REVIEW ARTICLE - NEUROSURGERY TRAINING

Evaluation of simulation models in neurosurgical training according to face, content, and construct validity: a systematic review

Shreya Chawla[1,2] · Sharmila Devi[1,2,3] · Paola Calvachi[1,4] · William B. Gormley[1] · Roberto Rueda-Esteban[1,5]

Abstract

Background Neurosurgical training has been traditionally based on an apprenticeship model. However, restrictions on clinical exposure reduce trainees' operative experience. Simulation models may allow for a more efficient, feasible, and time-effective acquisition of skills. Our objectives were to use face, content, and construct validity to review the use of simulation models in neurosurgical education.

Methods PubMed, Web of Science, and Scopus were queried for eligible studies. After excluding duplicates, 1204 studies were screened. Eighteen studies were included in the final review.

Results Neurosurgical skills assessed included aneurysm clipping ($n=6$), craniotomy and burr hole drilling ($n=2$), tumour resection ($n=4$), and vessel suturing ($n=3$). All studies assessed face validity, 11 assessed content, and 6 assessed construct validity. Animal models ($n=5$), synthetic models ($n=7$), and VR models ($n=6$) were assessed. In face validation, all studies rated visual realism favourably, but haptic realism was key limitation. The synthetic models ranked a high median tactile realism (4 out of 5) compared to other models. Assessment of content validity showed positive findings for anatomical and procedural education, but the models provided more benefit to the novice than the experienced group. The cadaver models were perceived to be the most anatomically realistic by study participants. Construct validity showed a statistically significant proficiency increase among the junior group compared to the senior group across all modalities.

Conclusion Our review highlights evidence on the feasibility of implementing simulation models in neurosurgical training. Studies should include predictive validity to assess future skill on an individual on whom the same procedure will be administered. This study shows that future neurosurgical training systems call for surgical simulation and objectively validated models.

Keywords Surgical; Surgical simulation · Neurosurgical simulation · Neurosurgical education · Neurosimulation · Residency training · Construct/content/face validity

---

✉ Roberto Rueda-Esteban
rj.rueda32@uniandes.edu.co

1   Computational Neuroscience Outcomes Center, Department of Neurosurgery, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

2   Faculty of Life Sciences and Medicine, King's College London, London, UK

3   Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

4   Department of Biomedical Informatics, Harvard Medical School, Boston, MA, USA

5   Universidad de los Andes School of Medicine, Bogotá, Colombia

Introduction

Simulation is an educational technique where a trainee interacts with an environment that either recreates or replicates a real-world clinical scenario in a risk-free atmosphere [28, 30, 37]. Simulation models for medical education include human and animal cadavers and live animal models,

synthetic/physical bench models, and virtual reality (VR) models [21, 30]. Simulation has been rapidly utilised in surgery, especially in high-income countries [65].

Neurosurgical training has been traditionally based on a model of apprenticeship ("see one, do one, teach one") [12, 16], where theoretical and practical learning take place in operating rooms (OR) [28, 30]. However, challenges associated with this apprenticeship system include financial constraints of the healthcare system and teaching in the OR [40], restrictions on working hours [14], worsening patient outcomes [52], and limited clinical exposure and operation opportunities due to increasing ethical and medico-legal constraints [20, 55]. Neurosurgeons perform a mean number of 223 cases a year, which are widely varied and unique in technical competence [23]. Together with the restrictions on training hours, trainees might not encounter a similar procedure often, limiting opportunities for surgical training. This has been exacerbated by the COVID-19 pandemic, as seen by a marked reduction in elective and non-essential neurosurgical cases and redeployment of neurosurgeons to the intensive care unit, overwhelming even affluent healthcare

systems [26, 56, 62, 67]. This has further reduced opportunities to perform neurosurgical procedures, impacting training [10]. Therefore, it is more relevant now than ever to continue developing simulation models to not only allow trainees to repeatedly perform procedures in a controlled environment, but also remotely (Fig. 1).

Medical simulation serves as an alternative for time-effective acquisition of skills [2, 21, 22], with the resultant shift of learning curve away from the patient [2, 21]. Simulation allows trainees to learn from and make errors in a safe environment as incorrect or technically demanding tasks can be performed to completion [17, 21, 30]. The efficacy of simulation on acquisition and development of technical skills is shown by improvement of objective performance metrics when skills learned from simulator-training are translated into the OR [16, 17, 21]. However, medical and surgical simulations must be evaluated as educational tools. Evaluation of simulation models is based on subjective and objective validation [68]. Face and content validity are types of subjective validity, evaluated by questionnaires. Face validity examines the realism of a simulator and difficulty level

Fig. 1 Overview of the diverse range of available neurosurgical training tools. These educational approaches can generally be subdivided into physical models (including both biological tissues and synthetic constructs) and simulated approaches (which include both virtual and augmented reality)
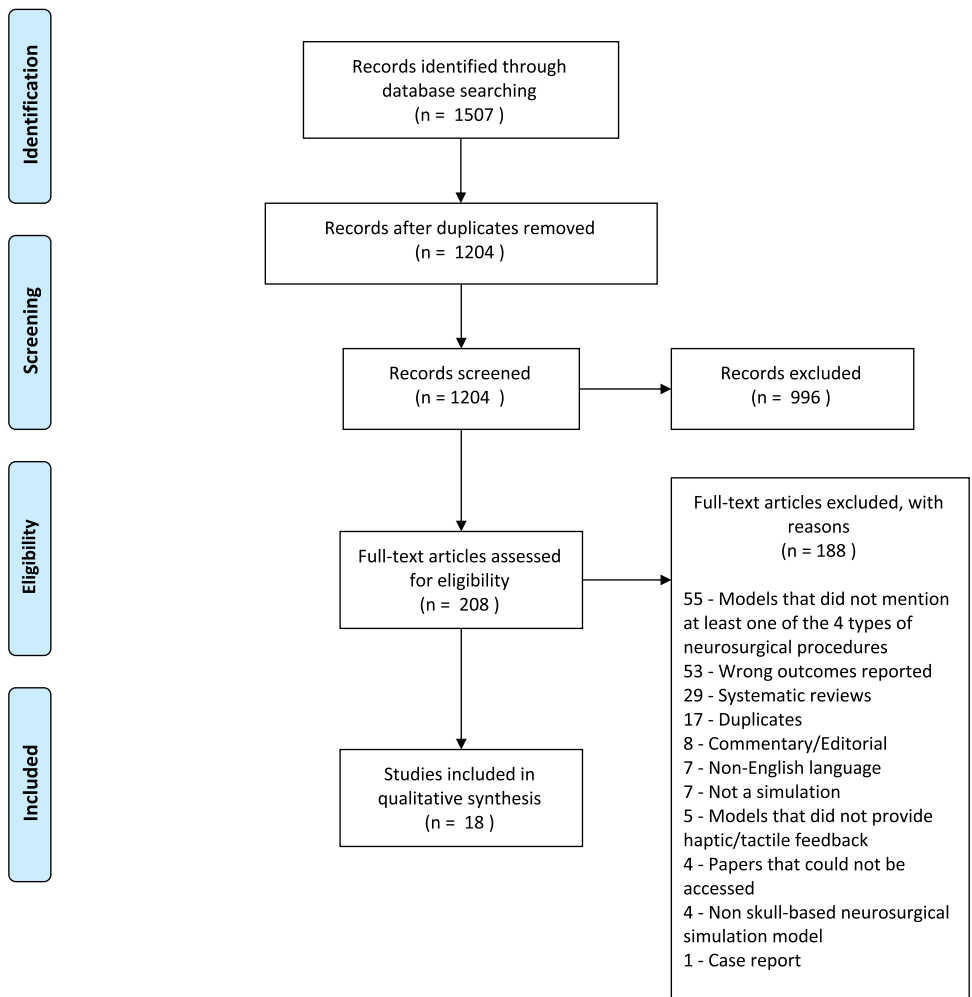
similarity in comparison to real training tasks, whilst content validity assesses the model's effectiveness during a specific skill training to improve participants' techniques [68]. Whilst these are subjective methods, construct validity is an objective method that considers the ability of a simulator to differentiate levels of skill competence [22, 47].

Previous reviews have explored developing neurosurgical simulation-based training [16, 30, 49], discussing the strengths and limitations of various simulation models. Simulations included in the reviews are also mostly limited to either physical or virtual reality simulators, showing that there is a gap in literature in assessing an extensive variety of simulator types. Furthermore, the type of neurosurgical procedures included is random, limiting our understanding of which neurosurgical skill or procedure can be best simulated by a specific simulator type. The ultimate goal of simulation is to ensure that skills learnt from the simulator can be transferred to the OR, inadequately explored in these reviews. Important work on simulation models is underway but current studies examining surgical simulation methods do not have a comprehensive approach to applying validation methods to assess simulator models. Whilst there is early cause for simulation to develop surgical skills, there is a gap in literature in establishing the use-case and quantifying the impact of surgical education on surgical skill acquisition and improvement.

This study, therefore, aims to reduce the gap in literature by reviewing simulation models in neurosurgical education and training for specific neurosurgical procedures: burr hole incision/craniotomy/craniectomy, aneurysm clipping, vessel suturing, skull-based tumour resection using face, content, and construct validity. These 4 skull-base neurosurgical procedures were chosen from documental analysis as they were most extensively evaluated in literature. These neurosurgical procedures are also part of the neurosurgical curriculum guidelines set by national associations including the Congress of Neurological Surgeons, American Association of Neurological Surgeons, General Medical Council, and Royal Australasian College of Surgeons.

Materials and methods

Search strategy

A comprehensive search was performed in PubMed, Web of Science, and Scopus, in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines up until September 2019 using predefined search terms. Search terminology for PubMed has been listed in Appendix 1. Identified abstracts and full-text papers were reviewed by two independent reviewers (SC, SD) against predefined inclusion and exclusion criteria to assess eligibility using Covidence® software [30]. A third reviewer (RR) resolved any discrepancy that occurred between the 2 authors after full-text screening. Our protocol has been published on PROSPERO under number: CRD42020141703.

Assessment of eligibility

Retrospective and prospective observational studies, randomised controlled trials, and case series that reported simulations used in skull-based neurosurgery education and/or training with reported outcomes were included. Case reports, systematic reviews, and articles written in languages other than English were excluded. Articles mentioning participants from any specialty and training level using the neurosurgical simulation models were included in the study. Articles that mentioned any neurosurgical simulation model for 4 such neurosurgical procedures (burr hole incision/craniotomy/craniectomy, aneurysm clipping, vessel suturing, and tumour resection) were selected. Studies that reported measurable outcomes with validated assessment tools of skill acquisition were included. Simulation models that did not report haptic feedback were excluded.

Data extraction

A pre-designed excel sheet was used to extract and organise data into categories by two independent authors (SC, SD). Data extracted included study characteristics (name of first author, year of publication, title, name of journal, study design), population characteristics, model characteristics (type of neurosurgical skill, neurosurgical subspecialty), and outcome evaluation (type of assessment tool, reported outcomes relating to face, content, and construct validity) (Tables 1 and 2).

Quality assessment

The Medical Education Research Study Quality Instrument (MERSQI) was used for a methodological evaluation of included studies [18, 53]. MERSQI is a reliable tool created for the critical appraisal of medical education research [18, 63]. The maximum number of points scored on the MERSQI scale is 18 points.

This tool evaluates study design, sampling, data type, validity of evaluation, data analysis, and study outcomes. The MERSQI tool was used due to the high validity of this tool, assessed using 3 criteria. These included correlating global quality ratings from 2 independent experts in medical education research, examining the association between MERSQI scores and the impact factor of the publishing journal, and performing a simple linear regression to measure the association between MERSQI scores and citation rate and impact factor. This assessment found that total MERSQI

Table 1 Study characteristics of included papers

No	Author, year	Country	Title	Simulation type	Procedure type
1	Aboud et al., 2015	USA	"Live cadavers" for training in the management of intra-operative aneurysmal rupture	Cadaver: lifelike model connected to a pump that sent artificial blood into the vessels	Aneurysm clipping
2	Alaraj et al., 2015	USA	Virtual reality cerebral aneurysms clipping simulation with real-time haptic feedback	VR: 3D VR with immersive touch platform	Aneurysm clipping, craniotomy
3	Aoun et al., 2015	USA	A pilot study to assess the construct and face validity of the Northwestern Objective Microanastomosis Assessment Tool	Synthetic vessel model	Vessel suturing (microanastomosis)
4	Ashour et al., 2016	USA	Navigation-guided endoscopic intraventricular injectable tumour model: cadaveric tumour resection model for neurosurgical training	Synthetic tumour model	Tumour resection
5	Belykh et al., 2016	USA	Low-flow and high-flow neurosurgical bypass and anastomosis training models using human and bovine placental vessels: a histological analysis and validation study	Animal model: placenta	Vessel suturing (mircoanastomosis)
6	Belykh et al., 2017	USA	Face, content and construct validity of an aneurysm clipping model using human placenta	Animal model: placenta	Aneurysm clipping
7	Craven et al., 2014	UK	Development of a modelled anatomical replica for training young neurosurgeons	Synthetic model: Modelled Anatomical Replica for Training Young Neurosurgeons (MARTYN)	Craniotomy, burr hole
8	De Oliveira et al., 2018	Canada, Brazil, USA	Learning brain aneurysm microsurgical skills in a human placenta model: predictive validity	Cadaver and animal model—placenta	Vessel suturing; aneurysm clipping
9	Gelinas-Phaneuf et al., 2014	Ireland	Assessing performance in brain tumour-resection using a novel virtual reality simulator	VR simulation—NeuroTouch	Tumour resection (meningioma)
10	Gmeiner et al., 2018	Austria	Virtual cerebral aneurysm clipping with real time haptic force feedback in neurosurgical education	VR simulation—aneurysm geometrics, MEDVIS 3D	Aneurysm clipping
11	Jaimovich et al., 2016	Argentina	Neurosurgical training with simulators: a novel neuroendoscopy model	Animal model—live rats	Tumour resection
12	Liu et al., 2017	China	Fabrication of cerebral aneurysm simulator with a desktop 3D printer	Synthetic model: 3D printed aneurysm	Aneurysm clipping
13	Mashiko et al., 2015	Japan	Development of three-dimensional hollow elastic model for cerebral aneurysm clipping Simulation enabling rapid and low-cost prototyping	Physical simulation—3D printed hollow and elastic aneurysm model	Aneurysm clipping
14	Muens et al., 2014	Germany	A neurosurgical phantom-based training system with ultrasound simulation	VR simulation–phantom-based training system	Craniotomy
15	Ryan et al., 2016	USA	Cerebral aneurysm clipping surgery simulation using patient-specific 3D printing and silicone casting	Physical simulation—3D printed model	Craniotomy, aneurysm clipping
16	Vloeberghs et al., 2007	UK	Virtual neurosurgery, training for the future	VR simulation—boundary elements	Tumour resection
17	Wong et al., 2014	Canada	Comparison of cadaveric and isomorphic virtual haptic simulation in temporal bone training	VR simulation—virtual isomorphic haptic model	Burr hole
18	Wang et al., 2018	China	3D printing of intracranial aneurysm based on intracranial digital subtraction angiography and its clinical application	Physical simulation—3D printed model	aneurysm clipping

scores were highly correlated with the median quality rating of the 2 independent experts ($p = 0.73$) and agreement between the 2 experts was excellent (ICC, 0.80; 95% CI, 0.49–0.85). The scores were significantly associated with a high 3-year citation rate and journal impact factor [53].

Results

A total of 1507 publications were initially identified across all searched databases. After removal of duplicates, 1204 studies were identified for title and abstract screening. Of these, 208 studies were eligible for full text screening and 18 studies were included in the final review (Fig. 2). General characteristics of the study and the type of simulation and procedure are presented in Table 1. Model characteristics assessed by face, content, and construct validity can be found in Table 2.

The studies conducted by Belykh et al. (2016) [12] and Belykh et al. (2017) [13] potentially used the same cohort to assess neurosurgical skills on the placenta. Despite this overlap, Belykh et al. (2016) [12] assessed vessel suturing, and Belykh et al. (2017) [13] assessed aneurysm clipping. Therefore, both studies were included in this review as the studies assessed different neurosurgical skills.

Study characteristics

Simulation models

The neurosurgical skills assessed were aneurysm clipping ($n = 6$) [1, 13, 27, 39, 42, 70], craniotomy and burr hole drilling ($n = 2$) [19, 46], tumour resection ($n = 4$) [7, 25, 33, 69] and vessel suturing—vascular anastomosis ($n = 1$) [22], vessel suturing—micro-anastomosis ($n = 2$) [6, 13]. Multiple neurosurgical procedures assessed by the studies included aneurysm clipping and craniotomy ($n = 1$) [1] and vessel suturing and aneurysm clipping ($n = 1$) [22] (Table 1). These skills were simulated on animal models ($n = 5$), synthetic models ($n = 7$) [6, 7, 19, 39, 42, 60, 70], and using VR ($n = 6$) [3, 25, 27, 46, 69, 72]. Animal models included cadavers ($n = 2$) [1, 22], animal placenta ($n = 3$) [12, 13, 22], and live rats ($n = 1$) [33].

Model validation

All studies assessed face validity, 11 studies assessed content validity [1, 3, 7, 22, 25, 27, 33, 39, 46, 60, 69], and 6 studies assessed construct validity of the models [6, 12, 13, 19, 25, 69]. Assessment of construct validity involved distinguishing the improvement of participants at different levels of experience and expertise. The mean MERSQI score of

included studies was 10.5. Complete results of the MERSQI tool can be found in Appendix 2.

Sub-group analysis

Aneurysm clipping

Aneurysm clipping in five included studies was simulated by a cadaver model [1], placenta model [13], 3D printed models ($n = 3$) [39, 42, 70], and one VR model [27] (Table 1). Belykh et al., 2017 [13] used the Objective Structured Assessment of Aneurysm Clipping Skills tool and Aneurysm Clipping Participant Survey, whilst other studies used either a 5-point [1, 27, 39] or a 4-point Likert scale questionnaire [42]. The scale used in all studies indicated that 1 referred to participants strongly disagreeing with the question, and the highest value indicating strong agreement.

All studies validated their models using face validity [1, 13, 27, 39, 42, 70] and 3 studies used both face and content validity [1, 27, 39] (Table 2). All studies conferred on participants either agreeing or strongly agreeing that the model was a true simulation of the aneurysm clipping in a surgical environment. Eighty to ninety-nine per cent of participants favourably reported that the model realistically simulated anatomy of aneurysm [1, 39], whilst 89–100% of participants agreed that simulation models to train aneurysm clipping should be integrated in neurosurgical training [1, 27, 39]. Content validity was assessed via questionnaires on the physical aspects of the simulation model. Respondents ranked their understanding of structure and location of an aneurysm favourably as either "excellent" or with a median score of 4.8 out of 5 [42]. Compared to existing models of live animals for micro-anastomosis training, the models included were rated as superior by the majority of participants (97–99%) [1]. Only Belykh et al. (2017) assessed construct validity, showing that the participants with the least experience scored 0–28 points, whilst the group with the intermediate experience scored 29–39 and the attending group was at 40–45 point-intervals using the model [13].

Craniotomy/burr hole drilling

Craniotomy and burr hole drilling were assessed using two VR models and one synthetic model [19]. The VR models included a phantom-based training system [46], and a virtual isomorphic haptic model [72]. Face validity was assessed by all studies, content validity was assessed in 1 study [46], and another study used construct validity to assess the proficiency of craniotomy/burr hole drilling among their participants [19] (Table 1). Using a 5-point Likert scale to assess face validity, participants found a high median visual and tactile realism with respondents ranking the tool as "useful"

Table 2 Model characteristics assessed by face, content, and construct validity

Author, year	Participant level, n	Assessment	Reported outcome	Scale		
				Face	Content	Construct
Aboud et al., 2015	Residents, 203; attendings, 89	Questionnaire	5-point Likert scale regarding accuracy and realism of model (1 = strongly disagree, 5 = strongly agree)	1.) The model was a true simulation of the conditions of live surgery on aneurysms a. 1.09% disagree b. 2.19% neutral c. 28.57% agree d. 68.13% strongly agree); 2.) This model promotes the acquisition of microsurgical skills a. 6.59% neutral, b. 21.97% agree, c. 71.42% strongly agree) 3.) This model offers benefits not available in existing training models a. 6.59% neutral b. 21.97% agree c. 71.42% strongly agree) 4.) This model could significantly improve current training in the management of intraoperative cerebrovascular complications a. 24.17% agree b. 75.82% strongly agree 5.) This model could add significantly to training in microneurosurgical techniques a. 26.37% agree b. 73.62% strongly agree 6.) This model will be a valuable addition to the medical device development and testing process a. 1.09% strongly disagree b. 2.19% disagree c. 7.69% neutral d. 23.07% agree e. 65.93% strongly agree	1.) The scenario of aneurysm clipping, and intraoperative rupture is realistic a. 7.69% neutral b. 30.76% agree c. 61.53% strongly agree 2.)This model is superior to existing models for cerebral revascularisation a. 1.09% disagree b. 4.39% neutral c. 27.47% agree d. 67.03% strongly agree 3.) This model could replace the use of live animals in microanastomosis training a. 3.29% strongly disagree b. 5.49% disagree c. 15.38% neutral d. 30.76% agree e. 45.05% strongly agree	

Table 2 (continued)

Author, year	Participant level, n	Assessment	Reported outcome	Scale — Face	Content	Construct
Alaraj et al., 2015	Residents, 17	Questionnaire	1. Dichotomous response items: Yes/No 2. 5-point Likert scales (5 = highest, 1 = lowest) 3. Free text responses	1.) The ITACS Immersive Touch Aneurysm Clipping Simulator is a useful tool a. 12% disagree b. 24% neutral c. 29% agree d. 35% strongly agree 2.) On the whole, the aneurysm simulator will help them in preparing for aneurysm clipping surgery if they have time to rehearse the same procedure on a patient-specific model a. 12% do not know b. 12% disagree c. 17% neutral d. 47% agree e. 12% strongly agree	1.) The ITACS can increase their understanding of aneurysm anatomy a. 18% disagree b. 18% neutral c. 29% agree d. 35% strongly agree 2.) Agreed the haptic sensation produced by the simulator is identical to the one encountered in real surgery a. 12% do not know b. 18% strongly disagree c. 29% disagree d. 29% neutral e. 12% agree 3.) Felt that the aneurysm simulation module would help define which approach should be used to access the aneurysm safely a. 23% disagree b. 6% neutral c. 59% agree d. 12% strongly agree 4.) The 3-D anatomy on the simulator represents the real anatomy a. 24% disagree b. 29% neutral c. 35% agree d. 12% agree 5.) Ability to operate the haptic stylus a. 6% strongly disagree b. 18% disagree c. 29% neutral d. 35% agree f. 12% agree	
Aoun et al., 2015	Postdoctoral research "experienced", 6 neurosurgical residents "exposed", 6 medical students "novice", 9	Northwestern Objective Microanastomosis Assessment Tool (NOMAT)	1.) Face validity: free response			1-mm microanastomosis: Mean NOMAT score (range) 1. Experienced: 47.3 (20–64) 2. Exposed: 26.0 (17–40) 3. Novice: 25.8 (16–36) 3-mm microanastomosis Mean NOMAT score (range) 1. Experienced: 47.8 (33–66) 2. Exposed: 45.0 (31–63) 3. Novice: 39.6 (25–50)

Table 2 (continued)

Author, year	Participant level, n	Assessment	Reported outcome	Scale		Construct
				Face	Content	
Ashour et al., 20,016	NA	Questionnaire used to generate score (%)	NA	1.) Usefulness in neurosurgical training 96% 2.) Utilisation of similar surgical approaches 98% us 3.) Use of instruments and microsurgical techniques 94%	1.) Consistency (of tumour) 90% 2.) Relation to adjacent neural and vascular structures 92% 3.) Challenges in resection 86% 4.) Radiographic visibility for adequate planning 82% 5.) Use of adjuncts to surgical resection 86% 6.) Skills improvement in dealing with real tumours 94%	NA
Belykh et al., 2016	Medical students, 10 Residents without experience "untrained", 3 Practicing neurosurgeons 10 Residents with microsurgical experience, 7	Bypass Participant Survey	1. Face: Bypass Participant Survey (1–20) 2. Content: Bypass Participant Survey (1–20) 3. Construct: NOMAT score	1.) Ability of the training model to replicate real bypass surgery: a) More than somewhat (score 13)–very well (score 20): 16/17 (94%) trained, 12/13 (92%) untrained b) Somewhat replicating (score 8–12): trained 1/17 (6%), 1/13 (8%) of untrained 2.) Difficulty of the task compared to real surgery: a) Nearly "the same" (score 5–15): 16/17 (94%) trained, 11/13 (85%) untrained 3.) Ability of the model to improve microsurgical techniques and instrument handling: a) Answered positively (score 15–20): 30/30 (100%) 4.) Model's ability to improve surgical technique when the skills were applied to patients (scores 15–20): a) Considered it able to improve: 30/30 (100%)	• NA (paper mentioned questions under "face validity" to be both "face and content validity")	Mean NOMAT 1) Untrained: 37.2 ± 7.0 2) Trained: 62.7 ± 6.1 Mean NOMAT of untrained group statistically significantly lower than the mean NOMAT score of the trained group
Belykh, et al., 2017	"Low-experience", 10 "Intermediate-experience", 7 "Attending", 10	Objective Structured Assessment of Aneurysm Clipping Skills (OSAACS) tool and Aneurysm Clipping Participant Survey	OSAACS: scale 1–5, total 45 points (1 worst, 5 best) Aneurysm Clipping Participant Survey 10-item measure; each question scored on a 20 point scale			Low experience: 22.9 ± 5.3 Intermediate experience: 32.8 ± 4.4 Attending: 43.25 ± 1.3

Table 2 (continued)

Author, year	Participant level, n	Assessment	Reported outcome	Scale		
				Face	Content	Construct
Craven et al., 2014	ST1 neurosurgical trainees and non-neurosurgeons "novices", 9 ST2 3 neurosurgical trainees or those who have done more than 10 burr holes and craniotomies "intermediates", 4 Trainees who have completed more than 30 craniotomies, 5	Questionnaire	5 Point Likert scale (1 = unrealistic and 5 = highly realistic)	1.) Median visual realism (overall) = 4 2.)Median tactile realism (overall) = 4 3.)Would participants recommend use of MARTYN model to your colleagues and whether it was useful a. Useful = 18, with 11 of these describing the model as "extremely useful" (or an equivalent term) for neurosurgical training		1. All individuals in the novice group experienced a significant increase in confidence to perform craniotomy after the training on the MARTYN model (confidence rating 1 before MARTYN model training to 3 after training) 2. All individuals in the intermediate group experienced a significant increase in confidence to perform craniotomy after the training on the MARTYN model (confidence rating 3.5 before MARTYN training to 4 after training) 2. Individuals in the "experienced" (senior registrars and above) group reported no significant increase in their self-assessed confidence despite MARTYN training in craniotomy (from 4 prior to MARTYN training to 5 post training)
Gelinas-Phaneuf, 2014	(Medical students, 10); (junior residents aka PGY1-3, 18); (senior residents PGY4 and above, 44)	Questionnaire	5-point Likert scale	Results in mean ± SD 1. Would use simulator if available in training program?: 3.9 ± 1.2 2. Visual realism 3.3 ± 1.0 3. Sensory realism 2.9 ± 1.05 4. Overall satisfaction 3.7 ± 0.94	1. Difficulty of challenge: 3.5 ± 0.94 2. Appropriate metrics: 3.7 ± 0.94	1) % of tumour removed a) Significant difference between medical students and junior residents ($p=0.034$) b) Significant difference between medical students and senior residents ($p=0.002$) c) No significant difference between junior and senior residents 2) Efficiency of ultrasonic aspiration a) Medical students under-performed b) No significant difference between the junior and senior residents in this metric of performance
Gmeiner et al., 2018	Residents, 4 Trained surgeons, 14	Questionnaire	Free response 5-point Likert scale; 1 = strongly disagree 5 = strongly agree	Many users considered simulation of head positioning (89%), simulation of craniotomy (94%), and realism of the 3D surgical situs (44%) as adequate 61% found the evaluation of results and scoring system realistic and helpful 17% agreed that real-life aneurysmclipping could be learned exclusively by virtual aneurysm clipping 94% of participants agreed that virtual aneurysm clipping simulator should be integrated in neurosurgical education 100% of participants stated that it should be used in daily routine before anuerysm clipping	89% of participants both agreed and strongly agreed that the virtual aneurysm clipping simulator improves anatomic understanding The clipping procedure itself was considered adequate by 22% and acceptable by 50% of participants 1/3rd of participants strongly agreed and agreed that the haptic interaction with the vessels and the aneurysm during virtual clipping was truly satisfactory	

Table 2 (continued)

Author, year	Participant level, n	Assessment	Reported outcome	Scale		Construct
				Face	Content	
Liu et al., 2017	Medical students, 4 Neurosurgeons, 6	Questionnaire	% Satisfaction	1) Size of simulator: 100% satisfaction 2) Haptic anatomy of the simulator: 80% satisfaction 3) Visual appearance of the simulator: 90% satisfaction 4) Teaching: 100% 5) Learning: 100% 6) surgical training: 90% 7) Would you use the simulator: 100%	Improving understanding the structure of the anuerysm's relationship to the parent artery: 90%	
Mashiko et al., 2015	Neurosurgeons, 18	Questionnaire	Poor/fair/good/ Excellent	Trained surgeon: 1.) Understanding of the structure of the aneurysm: (i) Solid: excellent (6); good (5); fair (1); poor (0) (ii) Elastic: excellent (8); good (4); fair (0); poor (0) 2.) Ease of handling [of the aneurysm model] (i) Solid: excellent (6); good (6); fair (0); poor (0) (ii) Elastic: excellent (7); good (5); fair (0); poor (0) Junior surgeon 1.) Understanding of the structure of the aneurysm: (i) Solid: excellent (4); good (2); fair (0); poor (0) (ii) Elastic: excellent (5); good (1); fair (0); poor (0) 2.) Ease of handling [of the aneurysm model] (i) Solid: excellent (5); good (1); fair (0); poor (0) (ii) Elastic: excellent (5); good (1); fair (0); poor (0)	NA	NA
Muens et al., 2014	Neurosurgery Residents (PGY2 – PGY7), 5	Questionnaire	5-point Likert scale (1 = best; 5 = worst)	Skin Question, average, (SD) Colour: 2.20 (1.10) Width: 1.40 (0.55) Haptics: 3.40 (0.55) Cutting sensation: 3.40 (1.14) Tear strength: 2.20 (0.45) Bone Colour: 1.40 (0.55) Haptics: 1.40 (0.55)	Skin Question, average, (SD) Removability from bone: 2 (1.22) Adhesive residue bone: 1.60 (0.89) Removability from muscle: 2.50 (0.58) Adhesive residues muscle: 2 (0.71) Skin suture: 2.60 (0.89) Bone Authentic drilling: 1.20 (0.45) Authentic milling: 1.20 (0.45)	NA

Table 2 (continued)

Author, year	Participant level, *n*	Assessment	Reported outcome	Scale		Construct
				Face	Content	
Ryan et al., 2016	Neurosurgery Residents (unspecified), 14	Questionnaire	5-point Likert scale (1 = worst; 5 = best)	Is the simulator clinically applicable?: 4.4 (4–5) Did it improve your understanding of the surgical view?: 4.5 (4–5) Did clip application seem realistic? 4.1 (3–5) Did the bone drill in a realistic manner?: 4.1 (2–5) Was the simulator useful to you?: 4.6 (4–5) Do you think your surgical skills would improve with practice using the simulator?: 4.4 (4–5)	Did the simulator improve your understanding of the aneurysm's relationship to the parent artery?: 4.4 (3–5)	
Vloeberghs et al., 2007	Neurosurgeons (Unspecified), 13	Questionnaire	5-point Likert scale (1 = best; 5 = worst)	Mean, (SD) In general, how easy was the simulator to use?: 1.31 (0.48) How realistic did the brain look whilst pushing'?: 2.15 (0.55) How realistic did pushing the brain feel?: 2.46 (0.5) 2 How easy was pushing the brain?: 2.08 (0.76) How realistic did the brain look whilst pulling?: 2.62 (0.65) How realistic did pinching the brain tissue feel? 2.77 (0.6) How easy was pulling the brain? 2 (0.74) How realistic did the brain look whilst cutting? 3 (0.82) How realistic did cutting the brain feel? 3.38 (0.51) How easy was cutting the brain? 2.38. (0.87) How realistic was the stereo viewing? 1.77 (0.73) How comfortable was the physical set-up? 1.54 (0.66)	Mean, (SD) How easy was pushing the brain?: 2.08 (0.76) How easy was pulling the brain? 2 (0.74) How easy was cutting the brain? 2.38. (0.87)	NA

Table 2 (continued)

Author, year	Participant level, n	Assessment	Reported outcome	Scale			
				Face	Content	Construct	
Wong et al., 2014	Otolaryngology residents, 10	Questionnaire	7-point Likert scale (1 = not beneficial; 7 = very beneficial)	Resident assessment of virtual model physical properties as compared to cadaveric bone (mean ± SD)	NA	NA	
				Overall similarity to cadaveric temporal bone 3.5 ± 1.8			
				Resident appraisal of virtual model educational value			
				This is an effective training instrument 5.4 ± 1.5			
				This instrument is an accurate reproduction of the temporal bone 5.7 ± 1.4			
				This instrument should be integrated into resident education 5.5 ± 1.4			
				This form of simulation can replace the cadaveric temporal bone lab 2.5 ± 2.3			
				This simulation provides a basis for appreciating the relative anatomy of temporal bone structures 6.1 ± 1.9			
				This simulated surgery improves confidence 5.3 ± 1.9			
				Increased exposure to this simulation would improve resident surgical performance 5.3 ± 1.8			
				Increased exposure to this simulation would improve resident comfort with actual patient surgery 5.4 ± 1.8			
				This simulation facilitates practice of skills across a range of anatomical and pathologic variations (sclerotic, low dura, disease) 5.6 ± 1.8			

Table 2 (continued)

Author, year	Participant level, n	Assessment	Reported outcome	Scale		Content	Construct
				Face			
Wang et al., 2018	Neurosurgery residents, 7 Standardisation training residents with other specialisation, 15	Questionnaire	5-point Likert scale (1 = lowest; 5 = highest)	Scores: mean (range) Is the simulation model clinically applicable? 4.4 (4–5) Did the simulation help you to comprehend the shape of the aneurysm? 4.8 (4–5) Did the simulation model help you to comprehend the location of the aneurysm? 4.8 (5–5) Did the simulation model help you to comprehend the direction of the aneurysm? 4.8 (4–5) Did the simulation model help you to comprehend the parent artery of the aneurysm? 4.7 (4–5) Did the simulation model improve your understanding on the aneurysmal therapeutic strategy of craniotomy clipping or endovascular coiling? 4.4 (4–5) Did the simulation model improve your understanding on the wide necked aneurysm and its therapeutic strategy? 4.3 (4–5) Did the simulation model improve your surgical skill? 4.1 (4–5) Did the simulation model improve your medical-patient communication skill? 4.6 (4–5) Was the simulation model training course useful to you? 4.6 (4–5)_		NA	NA
Jaimovich et al., 2016	Neurosurgeon trainees (mild/moderate neuro-endoscopy training), 29 Neurosurgeons (experienced in neuro-endoscopy), 4	Questionnaire	Rating scale: "Excellent", "Good", "Fair"	**Realism assessment** **Tumour biopsy:** Excellent: 75% Good: 25% Fair: 0% **Solid Tumour resection:** Excellent: 75% Good 25% Fair: 0%		Final skill assessment of improvement **Tumour biopsy:** Excellent: 50% Fair: 50% Fair: 0% **Solid tumour resection:** Excellent: 0% Fair: 100% Fair: 0%	NA

Table 2 (continued)

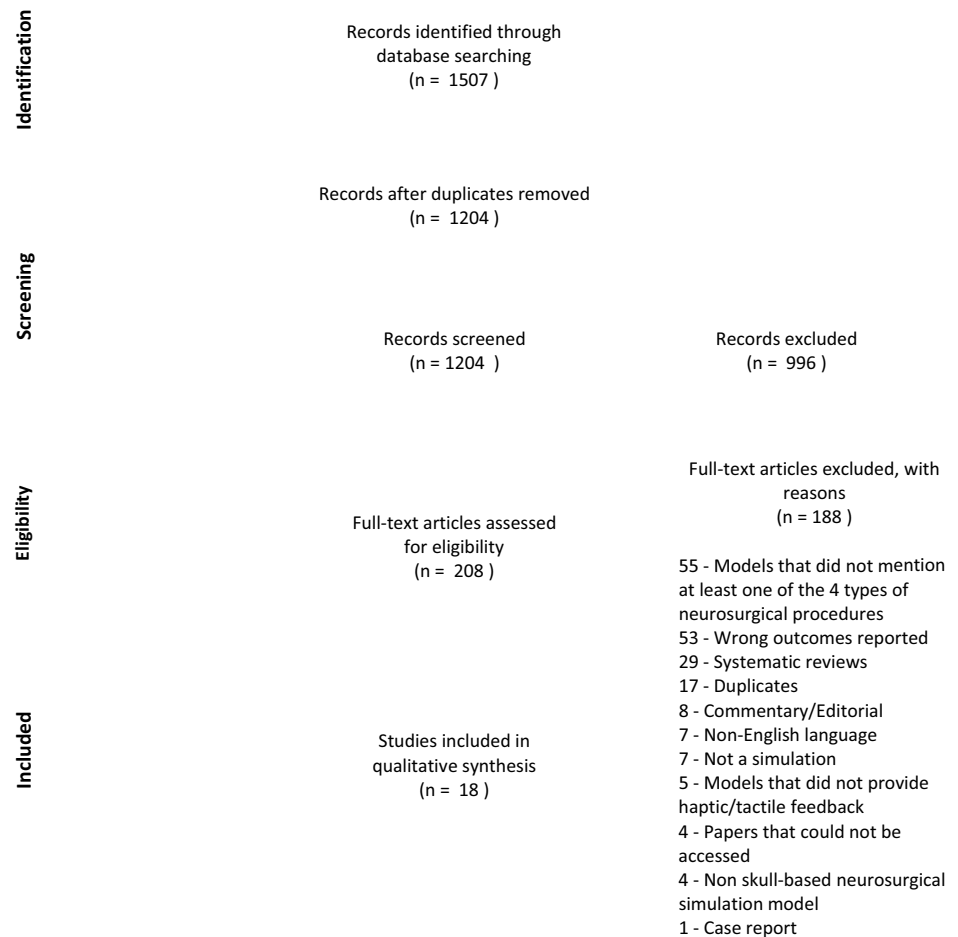| Author, year	Participant level, n	Assessment	Reported outcome	Scale		
				Face	Content	Construct
De Oliveira et al., 2018	Neurosurgery residents (PGY4), 3	Face: Questionnaire Content: Rating by neurosurgeons	5-point Likert scale (Face: 5 = exactly like, 4 = very similar, 3 = similar, 2 = little similarity, 1 = not similar) (Content: 1 = not able to do the task; 2 = poor technique (imprecise hand manoeuvres, not reaching end target); 3 = reasonable technique (imprecise hand manoeuvres proximally, with end target reached after many tries); 4 = good technique (precise hands movements with end target reached at first)	**Comparison of simulator with real surgery; mean (SD)** Aneurysm clipping: 4.0 (0.43) Aneurysm rupture management: 3.3 (0.49) Real surgery anatomical repro-duction during simulation: NA Time in mins to complete the entire simulation: 34.5 (12.3)	All assessors (neurosurgeons) evaluated the residents' performance of aneurysm clipping on placenta: 4 (0.0) Aneurysm clipping on cadaver: 4 (0.0) Aneurysm clipping on video: 1.67 (0.47)	NA

and "extremely useful" [19]. In terms of construct validity, the study demonstrated that the novice group experienced the greatest increase in confidence in performing craniotomy in comparison to the experienced group [19]. Upon assessment of content validity, the model scored highly on haptic feedback [46].

Aneurysm clipping and craniotomy

Studies that assessed both aneurysm clipping and craniotomy used VR [3] and 3D printed synthetic models [60] (Table 1). Both studies validated their models using questionnaires; Ryan et al. (2016) [60] used a Likert scale and Alaraj et al. (2015) [3] used a combination of yes/no dichotomous responses, a Likert scale, and free-text responses. Alaraj et al. (2015) [3] evaluated both face validity and content validity, whereas Ryan et al. (2016) [60] only considered face validity. Both studies found high ratings for "usefulness" and increasing understanding of the aneurysm (Table 2). Although the VR model was rated favourably for its use, it was limited by haptic realism as only 12% of participants agreed that it was realistic. Conversely, the patient-derived 3D printed aneurysm models [60] found higher ratings for realism of the artery (4.4) with lower ratings for clip application and bone drilling (4.1). This contrast between the two models is expected given the haptic realism of the 3D printed models.

Tumour resection

Tumour resection was assessed on a synthetic model [7], live rats [33], and two VR models [25, 69]. All studies used questionnaires to evaluate their models. Gelinas-Phaneuf et al. (2014) [25] and Vloeberghs et al. (2007) [69] used 5-point Likert scales. Ashour et al. (2016) [7] reported a percentage and Jaimovich et al. (2016) [33] used a rating scale of "excellent", "good", and "fair". All studies evaluated both face and content validity with Gelinas-Phaneuf et al. (2014) [25] also reporting construct validity. In Ashour's study, the synthetic model was rated > 90% for face validity and > 80% for content validity (Table 2) [7]. Similarly, realism of the live rat model was rated highly [33] (Table 2). Content valida-tion showed all participants upskilling after the use of the simulator (Table 2). Both studies using VR used the 5-point Likert scale and had favourable ratings for vis-ual realism. However, the models did not provide faith-ful sensory realism. In assessing Gelinas-Phaneuf's VR model, construct validity revealed a significant difference between medical students as compared to junior and sen-ior residents ($p < 0.05$) but no difference between junior and senior residents [25].

Fig. 2 PRISMA 2009 Diagram for included studies

PRISMA 2009 Flow Diagram

Identification

Records identified through
database searching
(n = 1507)

Screening

Records after duplicates removed
(n = 1204)

Records screened
(n = 1204)

Records excluded
(n = 996)

Eligibility

Full-text articles assessed
for eligibility
(n = 208)

Full-text articles excluded, with
reasons
(n = 188)

55 - Models that did not mention
at least one of the 4 types of
neurosurgical procedures
53 - Wrong outcomes reported
29 - Systematic reviews
17 - Duplicates
8 - Commentary/Editorial
7 - Non-English language
7 - Not a simulation
5 - Models that did not provide
haptic/tactile feedback
4 - Papers that could not be
accessed
4 - Non skull-based neurosurgical
simulation model
1 - Case report

Included

Studies included in
qualitative synthesis
(n = 18)

Vessel suturing

Aoun et al. (2015) [16] and Belykh et al. (2016) [12] assessed end-to-side micro-anastomosis techniques on synthetic and placenta models, respectively (Table 1). Both assessed face and construct validity. Participants from both studies agreed that the respective models were suitable in replicating the real surgical technique. Both studies used the Northwestern Objective Microanastomosis Assessment Tool to assess construct validity. The mean NOMAT score of the untrained group was significantly lower than that of the trained group in the studies by Aoun et al. (2015) [6] ($p = 0.02$) and Belykh et al. (2016) [12] ($p = 0.01$).

Tumour resection and aneurysm clipping

De Oliveira et al. (2019) [49] compared a placenta model with a cadaver model to evaluate vessel suturing and

aneurysm clipping skills (Table 1). The study used a questionnaire for face validity and ratings by neurosurgeons of the neurosurgery residents' performance for content validity. On assessment of face validity, the simulator scored > = 4 for microscope and microsurgical instruments handling (5), bipolar coagulation of bleeding microvessels (5), and aneurysm clipping (4). However, there were lower ratings for aneurysm rupture management (3.3) and aneurysm neck and dome dissection (3.8). Regarding content validity, there was no difference between the ratings received by residents (4) for the placenta and cadaver model (Table 2).

Discussion

Our systematic review highlighted early evidence of the feasibility and utility of using simulation models in neurosurgical training and education. Existing systematic reviews

[9, 16, 30, 44, 53, 59] have evaluated the utility and efficacy of simulation models in neurosurgical education and training. However, to our knowledge, there are no previous reviews assessing these 4 neurosurgical procedures/skills: craniotomy/burr hole drilling, aneurysm clipping, vessel suturing, and tumour resection using face, content, and construct validity. These are validation methods in medical education which provide a framework for evaluating the utility of simulation models. The MERSQI tool evidenced the quality of the studies included being moderate. The strengths of the included studies were high response rates and appropriateness of data analysis, whereas main weaknesses were failing to evaluate the simulations in more than one institution, study design, and poor validation of the qualitative and quantitative tools used for assessment. Face, content, and construct validity are discussed for all 4 simulation model types included in our study.

Human cadaver models

Both studies examining cadaver models found high ratings for face validity [1, 22]. Upon assessment of content validity, a majority of participants from Aboud et al. (2015) [1] found cadaveric intraoperative rupture to be realistic and superior to existing models for cerebral revascularisation. The cadaveric model outperformed the placental model for face validity [22] (although no differences were seen in content validity). Other studies comparing cadaveric models to physical and haptic simulators have shown they accrue the highest reported benefit for skill improvement [24]. Cadavers are known to simulate tissue dissection, bleeding, and pulsation with high fidelity [1, 54]. In line with literature, our findings demonstrate that cadavers are useful and effective for cranial procedures and manipulating soft tissues [57]. Despite their fidelity, cadaver models were the least commonly found simulation model in this review [1, 22]. This is in keeping with the decreasing prevalence of cadaver models in surgical training owing to high costs, low cadaver availability, and ethical issues [22, 49, 57].

Animal/tissue models (placenta)

Similar to cadavers, all three studies using animal or tissue models found favourable results for face validity. Placenta models [13, 22] successfully discriminated between competence levels, whilst the live rat models confirmed skill improvement after practicing with the model [22, 33]. Our results are consistent with literature findings that tissue and animal models are favoured due to neuroanatomical and neurovascular similarities [44]. Specifically, the large- and small-necked aneurysms present in placenta have been an excellent model in simulating microsurgical skills such as aneurysm clipping [41]. Furthermore, the resemblance of

brain tumours in rats to that in humans has served as a useful model for neuro-oncology procedures [38]. However, the paucity of animal models in our study reflects the downward trend of using animal models due to issues of animal rights, ethical concerns, and difficulty in procuring large numbers of animals [8, 54, 61].

Virtual reality (VR)

Six studies used VR models [3, 25, 27, 46, 69, 72]. Our study revealed favourable ratings for face validity comparing the appearance of the VR models to real surgery revealing the benefits of the model for anatomic understanding. However, participants of all six VR studies reported low haptic fidelity (Table 2). Therefore, haptic sensation was a key limitation to the success of VR models in our study, similarly found in robotic surgery where a lack of haptics limited surgical skill development [48]. Nevertheless, actuators, which apply force-feedback, are improving with rapid response times. This includes electroactive polymers, piezoelectric, electrostatic, and subsonic audio wave surface actuations that allow for improved sensory feedback [35, 73]. With the advancement of computer graphics, the visual realism of VR simulations continues to further improve and allows for realistic anatomical representations [9, 59]. Overall, VR models provide performance metrics that have been strongly correlated to skills in the OR [9, 36].

Synthetic models

Synthetic models (non-flesh) are traditionally low-fidelity simulations [9] used for fine motor specific skills such as suturing [6], tumour resection [7], aneurysm clipping [39, 42, 70], and burr holes [19, 60]. Synthetic models that simulated aneurysm clipping had high face validity and content validity in improving understanding of structure of the aneurysm in relation to the parent artery [39]. Similarly, synthetic models simulating craniotomy had high face validity (≥ 4) on a Likert scale [19] and provided most benefit to the novice group compared to experienced group [19], an objective indicator to favour the model. Ashour et al. [7] also showed significant skill improvement (94%). This is similarly reflected in literature where synthetic simulators have shown to translate skills from simulations to surgical performance [5, 29, 41]. Compared to aforementioned models, synthetic models have low overall costs [49, 71], are easily handled (avoiding the regulations around animal or cadaveric tissue) [49], and provide a safe, controlled environment. Potential limitations include the lack of repeated use of models, increased cost of synthetic models and the lack of realism compared to human cadaveric or animal models [9]. Despite this, advancements in 3D printing technology

allows for increasingly realistic and readily available simulation models. This makes training more accessible without significant financial investment.

COVID-19 and simulation models in neurosurgery

Development of novel learning modalities such as virtual reality, tele-simulations, and synthetic models have been accelerated during the COVID-19 pandemic [31]. However, many simulation models in neurosurgery are expensive and nascent in development. Nevertheless, there is a pressing need for these models due to the concerns of surgical training education which includes being exposed to fewer cases with complex operative techniques and limited patient interaction [44, 62, 66].

Based on our review and current literature findings on residents' concerns of surgical training in the pandemic, synthetic models, due to their visual and haptic realism, are the most convenient forms of simulation models to be used safely in remote settings. However, with the rapid advent of technology spurred on by the pandemic, VR models using haptic feedback technologies show promise. Despite the anatomical and haptic realness of cadaver and animal models, there is a downward trend in their use given procurement costs, storage, and ethical issues. This is reflected in our review where VR and synthetic models were most commonly used, with only two and three studies using cadaver or animal models, respectively.

These findings provide significant benefits for remote training in lower- and middle-income countries where maximising efficiency of surgical training is essential given over 95% of the population lacks access to basic surgical care [43]. These countries face barriers to incorporate simulated training including high costs, finding an appropriate training environment, and storage of cadaveric and animal material [15]. However, development of virtual learning platforms, remote tele-simulation [34], and the development of low-cost high-fidelity VR platforms [4, 50] could broaden opportunities for education. Our findings showed that synthetic and virtual platforms have been used to simulate key neurosurgical procedures, further demonstrating their applicability in low- and middle-income countries.

Strengths and limitations

Limitations of this review must be considered, many of which are related to restrictions in medical education research and the associated ethical barriers. Additionally, there is no consensus on the gold standard to assess medical and surgical education outcomes. Construct validity has

been perceived by few validity theorists to be "the whole of validity from a scientific point of view" [58]. However, only five studies assessed construct validity, whilst most studies included assessed face and content which are considered to be "subjective".

A major limitation of our paper is the lack of inclusion of predictive validity, considered an "ultimate" assessment to establish validity of an educational tool [47]. On post hoc analysis, we found that none of the included studies evaluated predictive validity. Predictive validity refers to the accuracy of prediction made by a model or test to confirm the future skill of an individual on whom the same model or test will be applied [47, 64]. However, assessing predictive validity is logistically impractical as it requires long-term follow-up. Furthermore, there is a lack of literature assessing simulation models according to the various validity types, possibly attributed to a lack of medical education tools in a surgical setting. This is also echoed by a recent systematic review [51] which calls for a prevalent use of validity scoring methodology to assess simulation tools and translation of these models in the OR. Additionally, there is a paucity of studies addressing all 3 validity scales to assess simulation models. Due to the heterogeneity in research methodology, study design, and types of outcomes reported, a meta-analysis was not conducted. This reveals the need for objective validation methodology for simulation methods.

However, the paper also had several strengths; according to our knowledge, this review is the first to address 4 major neurosurgical skills using various modalities of training, assessed by face, content, and construct validity. Existing reviews that discussed neurosurgical skills using face, content, and construct validity only cover a specific neurosurgical procedure or model type [11, 51]. Therefore, our review is the first to review these skills using an appraised method of validating simulation models. Furthermore, strong consensus in the literature shows the MERSQI tool as a robust tool of choice for critical appraisal of the medical education research included in this study [18, 63].

Future work should aim to develop a standardised approach to assessing neurosurgical simulation tools using face, content, and construct validity. Currently, there is a lack of consistent reporting of objective validation methods. A review of 83 studies on surgical simulation reported 60% targeted construct validity, 24% targeted concurrent validity, and 5% looked at predictive validity [47]. Similarly, our review found that although a majority of included studies reported subjective validation methods such as face ($n = 18$), and to a lesser extent, content validity ($n = 11$), there was a relative lack of reporting construct validity ($n = 6$). We, therefore, propose developing a simulation assessment template, which incorporates both subjective and objective validation assessments, adapted for various neurosurgical skills identified in this review. This will allow simulation

sites to collect adequate data to validate their models whilst saving them from time-intensive work to build new validity questionnaires for every simulation model.

Additionally, further work should examine how to facilitate the collection of data on predictive validity. Establishing predictive validity for simulators is essential given that the goal of simulation is the improved ability in the OR. This must therefore be assessed in studies evaluating simulation tools. However, doing this remains difficult for researchers to implement in practical terms. Possible options that may facilitate this are a digital survey form or a mobile app for surgeons or their supervisors to assess in-theatre performance that can then be compared to the simulator score.

Conclusion

This review assessed neurosurgical simulation models using face, content, and construct validity, and reported an increased use of simulation models in neurosurgical training. Whilst synthetic models are currently the most convenient and practical, especially during the COVID-19 crisis, VR models were found promising due the visual realism and improved haptic feedback technology. This also provides neurosurgical educators tools and assessment methods for simulation.

Although surgical simulation models receive generally positive feedback from trainees, comparing results among different studies were limited by the heterogeneity among studies. Moreover, studies examining simulation methods seldom use objective validation methods such as construct and predictive validity. Future work should examine how to facilitate the collection of objective validity, and aim to create a simulation assessment template that can be adapted for various neurosurgical simulation models.

Declarations

References

1. Aboud E, Aboud G, Al-Mefty O, Aboud T, Rammos S, Abolfotoh M, Hsu SPC, Koga S, Arthur A, Krisht A (2015) "Live cadavers" for training in the management of intraoperative aneurysmal rupture. J Neurosurg 123(5):1339–1346

2. Akhtar KSN, Chen A, Standfield NJ, Gupte CM (2014) The role of simulation in developing surgical skills. Curr Rev Musculoskelet Med 7(2):155–160

3. Alaraj A, Luciano CJ, Bailey DP, Elsenousi A, Roitberg BZ, Bernardo A, Banerjee PP, Charbel FT (2015) Virtual reality cerebral aneurysm clipping simulation with real-time haptic feedback. Oper Neurosurg 11(1):52–58

4. Alvarez-Lopez F, Maina MF, Saigí-Rubió F (2020) Use of a low-cost portable 3d virtual reality gesture-mediated simulator for training and learning basic psychomotor skills in minimally invasive surgery: development and content validity study. J Med Internet Res 22(7):e17491

5. Anastakis DJ, Regehr G, Reznick RK, Cusimano M, Murnaghan J, Brown M, Hutchison C (1999) Assessment of technical skills transfer from the bench training model to the human model. Am J Surg 177(2):167–170

6. Aoun SG, El Ahmadieh TY, El Tecle NE, Daou MR, Adel JG, Park CS, Batjer HH, Bendok BR (2015) A pilot study to assess the construct and face validity of the Northwestern Objective Microanastomosis Assessment Tool. J Neurosurg 123(1):103–109

7. Ashour AM, Elbabaa SK, Caputy AJ, Gragnaniello C (2016) Navigation-guided endoscopic intraventricular injectable tumor model: cadaveric tumor resection model for neurosurgical training. World Neurosurg 96:261–266

8. Aurich LA, Silva Junior LFM da, Monteiro FM do R, Ottoni AN, Jung GS, Ramina R (2014) Microsurgical training model with nonliving swine head. Alternative for neurosurgical education. Acta Cir Bras 29(6):405–409

9. Badash I, Burtt K, Solorzano CA, Carey JN (2016) Innovations in surgery simulation: a review of past, current and future techniques. Ann Transl Med 4

10. Bambakidis NC, Tomei KL (2020) Impact of COVID-19 on neurosurgery resident training and education. J Neurosurg 133(1):10–11

11. Barsom EZ, Graafland M, Schijven MP (2016) Systematic review on the effectiveness of augmented reality applications in medical training. Surg Endosc 30(10):4174–4183

12. Belykh E, Lei T, Safavi-Abbasi S, Yagmurlu K, Almefty RO, Sun H, Almefty KK, Belykh O, Byvaltsev VA, Spetzler RF (2016) Low-flow and high-flow neurosurgical bypass and anastomosis training models using human and bovine placental vessels: a histological analysis and validation study. J Neurosurg 125(4):915–928

13. Belykh E, Miller EJ, Lei T, Chapple K, Byvaltsev VA, Spetzler RF, Nakaji P, Preul MC (2017) Face, content, and construct validity of an aneurysm clipping model using human placenta. World Neurosurg 105:952-960.e2

14. Bina RW, Lemole GM, Dumont TM (2016) On resident duty hour restrictions and neurosurgical training: review of the literature. J Neurosurg 124(3):842–848

15. Campain NJ, Kailavasan M, Chalwe M, Gobeze AA, Teferi G, Lane R, Biyani CS (2018) An evaluation of the role of simulation training for teaching surgical skills in Sub-Saharan Africa. World J Surg 42(4):923–929

16. Cobb MI-PH, Taekman JM, Zomorodi AR, Gonzalez LF, Turner DA (2016) Simulation in neurosurgery—a brief review and commentary. World Neurosurg 89:583–586

17. Coelho G, Zanon N, Warf B (2014) The role of simulation in neurosurgery. Child's Nerv Syst 30(12):1997–2000

18. Cook DA, Reed DA (2015) Appraising the quality of medical education research methods: the medical education research study quality instrument and the Newcastle-Ottawa scale-education. Acad Med 90(8):1067–1076

19. Craven C, Baxter D, Cooke M, Carline L, Alberti SJMM, Beard J, Murphy M (2014) Development of a modelled anatomical replica for training young neurosurgeons. Br J Neurosurg 28(6):707–712

20. Dagi TF (2017) Seven Ethical Issues Affecting Neurosurgeons in the Context of Health Care Reform. Neurosurgery 80(4S):S83–S91

21. De Montbrun SL, MacRae H (2012) Simulation in surgical education. Clin Colon Rectal Surg 25(03):156–165

22. De Oliveira MMR, Ferrarez CE, Ramos TM, Malheiros JA, Nicolato A, Machado CJ, Ferreira MT, De Oliveira FB, De Sousa CFPM, Costa PHV (2017) Learning brain aneurysm microsurgical skills in a human placenta model: predictive validity. J Neurosurg 128(3):846–852

23. Dewan MC, Rattani A, Baticulon RE et al (2018) Operative and consultative proportions of neurosurgical disease worldwide: estimation from the surgeon perspective. J Neurosurg 130(4):1098–110

24. Gasco J, Holbrook TJ, Patel A, Smith A, Paulson D, Muns A, Desai S, Moisi M, Kuo Y-F, Macdonald B (2013) Neurosurgery simulation in residency training: feasibility, cost, and educational benefit. Neurosurgery 73(suppl_1):S39–S45

25. Gélinas-Phaneuf N, Choudhury N, Al-Habib AR, Cabral A, Nadeau E, Mora V, Pazos V, Debergue P, DiRaddo R, Del Maestro RF (2014) Assessing performance in brain tumor resection using a novel virtual reality simulator. Int J Comput Assist Radiol Surg 9(1):1–9

26. Germanò A, Raffa G, Angileri FF, Cardali SM, Tomasello F (2020) Coronavirus disease 2019 (COVID-19) and neurosurgery: literature and neurosurgical societies recommendations update. World Neurosurg 139:e812–e817

27. Gmeiner M, Dirnberger J, Fenz W, Gollwitzer M, Wurm G, Trenkler J, Gruber A (2018) Virtual cerebral aneurysm clipping with real-time haptic force feedback in neurosurgical education. World Neurosurg 112:e313–e323

28. Gnanakumar S, Kostusiak M, Budohoski KP, Barone D, Pizzuti V, Kirollos R, Santarius T, Trivedi R (2018) Effectiveness of cadaveric simulation in neurosurgical training: a review of the literature. World Neurosurg 118:88–96

29. Goova MT, Hollett LA, Tesfay ST, Gala RB, Puzziferri N, Kehdy FJ, Scott DJ (2008) Implementation, construct validity, and benefit of a proficiency-based knot-tying and suturing curriculum. J Surg Educ 65(4):309–315

30. Harrop J, Lobel DA, Bendok B, Sharan A, Rezai AR (2013) Developing a neurosurgical simulation-based educational curriculum: an overview. Neurosurgery 73(suppl_1):S25–S29

31. Hau H-M, Weitz J, Bork U (2020) Impact of the COVID-19 pandemic on student and resident teaching and training in surgical oncology. J Clin Med 9(11):3431

32. Innovation VH (2016) Covidence systematic review software

33. Jaimovich SG, Bailez M, Asprea M, Jaimovich R (2016) Neurosurgical training with simulators: a novel neuroendoscopy model. Child's Nerv Syst 32(2):345–349

34. Joos E, Zivkovic I, Shariff F (2021) Virtual learning in global surgery: current strategies and adaptation for the COVID-19 pandemic. Int J Surg Glob Heal 4(1):e42–e42

35. Kapoor S, Arora P, Kapoor V, Jayachandran M, Tiwari M (2014) Haptics-Touchfeedback technology widening the horizon of medicine. J Clin Diagn Res JCDR 8(3):294

36. Kundhal PS, Grantcharov TP (2009) Psychomotor performance measured in a virtual environment correlates with technical skills in the operating room. Surg Endosc 23(3):645

37. Lateef F (2010) Simulation-based learning: Just like the real thing. J Emerg Trauma Shock 3(4):348

38. Lee RMKW (1995) Morphology of cerebral arteries. Pharmacol Ther 66(1):149–173

39. Liu Y, Gao Q, Du S, Chen Z, Fu J, Chen B, Liu Z, He Y (2017) Fabrication of cerebral aneurysm simulator with a desktop 3D printer. Sci Rep 7:44301

40. M Bridges M, Diamond DL (1999) The financial impact of teaching surgical residents in the operating room. Am J Surg 177(1):28–32

41. Magaldi MO, Nicolato A, Godinho JV et al (2014) Human placenta aneurysm model for training neurosurgeons in vascular microsurgery. Clin Neurosurg 10(4):592–601

42. Mashiko T, Otani K, Kawano R, Konno T, Kaneko N, Ito Y, Watanabe E (2015) Development of three-dimensional hollow elastic model for cerebral aneurysm clipping simulation enabling rapid and low cost prototyping. World Neurosurg 83(3):351–361

43. Meara JG, Leather AJM, Hagander L et al (2015) Global Surgery 2030: Evidence and solutions for achieving health, welfare, and economic development. Lancet 386(9993):569–624

44. Morosanu CO, Nicolae L, Moldovan R, Farcasanu AS, Filip GA, Florian IS (2019) Neurosurgical cadaveric and in vivo large animal training models for cranial and spinal approaches and techniques—a systematic review of the current literature. Neurol Neurochir Pol 53(1):8–17

45. Moscote-Salazar LR, Agrawal A, Garcia-Ballestas E, Bohorquez-Rivero J, Rahman MM (2020) Letter to the Editor: "Human patient simulation: educational issues and practical implications in COVID-19 times." World Neurosurg 141:570

46. Müns A, Mühl C, Haase R, Möckel H, Chalopin C, Meixensberger J, Lindner D (2014) A neurosurgical phantom-based training system with ultrasound simulation. Acta Neurochir (Wien) 156(6):1237–1243

47. Nortwick SS Van, Lendvay TS, Jensen AR, et al (2010) Methodologies for establishing validity in surgical simulation studies. Surgery 147(5):622–630

48. Okamura AM (2009) Haptic feedback in robot-assisted minimally invasive surgery. Curr Opin Urol 19(1):102

49. Oliveira LM, Figueiredo EG (2019) Simulation training methods in neurological surgery. Asian J Neurosurg 14(2):364

50. Parham G, Bing EG, Cuevas A, Fisher B, Skinner J, Mwanahamuntu M, Sullivan R (2019) Creating a low-cost virtual reality surgical simulation to increase surgical oncology capacity and capability. Ecancermedicalscience. https://doi.org/10.3332/ecancer.2019.910

51. Patel EA, Aydin A, Cearns M, Dasgupta P, Ahmed K (2020) A systematic review of simulation-based training in neurosurgery, part 1: cranial neurosurgery. World Neurosurg 133:e850–e873

52. R Babu R, Thomas S, Hazzard M et al (2014) Morbidity, mortality, and health care costs for patients undergoing spine surgery following the ACGME resident duty-hour reform: Clinical article. J Neurosurg Spine 21(4):502–515

53. Reed DA, Cook DA, Beckman TJ, Levine RB, Kern DE, Wright SM (2007) Association between funding and quality of published medical education research. JAMA 298(9):1002–1009

54. Reznick R, Regehr G, MacRae H, Martin J, McCulloch W (1997) Testing technical skill via an innovative "bench station" examination. Am J Surg 173(3):226–230

55. Riella MC, Roy-Chaudhury P (2013) Vascular access in haemodialysis: strengthening the Achilles' heel. Nat Rev Nephrol 9(6):348–357

56. Robertson FC, Lippa L, Broekman MLD (2020) Task shifting and task sharing for neurosurgeons amidst the Covid-19 pandemic. J Neurosurg 133(1):5–7

57. Robison RA, Liu CY, Apuzzo MLJ (2011) Man, mind, and machine: the past and future of virtual reality simulation in neurologic surgery. World Neurosurg 76(5):419–430

58. Royal KD (2017) Four tenets of modern validity theory for medical education assessment and evaluation. Adv Med Educ Pract 8:567

59. Ruthenbeck GS, Reynolds KJ (2015) Virtual reality for medical training: the state-of-the-art. J Simul 9(1):16–26

60. Ryan JR, Almefty KK, Nakaji P, Frakes DH (2016) Cerebral aneurysm clipping surgery simulation using patient-specific 3D printing and silicone casting. World Neurosurg 88:175–181

61. Schöffl H, Froschauer SM, Dunst KM, Hager D, Kwasny O, Huemer GM (2008) Strategies for the reduction of live animal use in microsurgical training and education. Altern Lab Anim 36(2):153–160

62. Scullen T, Mathkour M, Maulucci CM, Dumont AS, Bui CJ, Keen JR (2020) Letter to the editor impact of the COVID-19 pandemic on neurosurgical residency training in New Orleans. World Neurosurg 139:718–719

63. Smith RP, Learman LA (2017) A plea for MERSQI: the medical education research study quality instrument. Obstet Gynecol 130(4):686–690

64. Sugiono M, Teber D, Anghel G, Gözen AS, Stock C, Hruza M, Frede T, Klein J, Rassweiler JJ (2007) Assessing the predictive validity and efficacy of a multimodal training programme for laparoscopic radical prostatectomy (LRP). Eur Urol 51(5):1332–1340

65. Sutherland L, Middleton P, Anthony A, Hamdorf J, Cregan P, Scott D, Maddern G (2006) Surgical simulation: a systematic review. Ann Surg 243(3):291–300

66. Tomlinson SB, Hendricks BK, Cohen-Gadol AA (2020) Innovations in neurosurgical education during the COVID-19 pandemic: is it time to reexamine our neurosurgical training models? J Neurosurg 133(1):14–15

67. Tsermoulas G, Zisakis A, Flint G, Belli A (2020) Challenges to Neurosurgery During the Coronavirus Disease 2019 (COVID-19) Pandemic. World Neurosurg 139:519

68. Varoquier M, Hoffmann CP, Perrenot C, Tran N, Parietti-Winkler C (2017) Construct, face, and content validation on Voxel-Man® Simulator for otologic surgical training. Int J Otolaryngol 2017

69. Vloeberghs M, Glover A, Benford S, Jones A, Wang P, Becker A (2007) Virtual neurosurgery, training for the future. Br J Neurosurg 21(3):262–267

70. Wang J-L, Yuan Z-G, Qian G-L, Bao W-Q, Jin G-L (2018) 3D printing of intracranial aneurysm based on intracranial digital subtraction angiography and its clinical application. Medicine (Baltimore) 97

71. Werz SM, Zeichner SJ, Berg B, Zeilhofer H, Thieringer F (2018) 3D printed surgical simulation models as educational tool by maxillofacial surgeons. Eur J Dent Educ 22(3):e500–e505

72. Wong D, Unger B, Kraut J, Pisa J, Rhodes C, Hochman JB (2014) Comparison of cadaveric and isomorphic virtual haptic simulation in temporal bone training. J Otolaryngol Neck Surg 43(1):31

73. Yu X, Xie Z, Yu Y, Lee J, Vazquez-Guardado A, Luan H, Ruban J, Ning X, Akhtar A, Li D (2019) Skin-integrated wireless haptic interfaces for virtual and augmented reality. Nature 575(7783):473–479

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Comments

The authors should be commended for their effort and for drawing attention to the face, content and construct validity elements of simulation studies in neurosurgery. As in most surgical specialties, hundreds of various training tricks and techniques are published, ranging from simple exercises to highly complex simulators. Simply publishing a concept paper on a simulation idea is no longer sufficient. In the absence of validation studies, the question must always be: Why is this particular simulator/training technique necessary? Also, without validation studies, most of these simulators are no more than very expensive toys that give a training programme a "high-tech" edge.

This review further exposes the most important and likely the most painful aspect of validation studies: none of them assesses predictive validity. Predictive validity refers to the ability of a simulation to lead to improved skill intra-operatively and ultimately lead to better patient outcomes. While it is evident that training alone can improve skill, a scientific manuscript claiming the validity of a certain technique claims improved skill and patient outcomes *through training with that particular technique/simulator*. The question is whether this effect relates to the simulator or technique itself or simply to the fact that one trains more. Until better studies appear, one must remain skeptical of different kinds of simulators that have not been validated and we must not lose sight of the fact that predictive validity in neurosurgical simulators is abysmal.

Victor Volovici.
Rotterdam, The Netherlands.