virtual reality (AR and VR respectively) broadly refer to advanced visualization systems enabling users to immerse in and meaningfully interact with computer-generated environments. According to Milgram, AR and VR exist on a spectrum wherein reality and virtuality share an inverse relationship.[1] VR

Corresponding author: Youngkyung Jung, McMaster University, 1280 Main St West, Hamilton ON, Canada. Email: jessica.jung@medportal.ca

Cite this article: Jung Y, Muddaluru V, Gandhi P, Pahuta M, and Guha D. (2024) The Development And Applications Of Augmented And Virtual Reality Technology In Spine Surgery Training: A Systematic Review. *The Canadian Journal of Neurological Sciences* **51**: 255–264, https://doi.org/10.1017/cjn.2023.46

is a three-dimensional fully digitized, simulated, and immersive environment either emulating a real or imaginary world. Comparatively, AR overlays virtual elements onto the physical environment to enhance and supplement certain elements of the real world. Mixed reality is proposed to be an extension of AR, wherein users can interact with and extensively manipulate both the physical and virtual environments.

AR and VR technologies were developed in the 1960s with Sutherland creating the first head-mounted display using ultrasound tracking to create a 3D virtual world.[2] Subsequent significant developments in AR and VR technology include the development of electromagnetic trackers to manipulate the distance between objects in a virtual world, increasingly sophisticated graphics, and haptic feedback. There is a growing interest in using these technologies as a way to optimize surgical exposure without the typical cost and resource constraints of traditional cadaver labs, with early studies demonstrating general improvements in speed and accuracy with VR/AR-guided training compared to traditional didactic methods.[3,4]

Despite the growing momentum surrounding the potential for AR and VR to advance surgical training, there is limited evidence available regarding its effectiveness, validity, and outcome measurement. There is a significant advantage to adopting these technologies to spine procedures. Fundamental procedures such as pedicle screw placement and endoscopic spine approaches have a very steep learning curve and are often left to the purview of senior operators. Spine surgery training also consists of residency training in either a neurological or orthopedic surgery program, often followed by a spine fellowship. These two separate training models with a heterogeneity of case breadth and variety pose a unique challenge in spine surgery training[5]. Neurosurgical and spine procedures are optimal for incorporating AR and VR given the limited inherent mobility of the brain and spine. This inertia facilitates tracking and displaying of images, giving rise to the ubiquity of neuronavigation for multiple purposes; in the spine, used commonly to guide instrumentation placement, and more recently decompression, tumor resection, deformity correction, and interbody grafting.[6–8]

This must be weighed against the inherent limitations to AR and VR, including cost, physical discomfort (i.e., cybersickness including eye strain, vertigo, and nausea), latency in visual displays or tactile feedback, and limited realism.[9–11] Moreover, AR/VR is not a homogenous entity and needs to be tailored for a specific purpose. For example, laparoscopic simulators are fairly well-established, while similar training models in orthopedic or ophthalmology remain poorly developed.[12]

The aim of this systematic review is to evaluate the current landscape of advanced visualization techniques in the context of spine surgical education and identify gaps that may be addressed with existing and emerging technologies.

Methods

Search Strategy

A systematic search was conducted on May 13th 2022 of published articles available across four electronic databases: PubMed, Web of Science, Medline, and Embase. Search strategies were refined in collaboration with an institutional librarian, with the following terms for OVID Medline: ("augmented reality" OR "virtual reality OR haptic technology" OR "computer-assisted instruction" OR "stimulation training") AND ("spine surgery" OR "spinal diseases") AND ("surgery training" OR "surgical education" OR "internship and residency"); these terms were translated by our librarian to meet the requirements of the other databases. Search included inception through May 13, 2022. To optimize the yield of our search, no restrictions for publication status, year, nor language were placed.

Eligibility Criteria

We included all primary studies evaluating the role of virtual, augmented, or mixed reality in training surgical residents across spine procedures. Studies that did not involve surgical trainees or junior surgeons were excluded. There were no exclusion criteria based on the type of study, training modality, surgical procedure, nor outcomes collected. Our primary objective was to assess the current landscape of AR/VR in spine surgical education and its feasibility as a training tool for the range of spinal procedures.

All eligible articles were screened in two stages. Initial screening of abstracts was performed by VM and YJ to exclude overtly irrelevant articles. Discrepancies at this phase were automatically forwarded to the next stage to ensure all relevant articles were included. This was followed by full-text screening of all articles by PG and VM; discrepancies at this stage were solved with consensus and oversight by a third reviewer (YJ) as required.

Data Extraction

Data extraction was performed by two independent reviewers. Given the aim of this paper was to assess the role of AR/VR technology in surgical education, the Medical Education Research Study Quality Instrument (MERSQI) score was calculated for each study. The MERSQI score is a 10-item scale designed in 2007 to assess the quality of reporting of literature in medical education.[13] Scores for each item are totaled into a final MERSQI score, ranging from 3 to 18. This instrument was specifically designed to provide an objective means of evaluating the quality of study design, data analysis, and outcome measurements.

Within each paper, the following data were also extracted for analysis; type of study, number of participants, type of AR/VR device used, surgical procedure analyzed, outcomes, available feedback. For the purposes of bibliometric evaluation, the number of citations, citations per year, and funding sources were also recorded. Where applicable, a meta-analysis of standardized mean differences was conducted to evaluate the impact of AR/VR-guided training on operative performance. To this end, a random-effects model was used. Calculations and forest plots were generated using MedCalc statistical software (MedCalc Software Ltd).

Results

Overview of Results

The literature search initially yielded 7259, of which 510 were duplicates. A total of 6752 publications were screened, and 12 were selected and four additional studies were manually added from a search of references (Figure 1). The studies collectively included 348 participants and 3 study designs (single group cross-sectional, n = 6; randomized control trial, n = 6; validation study, n = 4). Virtual reality platforms were most common (n = 11),[14–20] followed by augmented reality (n = 4),[21–24] with only one study focusing on mixed reality (n = 1).[25] The most common procedure was pedicle screw insertion (n = 7),[14–17,21–23] followed by discectomy (n = 4),[25–28] laminectomy (n = 3),[19,20,29] percutaneous spinal needle placement (n = 1),[24] and lateral lumbar access (n = 1).[18] A total

https://doi.org/10.1017/cjn.2023.46 Published online by Cambridge University Press

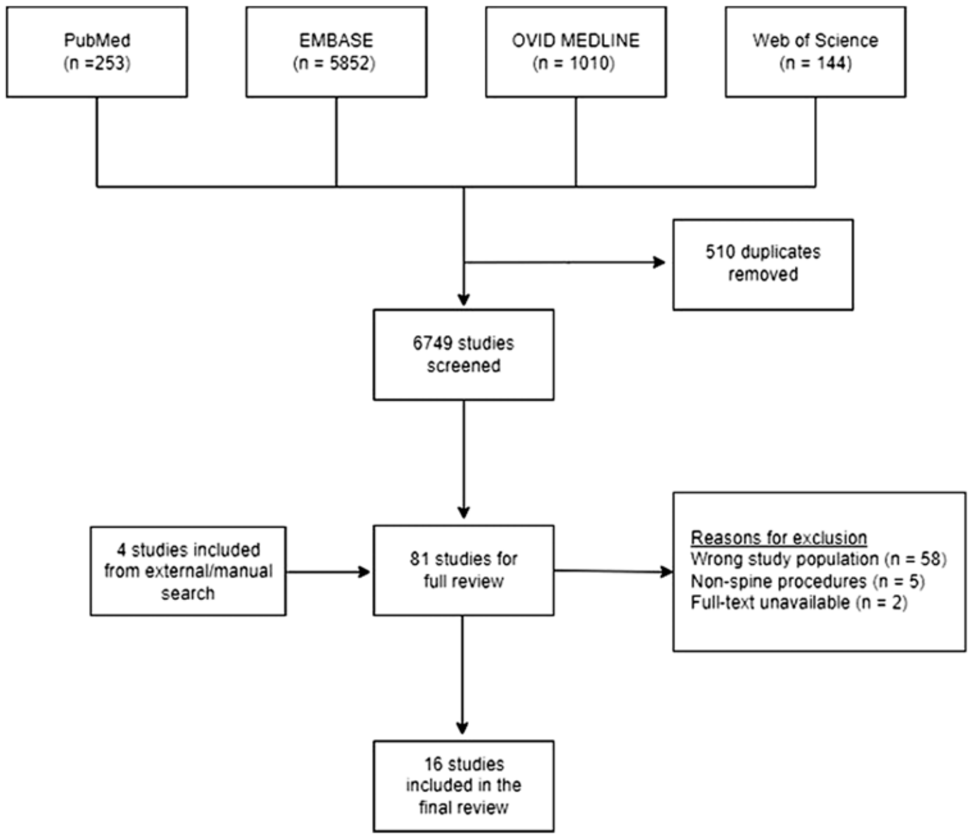


Figure 1: PRISMA flow diagram.

of 11 unique AR/VR systems were evaluated, and individual study results are outlined below and highlighted in Table 1.

MERSQI Scores

The methodological rigor of the studies varied, with MERSQI scores ranging from 7.5 to 14.5, with a mean score of 12.1 + 1.8 (Table 2). Notably, most studies scored poorly within the sampling section, as most studies were single-center and did not specify a response rate. All selected publications scored a 1.5 on the outcome category, as they all focused on operative skills without direct patient involvement. Most studies (n = 14) had objective outcomes, with only two studies focusing on subjective self-assessments.[19,20]

Bibliometric Data

The bibliometric data collected are available in Table 3. Across these 16 studies, the most common journal of publication was *Neurosurgery* (n = 3) and *World Neurosurgery* (n = 3), followed by *Operative Neurosurgery* (n = 2), *Neurosurgical Focus* (n = 2), and *Archives of Orthopedic and Trauma Surgery* (n = 1), *Computers in Biology and Medicine* (n = 1), *Journal of Orthopedic Surgery and Medicine* (n = 1), *North American Spine Society Journal* (n = 1), and *The Journal of Bone and Joint Surgery* (n = 1). The number of citations (at time of writing) ranged from 2 to 68, with an average of 39.5 + 40.3. The average yearly citation rate was 4.0 + 1.8. Overall, the total number of citations appears to loosely correlate with the age of the paper.

Training Modalities

The virtual surgical training system (VSTS) is a custom VR platform created by Hou et al. (2018),[14,15] to optimize training of thoracic pedicle screw placement and fixation in junior trainees. The software reconstructs patient CT scans into manipulable 3D models of the spine displayed on a screen without the use of a headset. This system provides haptic feedback to the user, through a connected external handle. This technology was also used by Hou et al. (2018)[14] in a similar study of the cervical spine and Shi et al. (2018)[16] for lumbar pedicle screw placements. All three studies demonstrated improved accuracy of pedicle screw placements on cadaver models for trainees who underwent a practice trial with the VSTS compared to a standard teaching session. Across all three studies, pedicle screw placement was evaluated based on CTs by three independent observers. Hou et al. (2018) and Shi et al. (2018) used a four-point grading scale to quantify the extent of pedicle screw breach (grade I: no breach, grade II: 0–2 mm breach, grade III: 2–4 mm breach, grade IV: > 4 mm breach) for thoracic and lumbar screw placement. For cervical screw placement, Hou et al. (2018) used three-point grading scale (grade I: no breach, grade II: < 50% screw diameter breach, grade III: > 50% screw diameter breach). At the same institution as these three studies, Xin et al. (2018)[17] developed a novel Immersive Virtual Reality Surgical Simulator (IVRSS). Unlike the VSTS, the IVRSS uses VR glasses to recreate a practice model, instead of a display screen. This study also measured screw placement on CTs by three independent observers, based on a similar four-point grading scale (grade I: no breach, grade II: < 25% screw diameter breach without violation of the anterior vertebra, grade III: 25%–50% screw diameter breach without violation of the anterior vertebra, grade IV: > 50% screw diameter breach and/or violation of the anterior vertebra). Compared to the control group, surgical trainees randomized to practice sessions with the IVRSS had significantly improved pedicle screw accuracy and lower failure rates on cadaver specimens.

https://doi.org/10.1017/cjn.2023.46 Published online by Cambridge University Press

Table 1: Study design and outcomes

Study	Discipline	Study design	N	Training level	Platform	AR/VR	Procedure	Data processing	Results
Luciano et al. 2011[23] USA	Neurosurgery	Single group cross-sectional	51 (no control)	Fellows and residents	ImmersiveTouch	Augmented reality	T9-11 pedicle screw placement	Final placements based on parameters set by attending neurosurgeons who underwent trial sessions on the simulator. Results (i.e., distance from target and failure to place pedicle screw altogether) calculated by program	After one practice session, failure rate of 16.9% on a virtual model (average of 6 attempts)
Luciano et al., 2013[24] USA	Neurosurgery	Single group cross-sectional	63 (no control)	Fellows and residents	ImmersiveTouch	Augmented reality	Percutaneous spinal needle placement	Final placements based on pre-defined targets. Results (i.e., distance from target and total fluoroscopic time) calculated automatically by the program	After one practice session, failure rate of 7.93%, and improvement in both trajectory of needle placement and time taken after 2 attempts ($p = .02$)
Hou et al., 2018[14] China	Orthopedic surgery	Randomized control trial	10 residents	Junior residents	VSTS	Virtual Reality	Cervical screw placement (bilateral C3-C6 pedicle screw instrumentation)	Post-training cadaveric cervical screw placements rated by 3 independent staff	Reduced screw penetration rates on a cadaver model after VR training compared to controls (62.5 and 10%, $p < 0.05$)
Hou et al., 2018b[15] China	Orthopedic surgery	Randomized control trial	10 (simulation training = 5, control = 5)	PGY1-PGY2	VSTS	Virtual reality	Thoracic pedicle screw T6-12 on cadavers after virtual simulation training with haptic feedback	Post-training cadaveric cervical screw placements rated by 3 independent staff	Greater rate of grade I pedicle screw placement on cadavers with stimulation training (92.86 vs 70%, $p < .05$)
Shi et al., 2018[16] China	Orthopedic surgery	Randomized control trial	10 (n = 5 stimulation, n = 5 control)	Residents	VSTS	Virtual reality	Pedicle screw implantation	Post-training cadaveric cervical screw placements rated by 3 independent staff	Stimulation group had lower rate of pedicle breach (12.5% and 37.5%, $p < 0.05$) and greater rate of grade I or II screw placement (100 and 85%, $p < 0.05$)
Xin et al., 2018[17] China	Orthopedic surgery	Randomized control trial	24 (n = 12 VR, n = 12 control)	Residents	IVRSS-PSP	Virtual reality	Pedicle screw placement (T11-L4)	Post-training cadaveric cervical screw placements rated by 3 independent staff	Greater rate of pedicle screw placement without breaching anterior wall of vertebrae (89.6 vs 60.4%) on cadavers after virtual reality training compared to a traditional teaching session
Bissonnette et al., 2019[29] Canada	Orthopedic surgery and neurosurgery	Validation study	41 (no control)	Senior and junior residents	NeuroVR	Virtual reality	L3 hemilaminectomy	Position, angle, and force applied with surgical instruments, and volume of removed tissue automatically captured by the platform	An artificial neural network was able to distinguish junior and senior trainees with 97.6% accuracy
Yu et al., 2019[25] China	Orthopedic surgery	Randomized control trial	60 (n = 30 stimulation, n = 30 control)	Residents	3D Slicer platform	Mixed reality	PTED	Not specified	Reduced fluoroscopy time and accuracy of PTED on a 3D-printed model with mixed reality training compared to standard 2D learning
Luca et al., 2020[18] Italy	Orthopedic surgery	Single group cross-sectional	10 (no control)	Senior surgeons and residents	Oculus Rift S and 3D Systems Touch Haptic Device	Virtual reality	Lateral lumbar access	Individualized performance report generated by the software program outlining error –authors/programmers pre-define correct answers. Wrong decisions during the simulation can also result in intraoperative complications	Across 2 attempts, significant reduction in the number of errors (5.2 to 1.8). No correlation to experience

Dennler et al., 2020[22] Switzerland	Orthopedic surgery and neurosurgery	Randomized control trial	4 (n = 2 junior attending, n = 2 senior attending)	Junior and senior attendings	Microsoft Hololens	Augmented reality	Pilot holes for pedicle screws in lumbar vertebra (comparing free-hand technique and augmented reality technique)	Pedicle screw placements manually assessed based on CT images	No difference in AR training for experienced surgeons, but improvement (i.e., more central drill placement) for pedicle screw placement with AR assistance compared to free-hand screw placement for novice surgeons
Ledwos et al., 2020[26] Canada	Orthopedic surgery and neurosurgery	Validation study	21 residents (no control)	Senior and junior residents	Sim-Ortho	Virtual reality	C4-5 anterior cervical discectomy and fusion	Automated feedback based on specific objectives set at each step of the procedure	More senior trainees interacted with the disc nucleus, removed more tissue, and greater force on the annulus compared to junior trainees
Mirchi et al. 2020[27] Canada	Orthopedic surgery and neurosurgery	Validation study	21 residents (no control)	Senior and junior residents	Sim-Ortho	Virtual reality	Anterior cervical discectomy and fusion	Automated feedback based on specific objectives set at each step of the procedure	An artificial neural network was used to analyze 15 junior and senior residents on their surgical performance on a simulator. The training module identified the seniority level of a trainee with 83.3% accuracy
Alkadri et al., 2021[28] Canada	Orthopedic surgery and neurosurgery	Validation study	23 residents (no control)	Senior and junior residents	Sim-Ortho	Virtual reality	C4-5 anterior cervical discectomy and fusion	Automated feedback based on specific objectives set at each step of the procedure	An artificial neural network was used to analyze 15 junior and senior residents on their surgical performance on a simulator. The training module identified the seniority level of a trainee with 80% accuracy
Chen et al., 2021[19] Canada/ USA	Orthopedic surgery and neurosurgery	Single pre-test and post-test	28 (no control)	Senior and junior residents	Samsung Odyssey + OpenVR and SlicerVirtualReality plugin	Virtual reality	Single-level spinal decompression	Not specified	Post-test scores on an anatomy exam improved significantly for junior trainees (PGY 1-2), with no difference for senior learners (PGY3-fellows). Most (89%) residents agreed this model was useful for learning anatomy, whilst fewer believed it was a useful tool for technical training (33%)
Knafo et al., 2021[20] France	Neurosurgery	Single group cross-sectional	12 (no control)	Senior and junior residents	NeuroTouch	Virtual reality	Lumbar hemilaminectomy	Trainees evaluated based on built-in assessments. Performance metrics include volume of L3 vertebra removed. Marks deducted for spinal cord injury, removal of other tissue (e.g., flavium ligament, other vertebral bodies), excessive force, excessive duration, blood loss	No correlation between self-assessment and knowledge with technical skills as measured by NeuroTouch simulator
Yanni et al., 2021[21] USA	Neurosurgery and orthopedic surgery	Single group cross-sectional	11 (no control)	Staff surgeons and residents	SpineAR	Augmented reality	L2-L5 pedicle screw placement	Instrumentation manually assessed by an independent staff neuroradiologist	98.4% accuracy of free-hand pedicle screw placement with AR technology on 3D printed phantom models after a training session. No correlation to experience

IVRSS = immersive virtual reality surgical simulator; PGY = post-graduate year; PTED = percutaneous transforaminal endoscopic discectomy; VSTS = virtual surgical training system.

NeuroTouch/NeuroVR (CAE Inc., Montreal, QB, Canada) is a commercially available neurosurgical VR simulator. Bissonnette et al. (2019)[29] performed a validation study of the NeuroVR system by creating an artificial intelligence algorithm that was able to distinguish senior vs junior trainee performance across several metrics (i.e., procedure duration, force applied on dura, control over the drill/burr). The algorithm was able to correctly identify trainees with a 97.6% accuracy. Knafo et al. (2021)[20] demonstrated that performance across several procedures (lumbar hemilaminectomy, meningioma resection, and ETV) was unrelated to a trainee's self-assessment of their surgical ability and knowledge. Performance metrics included volume of L3 vertebra removed, and points were deducted for removal of vertebrae other than L3, injury/removal of neighboring structures (spinal cord, dura, other tissues), and excessive blood loss. These outcomes were pre-programmed within the NeuroTouch software. Performance scores were unrelated to seniority and a trainee's self-evaluation of their abilities. A major limitation of this study was the lack of a control group.

Several studies used commercially available VR or AR headsets. Dennler et al. (2020)[22] combined the Microsoft Hololens (Microsoft, Redmond, WA, USA), an AR headset, with Unity (Unity Technologies, San Francisco, CA, USA), a commercially available software package, enabling participants to rotate and translate a 3D sawbone model of the lumbar vertebrae to facilitate pedicle screw placements. The results were graded based on CT scans obtained of the models after pedicle screw insertion; the authors used Phönix-PACS software (GmbH, Freiburg, Germany) to calculate the accuracy of screw placements. For the junior surgeons, there was a lower rate of pedicle screw perforations with AR-guided placement compared to free-hand attempts. Luca et al. (2020)[18] used the Oculus Rift VR headset (Oculus VR, Irvine, CA, USA) and its built-in software development kit system, combined with a robotic arm with haptic feedback to mimic lateral lumbar access. Trainees were evaluated on competency in setting up the OR, bleeding control, manipulation of soft tissues, and target accuracy through this VR system. Using the Oculus Rift's software development kit, three senior spine surgeons created a pre-defined pathway of correct options, including appropriate OR set-up, identifying the correct anatomic level with fluoroscopy, and identification of the tissue layers. The system automatically created a score sheet based on a trainee's performance and choices compared to those of their senior colleagues. There was no difference between residents and senior surgeons in respect to the mean number of errors. Chen et al. (2021)[19] opted to use Samsung Odyssey (Samsung, Seoul, South Korea) over the Oculus headsets as it relies on inside-out tracking instead of conventional position tracking. This facilitates faster set-up as there are no external cameras and the device is essentially a self-contained unit. This was combined with 3D Slicer (open-source), an image analysis software capable of developing 3D models based on CT and MRIs. Junior residents had significant improvements in anatomical test scores after performing a virtual laminectomy and lateral recess decompression, whereas there were no changes for senior residents, suggesting VR enhances understanding of surgical anatomy in the early stages of training. Yu et al. (2019)[25] also used the 3D Slicer platform to create 3D representations of the lumbar spine based on CT scans. Compared to trainees who were taught with plain films, residents who learnt percutaneous transforaminal endoscopic discectomy had shorter procedure times.

A team at the University of Chicago used an in-house AR system, ImmersiveTouch (Industrial Virtual Reality Institute, Chicago, IL, USA) to evaluate pedicle screw placement[23] and percutaneous spinal fixation.[24] ImmersiveTouch is a projection-based AR system with hardware capable of providing haptic feedback. Accuracy (i.e., distance to target or correct level) improved with repeated attempts for both procedures.

Yanni et al. (2021)[21] adapted SpineAR (Surgical Theater, Inc., Beachwood, OH, USA), an intraoperative navigation tool combined with a headset to allow for overlay of AR-guided pedicle screw trajectory with the operative field, optimizing surgical efficiency by eliminating the need for the operator to continuously shift their gaze from an external monitor to the operative site. Both junior learners and attendings had higher rates of Gertzbein-Robbins grade 0 (i.e., no cortical breach) or 1 (i.e., 0–2 mm cortical breach) screw placements compared to literature rates of free-hand placements.

A series of validation studies for a novel VR tool, the Sim-Ortho (OSSimTech, Montreal, Canada, and AO Foundation, Davos Switzerland) from McGill. Ledwos et al. (2020) demonstrated reasonable validity of the device for a one-level anterior cervical discectomy and fusion (ACDF). Adequate face and content validity were demonstrated across resident ratings of simulator realism. Construct validity was assessed by analyzing the performance differences between junior and senior trainees; the latter interacted with the disc for a greater proportion of time throughout the operation and removed more target tissue, demonstrating Sim-Ortho's potential as a surgical education tool. Some noted limitations of the Sim-Ortho were its omission of certain parts of a standard ACDF (e.g., soft tissue exposure) and lack of realism across certain structures (e.g., the posterior longitudinal ligament). Subsequent studies from the same institution aimed to use the Sim-Ortho for specific portions of an ACDF, such as the discectomy[27] and annulus incision.[28] Both studies demonstrated an artificial intelligence algorithm could be trained to identify senior and junior trainees with 80 and 83.3% accuracy using pre-defined metrics such as safety, efficiency, and appropriate usage of surgical instruments. This demonstrates a role for surgical simulation to not only train residents, but also provide objective feedback as to their performance against a standardized cohort.

Simulator Feedback

The 16 studies varied significantly in the type and extent of feedback provided to the trainee, as outlined in Table 1. Across the 11 different simulators used across the studies, five provided feedback on pedicle screw placement (i.e., ImmersiveTouch, VSTS, IVRSS-PSP, Microsoft HoloLens, and SpineAR). For several, The ImmersiveTouch program provided automatic calculations for pedicle screw placements, whereas for the other modalities, the results were manually obtained based on post-instrumentation CTs. The validation studies for Sim-Ortho focused on more nuanced details such as volume of tissue removed, contact with structures, instrument tip path length (as a proxy for efficient intraoperative movements), force used with drilling – all automatically provided by the built-in software. The 3D Slicer primarily compared trainees based on overall fluoro exposure time. In Chen's study, a plugin software was used to mimic posterior lumbar decompression, and an emphasis was placed on the exposure and handling of soft tissue. All simulators provided haptic and

Table 2: MERSQI scores

Study	Study Design	Sampling		Type of Data	Validity of evaluation instrument			Data analysis		Outcomes	Total score
		Institution	Response rate		Internal structure	Content	Relationships to other variables	Appropriateness	Complexity		
Luciano et al., 2011[23] USA	1	0.5	0.5	3	1	1	0	1	2	1.5	11.5
Luciano et al., 2013[24] USA	1	0.5	0.5	3	1	1	0	1	2	1.5	11.5
Hou et al., 2018[14] China	3	0.5	0.5	3	1	1	0	1	2	1.5	13.5
Hou et al., 2018b[15] China	3	0.5	0.5	3	1	1	0	1	2	1.5	13.5
Shi et al., 2018[16] China	3	0.5	0.5	3	1	1	1	1	2	1.5	14.5
Xin et al., 2018[17] China	3	0.5	0.5	3	1	1	0	1	2	1.5	13.5
Bissonette et al., 2019[29] Canada	1	1.5	0.5	3	1	1	0	1	2	1.5	12.5
Yu et al., 2019[25] China	3	0.5	0.5	3	1	1	1	1	2	1.5	14.5
Dennler et al., 2020[22] Switzerland	1	0.5	0.5	3	1	1	1	1	2	1.5	12.5
Ledwos et al., 2020[26] Canada	1	1.5	0.5	3	1	1	0	1	2	1.5	12.5
Luca et al., 2020[18] Italy	1	0.5	0.5	3	1	0	1	1	0	1.5	9.5
Mirchi et al., 2020[27] Canada	1	1.5	0.5	3	1	1	0	1	2	1.5	12.5
Alkadri et al., 2021[28] Canada	1	1.5	0.5	3	1	1	0	1	2	1.5	12.5
Chen et al., 2021[19] USA	1	0.5	0.5	1	1	0	1	1	0	1.5	7.5
Knafo et al., 2021[20] France	1	0.5	0.5	1	1	0	1	1	2	1.5	9.5
Yanni et al,. 2021[21] USA	1	0.5	0.5	3	1	1	1	1	2	1.5	12.5

visual feedback, with the Sim-Ortho uniquely incorporating auditory feedback to further optimize the realism of the scenarios.

The five studies analyzing pedicle screw placement offered enough data to allow for pooled analysis; all of these studies were non-randomized comparisons between AR/VR-trained groups and controls. All studies demonstrated increased accuracy of pedicle screw placements (i.e., reduced frequency of pedicle screw breach/perforations), with AR/VR-trained residents (OR 5.05, 95% CI 2.93–8.68). (Figure 2). Due to the disparity across studies in the grading scales used, the data are represented as odds of no breach vs any breach. For heterogeneity, $I^2 = 0\%$.

Discussion

This review identified 12 studies evaluating the role of AR/VR technologies in spine education. While a collective analysis of the data was limited by the heterogeneity of outcome measurements, procedure types, and AR/VR platforms across the study, there are several findings that suggest AR/VR may serve as a useful tool in surgical education.

Of note, AR/VR technologies are a heterogeneous entity, with no robust comparisons between the different available modalities in respect to surgical training. Four of sixteen studies evaluated AR systems, six focused on VR, with 1 study using a mixed reality system. Generally within neurosurgery, VR technologies have been used more frequently for training, and AR has traditionally used for intraoperative image enhancement, due to the ability to overlay images onto the operative field.[30] However, as demonstrated by several studies in this review, there is a role for AR-guided training in spine surgery. This becomes particularly relevant, as immersive VR systems are often time-intensive to set up and suffer from a lack of realism; this is countered in an AR environment, which allows for the overlay of visual data onto existing surroundings.

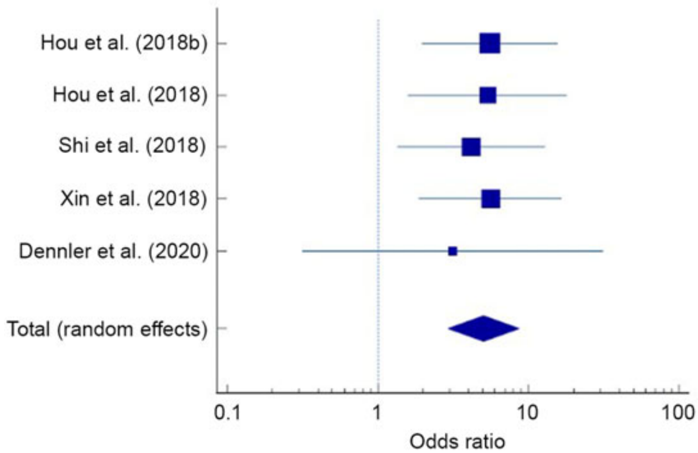
A general advantage of AR/VR is the ability to refine surgical skills in a low-stakes environment while generating customizable surgical scenarios and automated objective feedback. This provides

https://doi.org/10.1017/cjn.2023.46 Published online by Cambridge University Press

Table 3: Bibliometric data

Study	Location	Journal	Citations (total)	Citations/ year	Funding source
Luciano et al., 2011[23]	University of Illinois at Chicago, USA	Neurosurgery	68	5.23	AANS Young Neurosurgeons Committee, NIH, NIBIB
Luciano et al., 2013[24]	University of Illinois at Chicago, USA	Neurosurgery	41	3.73	AANS Young Neurosurgeons Committee, NIH, NIBIB
Bissonnette et al., 2019[29]	Neurosurgical Simulation & Artificial Intelligence Learning Centre, Department of Neurosurgery, Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada	The Journal of Bone and Joint Surgery	38	7.6	Di Giovanni Foundation, The Montreal Neurological Institute and Hospital and the McGill Department of Orthopaedics
Mirchi et al., 2020[27]	Neurosurgical Simulation and Artificial Intelligence Learning Centre, Department of Neurology and Neurosurgery, Montreal Neurological Institute and Hospital, McGill University. Montreal, Canada	Operative Neurosurgery	26	6.5	AO Foundation, Di Giovanni Foundation, Montreal Neurological Institute and Hospital, Department of Orthopaedic Surgery at McGill University
Xin et al., 2018[17]	Orthopaedic Oncology Center, Department of Orthopaedics, Changzheng Hospital, Shanghai, China	World Neurosurgery	25	5	None disclosed
Dennler et al., 2020[22]	Laboratory for Biomechanics, University Hospital Balgrist, University of Zürich, Zurich, Switzerland.	Journal of Orthopedic Surgery and Research	22	5.5	Balgrist University Hospital
Ledwos et al., 2020[26]	Neurosurgical Simulation and Artificial Intelligence Learning Centre, Department of Neurology & Neurosurgery, Montreal Neurological Institute and Hospital, McGill University	Operative Neurosurgery	15	5	AO Foundation, Di Giovanni Foundation, Montreal Neurological Institute and Hospital, Department of Orthopaedic Surgery at McGill University
Hou et al., 2018[14]	Changzheng Hospital, Second Military Medical University. Shanghai, China	Archives of Orthopaedic and Trauma Surgery	16	2.67	National Natural Science Foundation of China, Shanghai Natural Science Fund, Shanghai Education Science Research Project
Shi et al., 2018[16]	Changzheng Hospital, Second Military Medical University. Shanghai, China	World Neurosurgery	16	2.67	National Natural Science Foundation of China, Shanghai Natural Science Fund, Shanghai Education Science Research Project
Hou et al., 2018b[15]	Changzheng Hospital, Second Military Medical University. Shanghai, China	Neurosurgery	12	2	National Natural Science Foundation of China, Shanghai Natural Science Fund, Shanghai Education Science Research Project
Luca et al., 2020[18]	Spine Unit III, IRCCS Istituto Ortopedico Galeazzi, Milan, Italy	World Neurosurgery	11	2.75	None disclosed
Yu et al., 2019[25]	Orthopedic Department, Shanghai Tenth People's Hospital. Shanghai, China	World Neurosurgery	10	2	None disclosed
Yanni et al., 2021[21]	Not specified	Neurosurgical Focus	8	2.67	Surgical Theater Inc.
Alkadri et al., 2021[28]	Musculoskeletal Biomechanics Research Lab; Neurosurgical Simulation and Artificial Intelligence Learning Center, McGill University, Montreal, Canada	Computers in Biology and Medicine	7	2.33	Natural Sciences and Engineering Research Council of Canada Collaborative Research Development Grant, AO Foundation, Di Giovanni Foundation, Montreal Neurological Institute and Hospital, Department of Orthopaedic Surgery at McGill University
Chen et al., 2021[19]	Orthopedic Biomechanics Laboratory, Sunnybrook Health Sciences Centre. Toronto, Canada	North American Spine Society Journal	6	3	Feldberg Chair in Spinal Surgery at Sunnybrook Health Sciences
Knafo et al., 2021[20]	Assistance Publique – Hôpitaux de Paris, France	Neurosurgical Focus	2	0.67	None disclosed

AANS = American Association of Neurological Surgeons; NIH = National Institute of Health; NIBIB = National Institute of Biomedical Imaging & Bioengineering



OR	95% CI	Weight (%)
5.57	[1.96; 15.82]	27.03
5.4	[1.60; 18.21]	19.92
4.2	[1.35; 13.10]	22.85
5.63	[1.89; 16.79]	24.68
3.16	[0.32; 31.78]	5.53
5.05	[2.93; 8.68]	100

Figure 2: Random effect model forest plot for AR/VR-guided training versus control for pedicle screw placement accuracy.

trainees with concrete techniques to improve upon and allows for comparisons over time. In addition, AR/VR technologies allow trainees to focus on specific skills; for example, Luca and Knafo's models specifically incorporated bleeding control and handling of soft tissue into their metrics, while several other studies simply looked at final accuracy (e.g., pedicle screw placement). Other companies such as the OculusRift have created software development kits that enable users to create their own scenarios, theoretically allowing incorporation of real-world complications such as CSF leak, dural tears, patient wakening, etc. Furthermore, based on the validation studies performed for the Sim-Ortho model, there may also be a role for AR/VR technologies in providing objective feedback for residents, to help them identify if they are performing at a standard appropriate for their training year.[26–29] In another study by Winkler-Schwartz et al. (2019),[31] machine learning algorithms were used to identify junior trainees, senior trainees, and staff surgeons with remarkable accuracy, even across complex surgical tasks.[31]

An interesting concept raised by these studies is the growing role of artificial intelligence in surgical education. Winkler-Schwartz et al. (2019b) outline a 20-point scale to help determine quality of literature in this area – they specify in their review, many studies do not explain the educational relevance of a chosen metric.[32] While AR/VR technologies offer the ability to objectively quantify technical skills that were previously difficult to manually assess, it is important relevant measures are chosen – for example, eye movements, a common outcome used in AR/VR studies may not be particularly relevant, as it is not a teachable or trainable facet of surgical performance. Rather, metrics directly impacting surgical outcomes, such as duration, bimanual dexterity, volume of target tissue removed, etc., may be more useful.

Across the 16 studies, the commonly analyzed procedure was pedicle screw insertion, likely reflective of the significant morbidity associated with pedicle screw breach (i.e., fracture, nerve injury). While the studies used different grading systems to evaluate outcomes, when dichotomized into no breach vs breach, the pooled results of the 5 studies offering controls demonstrate a benefit of AR/VR-guided training. While this suggests AR/VR technologies can help trainees achieve minimal competency, these studies varied significantly in respect to case complexity and realism (e.g., soft tissue handling, hemostasis), making it difficult to determine if these findings will translate to enhanced operative performance.

Three studies demonstrated a decreased number of errors after a few practice runs in the same sitting, for lateral lumbar access[18], thoracic pedicle screw placement,[23] and percutaneous spinal fixation,[24] indicating that lack of experience with AR/VR is not a significant barrier to its use. Across the 5 studies that randomized participants into two groups (i.e., AR/VR teaching vs standard didactic lecture) prior to a surgical task, all demonstrated an improvement in respect to accuracy[14–17,22] or time[25] with VR/AR. This improvement may be facilitated by an enhanced understanding of surgical anatomy with AR/VR technology.

Conclusion

In this systematic review, limited evidence suggests AR/VR platforms are a useful tool for enhancing proficiency in various spine procedures, particularly pedicle screw insertion. However, the cost-benefit, translation to clinical practice remains unclear. As AR and VR technologies rapidly advance, further research will be necessary to reassess their role in surgical education.

Specific areas of future interest include the role of increasingly sophisticated automated feedback, the role of artificial intelligence, and algorithms to help analyze performance and determine the procedures and specific steps most amenable to being taught via AR/VR tools.

Conflicts of Interest. The authors of this manuscript have no conflicts of interest to disclose.

Author Contributions. YJ: conceptualization, methodology, writing-original draft, writing-reviewing & editing.

VM: conceptualization, methodology, writing-reviewing & editing.

PG: conceptualization, methodology, writing-reviewing & editing.

DG: conceptualization, methodology, writing-original draft, writing-reviewing & editing, visualization, supervision.

PG: conceptualization, methodology, writing-reviewing & editing.

DG: conceptualization, methodology, writing-original draft, writing-reviewing & editing, visualization, supervision.

References

1. Milgram P, Takemura H, Utsumi A, Kishino F. Augmented reality: A class of displays on the reality-virtuality continuum. Telemanipulator and Telepresence Technologies; 1994. pp. 2351.

https://doi.org/10.1017/cjn.2023.46 Published online by Cambridge University Press

2. Sutherland I. The ultimate display. In: Proceedings of the IFIPS Congress 65. *IFIP. CUMINCAD*, New York, 1965, pp. 506–508. http://papers.cumincad.org/cgi-bin/works/Show, accessed December 1, 2022.
3. Mao RQ, Lan L, Kay J, et al. Immersive virtual reality for surgical training: A systematic review. J Surg Res. 2021;268:40–58. DOI: 10.1016/j.jss.2021.06.045.
4. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance: Results of a randomized, double-blinded study. Ann Surg. 2002;236:458–463. DOI: 10.1097/00000658-200210000-00008 discussion 463-464.
5. Daniels AH, Ames CP, Garfin SR, et al. Spine surgery training: Is it time to consider categorical spine surgery residency? Spine J. 2015;15:1513–8. DOI: 10.1016/j.spinee.2014.08.452.
6. Wacker FK, Vogt S, Khamene A, et al. An augmented reality system for MR image-guided needle biopsy: Initial results in a swine model. Radiology. 2006;238:497–504. DOI: 10.1148/radiol.2382041441.
7. Tagaytayan R, Kelemen A, Sik-Lanyi A, C. Augmented reality in neurosurgery. Arch Med Sci. 2018;14:572–8. DOI: 10.5114/aoms.2016.58690.
8. Wu J-R, Wang M-L, Liu K-C, Hu M-H, Lee P-Y. Real-time advanced spinal surgery via visible patient model and augmented reality system. Comput Methods Programs Biomed. 2014;113:869–81. DOI: 10.1016/j.cmpb.2013.12.021.
9. Hughes CL, Fidopiastis C, Stanney KM, Bailey PS, Ruiz E. The psychometrics of cybersickness in augmented reality. Frontiers in Virtual Reality; 2020. 1. https://www.frontiersin.org/articles/10.3389/frvir.2020.602954, accessed December 18, 2022.
10. Brunnström K, Dima E, Qureshi T, Johanson M, Andersson M, Sjöström M. Latency impact on quality of experience in a virtual reality simulator for remote control of machines. Signal Process Image Commun. 2020;89:116005. DOI: 10.1016/j.image.2020.116005.
11. Grosch AS, Schröder T, Schröder T, Onken J, Picht T. Development and initial evaluation of a novel simulation model for comprehensive brain tumor surgery training. Acta Neurochir. 2020;162:1957–65. DOI: 10.1007/s00701-020-04359-w.
12. Cao C, Cerfolio RJ. Virtual or augmented reality to enhance surgical education and surgical planning. Thorac Surg Clin. 2019;29:329–37. DOI: 10.1016/j.thorsurg.2019.03.010.
13. Reed DA, Cook DA, Beckman TJ, Levine RB, Kern DE, Wright SM. Association between funding and quality of published medical education research. JAMA. 2007;298:1002–1009. DOI: 10.1001/jama.298.9.1002.
14. Hou Y, Shi J, Lin Y, Chen H, Yuan W. Virtual surgery simulation versus traditional approaches in training of residents in cervical pedicle screw placement. Arch Orthop Trauma Surg. 2018;138:777–82. DOI: 10.1007/s00402-018-2906-0.
15. Hou Y, Lin Y, Shi J, Chen H, Yuan W. Effectiveness of the thoracic pedicle screw placement using the virtual surgical training system: A cadaver study. Oper Neurosurg (Hagerstown). 2018;15:677–85. DOI: 10.1093/ons/opy030.
16. Shi J, Hou Y, Lin Y, Chen H, Yuan W. Role of visuohaptic surgical training simulator in resident education of orthopedic surgery. World Neurosurg. 2018;111:e98–e104. DOI: 10.1016/j.wneu.2017.12.015.
17. 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;S1878-8750:32913–9. DOI: 10.1016/j.wneu.2018.12.090.
18. Luca A, Giorgino R, Gesualdo L, et al. Innovative educational pathways in spine surgery: Advanced virtual reality-based training. World Neurosurg. 2020;140:674–80. DOI: 10.1016/j.wneu.2020.04.102.
19. Chen T, Zhang Y, Ding C, et al. Virtual reality as a learning tool in spinal anatomy and surgical techniques. N Am Spine Soc J. 2021;6:100063. DOI: 10.1016/j.xnsj.2021.100063.
20. Knafo S, Penet N, Gaillard S, Parker F. Cognitive versus virtual reality simulation for evaluation of technical skills in neurosurgery. Neurosurg Focus. 2021;51:E9. DOI: 10.3171/2021.5.FOCUS201007.
21. Yanni DS, Ozgur BM, Louis RG, et al. Real-time navigation guidance with intraoperative CT imaging for pedicle screw placement using an augmented reality head-mounted display: A proof-of-concept study. Neurosurg Focus. 2021;5:E11. DOI: 10.3171/2021.5.FOCUS21209.
22. Dennler C, Jaberg L, Spirig J, et al. Augmented reality-based navigation increases precision of pedicle screw insertion. J Orthop Surg Res. 2020;15:174. DOI: 10.1186/s13018-020-01690-x.
23. 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:ons14–19;. DOI: 10.1227/NEU.0b013e31821954ed discussion ons19.
24. Luciano CJ, Banerjee PP, Sorenson JM, et al. Percutaneous spinal fixation simulation with virtual reality and haptics. Neurosurgery. 2013;72 Suppl 1:89–96. DOI: 10.1227/NEU.0b013e3182750a8d.
25. Yu H, Zhou Z, Lei X, Liu H, Fan G, He S. Mixed reality-based preoperative planning for training of percutaneous transforaminal endoscopic discectomy: A feasibility study. World Neurosurg. 2019;129:e767–e775. DOI: 10.1016/j.wneu.2019.06.020.
26. Ledwos N, Mirchi N, Bissonnette V, et al. Virtual reality anterior cervical discectomy and fusion simulation on the novel sim-ortho platform: Validation studies. Oper Neurosurg (Hagerstown). 2020;20:74–82. DOI: 10.1093/ons/opaa269.
27. Mirchi N, Bissonnette V, Ledwos N, et al. Artificial neural networks to assess virtual reality anterior cervical discectomy performance. Oper Neurosurg (Hagerstown). 2020;19:65–75. DOI: 10.1093/ons/opz359.
28. Alkadri S, Ledwos N, Mirchi N, et al. Utilizing a multilayer perceptron artificial neural network to assess a virtual reality surgical procedure. Comput Biol Med. 2021;136:104770. DOI: 10.1016/j.compbiomed.2021.104770.
29. Bissonnette V, Mirchi N, Ledwos N, Alsidieri G, Winkler-Schwartz A, Del Maestro RF. Artificial intelligence distinguishes surgical training levels in a virtual reality spinal task. J Bone Joint Surg Am. 2019;101:e127. DOI: 10.2106/JBJS.18.01197.
30. Liu T, Tai Y, Zhao C, et al. Augmented reality in neurosurgical navigation: A survey. Int J Med Robot Comput Assist Surg. 2020;16:e2160–20. DOI: 10.1002/rcs.2160.
31. Winkler-Schwartz A, Bissonnette V, Mirchi N, et al. Artificial intelligence in medical education: best practices using machine learning to assess surgical expertise in virtual reality simulation. J Surg Educ. 2019;76:1681–90.
32. Winkler-Schwartz A, Yilmaz R, Mirchi N, et al. Machine learning identification of surgical and operative factors associated with surgical expertise in virtual reality simulation. JAMA Netw Open. 2019;2:e198363–e198363. DOI: 10.1001/jamanetworkopen.2019.8363.