Best Practices Using Ex Vivo Animal Brain Models in Neurosurgical Education to Assess Surgical Expertise

Ahmad Alsayegh[1,2], Mohamad Bakhaidar[1,2], Alexander Winkler-Schwartz[1], Recai Yilmaz[1], Rolando F. Del Maestro[1]

■ BACKGROUND: **Ex vivo animal brain simulation models are being increasingly used for neurosurgical training because these models can replicate human brain conditions. The goal of the present report is to provide the neurosurgical community interested in using ex vivo animal brain simulation models with guidelines for comprehensively and rigorously conducting, documenting, and assessing this type of research.**

■ METHODS: **In consultation with an interdisciplinary group of physicians and researchers involved in ex vivo models and a review of the literature on the best practices guidelines for simulation research, we developed the "ex vivo brain model to assess surgical expertise" (EVBMASE) checklist. The EVBMASE checklist provides a comprehensive quantitative framework for analyzing and reporting studies involving these models. We applied The EVBMASE checklist to the studies reported of ex vivo animal brain models to document how current ex vivo brain simulation models are used to train surgical expertise.**

■ RESULTS: **The EVBMASE checklist includes defined subsections and a total score of 20, which can help investigators improve studies and provide readers with techniques to better assess the quality and any deficiencies of the research. We classified 18 published ex vivo brain models into modified (group 1) and nonmodified (group 2) models. The mean total EVBMASE score was 11 (55%) for group 1 and 4.8 (24.2%) for group 2, a statistically significant difference ($P = 0.006$) mainly attributed to differences in the simulation study design section ($P = 0.003$).**

■ CONCLUSIONS: **The present findings should help contribute to more rigorous application, documentation, and assessment of ex vivo brain simulation research.**

INTRODUCTION

Surgical simulation is an educational technique that places learners in an interactive physical or virtual learning environment that recreates or replicates different scenarios to teach defined techniques or skills. This immersive learning method has the potential to completely envelop an individual in an active and dynamic learning process. Surgical simulation has several advantages, including the following: 1) learners can be placed in comprehensive real-world clinical scenarios without risk to patient safety; 2) it allows self-directed learning, including unlimited rehearsal and opportunities to fail; 3) it can provide quantitative feedback on surgical performance not available in the operating room environment; and 4) it decreases the need for the presence of instructors. The time-focused apprenticeship paradigm of neurosurgical bimanual psychomotor training is evolving to one based on quantifiable competency.[1] Technical skills performance competency in surgery is achieved when the learner, using the appropriate surgical techniques, can safely and efficiently perform a variety of procedures common to that specialty.[2]

Various neurosurgical training models have been created to develop microsurgical skills and improve control of microsurgical instruments. These models can be synthetic, virtual, or biological in nature, the latter is obtained from human and animal cadavers.[3-5] Although virtual reality simulation is continuously

Key words
- Animal
- Best practices
- Brain
- Ex vivo
- Neurosurgery
- Simulation

Abbreviations and Acronyms
EEG: Electroencephalography
EVBMASE: Ex vivo brain model to assess surgical expertise
MRI: Magnetic resonance imaging

From the [1]Neurosurgical Simulation and Artificial Intelligence Learning Centre, Department of Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada; and [2]Division of Neurosurgery, Department of Surgery, Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia

To whom correspondence should be addressed: Ahmad Alsayegh, M.B.B.S. [E-mail: aalsayegh@kau.edu.sa]

Citation: World Neurosurg. (2021) 155:e369-e381.
https://doi.org/10.1016/j.wneu.2021.08.061

Journal homepage: www.journals.elsevier.com/world-neurosurgery

Available online: www.sciencedirect.com

evolving and can be used to train surgical techniques, organic materials provide high tissue fidelity to enhance technical skills instruction. A common microsurgical training method uses in vivo rat vessels. Examples of this model include dissection and anastomosis of cervical and femoral vessels,[6] microdissection of abdominal arteriovenous structures,[7] clipping of microsurgically induced aneurysms,[8] side-to-side anastomosis between the internal and external carotid arteries,[9] and between the femoral artery and vein.[10] The advantage of these models is the similarity of rat vessels to human cerebral arteries in texture, pulsation, and coagulability.[11] The disadvantages include the ethical considerations when using live animals for experimentation, limited animal housing capabilities, and high cost.[11] Other models have used tissues harvested from animal cadavers. The vessels from chicken wings have been used to train microvascular anastomotic techniques and aneurysm clipping.[12] The suitability of biological tissues acquired from other animal cadavers for microvascular anastomosis was also demonstrated using turkey carotid arteries,[4] porcine coronary arteries,[13] and bovine hearts.[14] Brain tumor simulation scenarios have been reported less often. Oliveira et al.[15] proposed a human placenta brain tumor model for neurosurgical training.

Human cadavers have been used for centuries to teach anatomy and surgical techniques.[16] These platforms allow learners the ability to navigate through the various bodily layers and structures in 3-dimensional space, especially when the simulation scenario involves complex anatomy such as temporal bone surgery.[17] The major disadvantages of using human cadavers include their high cost, ethical controversies, low availability, and poor tissue compliance.[17] Ex vivo animal brain simulations have been increasingly used for neurosurgical training. As these models are not an exact anatomical replica of the human brain, they should not be used to study detailed human cerebral anatomy. However, they can be used to mimic human brain conditions such as cerebral tumors and are useful for developing microsurgical techniques used in operative procedures performed commonly by neurosurgeons. Although challenging, the goal of these simulation scenarios and models is to improve neurosurgical trainee psychomotor bimanual skill performance and to assess whether the training enhances resident skills in the operating room.

A need exists to develop ex vivo models that can accurately quantitate technical psychomotor skills and operative results in realistic simulated operative settings. These models must possess both visual and tactile reality coupled with advanced quantitation. Although different exvivo animal platforms are available and various methods are available to assess these frameworks, the documentation and, therefore, the reproducibility of these platforms has been inconsistent. Thus, a need exists to develop standardized best practices criteria for objective reporting of surgical simulation studies.

In the present report, we aimed to provide the neurosurgical investigative community interested in using ex vivo animal brain simulation models with guidelines for comprehensively and rigorously conducting and documenting this type of research. First, in consultation with an interdisciplinary group of investigators and physicians involved in ex vivo models and a review of the literature on the best practices guidelines for simulation research, we developed the ex vivo brain model to assess surgical expertise (EVBMASE) checklist. Second, we applied the EVBMASE checklist to reported studies to score and compare between ex vivo animal brain simulation models used in surgical expertise training. The framework complements the existing guidelines for best practices in reporting experimental designs in medical education.[18-21] To the best of our knowledge, ours is the first attempt to create a conceptual structure to ensure the quality of ex vivo animal brain simulation studies that assess surgical skills. The present report will help to contribute to a more rigorous application and thorough documentation of this research field.

METHODS

EVBMASE Checklist

In consultation with an interdisciplinary group of physicians and researchers involved in ex vivo models and a review of the literature on best practices guidelines for simulation research,[19-21] we developed the EVBMASE checklist composed of 20 key elements for reporting studies of ex vivo brain simulation models to assess technical skills. The definitions of the relevant terms outlined in the present section are listed in **Table 1**.[22] The key elements of this checklist were divided into 3 sections: simulation study design, data structure, and discussion quality (**Table 2**).

Simulation Study Design

The simulation study design section contains 10 elements: literature review, comprehensive research platform, operative realism, ex vivo animal simulation model description, simulation properties of the model described, validation of the model, simulated tasks description, purpose of assessment tools used, expertise definition, and definition of level of expertise in each group.

Literature Review. A current and relevant literature review on the previous use of similar ex vivo animal brain simulation models should be available to the reader. An attempt should be made to situate the current study in the context of any previous reports.

Comprehensive Research Platform. The neuro-oncological patient clinical model consists of 3 distinct phases. First, the preoperative assessment phase should include the evaluation of the various locations and intrinsic properties attributed to the tumor using preoperative computed tomography and magnetic resonance imaging (MRI). The findings will result in subsequent surgical planning. The second phase is the operative phase in which the actual intervention occurs. Finally, the third phase is the postoperative assessment phase in which the outcomes of the operative procedure are determined using postoperative computed tomography and MRI. An attempt should be made to replicate all 3 components of this arc in the ex vivo simulation with the goal of obtaining quantifiable data from each step (**Figure 1**). This will enhance the ability of the model, not only to provide training opportunities for formative and summative assessments of the learners, but also to provide new insights into the composites of surgical expertise.

Operative Realism. One step in developing a comprehensive ex vivo research platform is to place the learner in realistic operating room environments involving the use of appropriate surgical

Table 1. Definitions in the Context of Ex Vivo Animal Model Simulation

Keyword	Definition
Simulation	A technique to replace or amplify real experiences with guided experiences, often immersive in nature, that evoke or replicate substantial aspects of the real world task in a fully interactive manner
Ex vivo	Experimentation performed on tissue derived from an organism in an external environment with minimal alteration of its natural conditions
Metric	A standard set of measurements by which a plan or process can be assessed and that quantifies these elements of performance
Validity	The extent to which a test measures what it claims to measure or the extent that a platform simulates a real world task
Face validity	The extent to which a test is subjectively viewed to mimic that what it intends to measure or the extent of a simulator's realism and appropriateness when compared with the actual task
Content validity	The extent to which items on a test such as skills are representative to the entire domain the test aims to measure or the extent to which a simulator's content is representative of the knowledge or skills that must be learned in the real environment
Construct validity	The extent to which a testing instrument identifies the quality, ability, or trait it was designed to measure or the ability of the simulator to differentiate experienced from inexperienced surgeons
Concurrent validity	The extent to which the test scores correspond to the scores on the benchmark test that measures the same construct
Predictive validity	The extent to which the scores on a test are predictive of the actual performance
Sensitivity	$\frac{\sum \text{True Positive}}{\sum \text{True Positive} + \sum \text{False Negative}}$; a measure of how many positive condition predictions are actually true-positive results
Specificity	$\frac{\sum \text{True Negative}}{\sum \text{True Negative} + \sum \text{False Positive}}$; a measure of how many negative condition predictions are actually true-negative results

Some definitions were modified from Gallagher and O'Sullivan.[22]

instruments and an operating microscope. An ex vivo brain simulation encased in a realistic skull replica that has been draped appropriately will help immerse the student in a dynamic learning process.

Ex Vivo Animal Simulation Model Description. Several ex vivo animal simulation models using bovine, ovine, and porcine platforms are available. Significant differences exist in the brain size of the 3 common ex vivo brains used, with the average bovine brain weight double that of ovine and porcine brains yet less than one half the weight of the adult human brain.[23] These differences in brain size can influence the model's utility and the ability to obtain face and content validity. The reason that one model over another is used should be documented. Where available, previous reports outlining these items can be cited instead.

Simulation Properties of the Model Described. The simulation properties of the ex vivo animal simulation model used should be carefully outlined. If a specific human pathology is being simulated, whether attempts had been made to replicate the human condition should be documented. If a particular human tumor or tumor site is being simulated, whether efforts had been made to best approximate the color, tactile properties, and location of that type of human tumor should be determined.

Model Validation. The educational utility of any simulation platform is enhanced by the subjective and objective validity assessments of the platform. Questionnaires are used to assess subjective validity, which includes face and content validity. Subjective validity assessments are assessed using 5- or 7-point Likert scales. However, the number that constitutes sufficient validity when using these scales has not been defined. It might be reasonable to consider standardizing the 7-point Likert scale in the use of ex vivo simulations to improve documentation and reproducibility. Thus, sufficient validity for the overall procedure and specific tasks would be deemed valid if a median score of $\geq$4.0 on a 7-point Likert scale were achieved.[24] Objective validity includes construct, concurrent, and predictive validity. Construct validity is assessed by comparing the surgical performance between the "expert" and "less skilled" groups. Concurrent validity assesses the degree that the simulation model results correlate with the reference standard for that procedure. Predictive validity assesses the question of whether the simulation predicts the future performance for an equivalent human operation. Objective validity can be assessed using a priori metrics established independently for each step and using statistical methods based on these data.

Simulated Tasks Description. A variety of simulated scenarios are available on any ex vivo simulation platform. Therefore, an adequate description of the surgical tasks to be performed should be provided.

Purpose of Assessment Tools Utilized. A variety of assessment tools are available to assess any metric the investigator might want to measure using an ex vivo simulation. The assessment tools should

Table 2. Ex Vivo Brain Model to Assess Surgical Expertise (EVBMASE) Checklist*

Component No.	Description
Simulation study design (10 points possible)	
1	Is relevant literature on the use of ex vivo animal brain simulation models provided?
2	Is the simulation comprehensive?
3	Does the simulation occur in a realistic operative environment?
4	Is the ex vivo brain simulation model described?
5	Are the simulation properties of the model adequately described?
6	Are efforts made to validate the ex vivo brain simulation model?
7	Are the simulation tasks to be performed outlined?
8	Are the purposes of assessment tools to be use outlined?
9	Is a definition of expertise in the simulation provided?
10	Is a definition of each group of expertise provided?
Data structure (5 points possible)	
11	Are the metrics used as assessment tools defined?
12	Are the purposes of each metric outlined (i.e., are these to assess performance, efficiency and/or progress)?
13	Is the sample size clearly stated (including the number of groups and the number of participants in each group)?
14	Are the correct statistical methods used?
15	Is the sensitivity and specificity mentioned?
Discussion quality (5 points possible)	
16	Is the educational application of simulation model in the context of surgical simulation outlined?
17	Are efforts made to explain the educational rationale of the simulation model used?
18	Is the use of this simulation model discussed as a formative and summative assessment tool?
19	Are methodological limitations discussed, including those pertaining to any of the above points?
20	Are the future directions discussed?

*One point given for each positive answer.

be outlined such that other investigators can reproduce these tools in their studies.

Expertise Definition. When using ex vivo brain simulation to assess technical skills, determining what constitutes "expert" performance is critical to interpreting the results of the study. If a neurosurgical participant is considered an "expert," their patient practice should either focus on the area that the simulation measures or data should be provided that indicates that this "expert" is very knowledgeable concerning the procedure and performs the operation frequently.

Definition of Level of Expertise in Each Group. A clear definition of the abilities of each group is critically important. Specifically, the definitions should include which participants constitute the "expert" group as described in the previous section and, in particular, the reasons the specific individuals were chosen for the "less skilled" group. Care should be taken to identify the participant learner group according to the year of neurosurgical training (i.e., junior, senior resident, fellow) or skill level using other defined criteria.

Data Structure

The data structure section contains 5 elements: metrics used, purpose of the metrics, sample size, statistical methods, and sensitivity and specificity.

Metrics Used. The metrics used as assessment tools should be clearly outlined. The origin of these metrics should be described, including whether these metrics were the result of a consensus of "experts" derived from the literature or from virtual reality and artificial intelligence technology.

Purpose of the Metrics. The metrics can be used to assess performance expertise, efficiency, and progress of the technical psychomotor task being measured during the simulation. The purpose for the choice of a particular metric should be outlined so that the reader understands why that metric and not another was used to assess that particular skill. The time required to complete a procedure on a simulation platform is frequently used to assess expertise. However, studies of virtual reality neurosurgical performance have highlighted that metrics of safety and efficiency are essential to expertise performance.[25] More focus on the technical

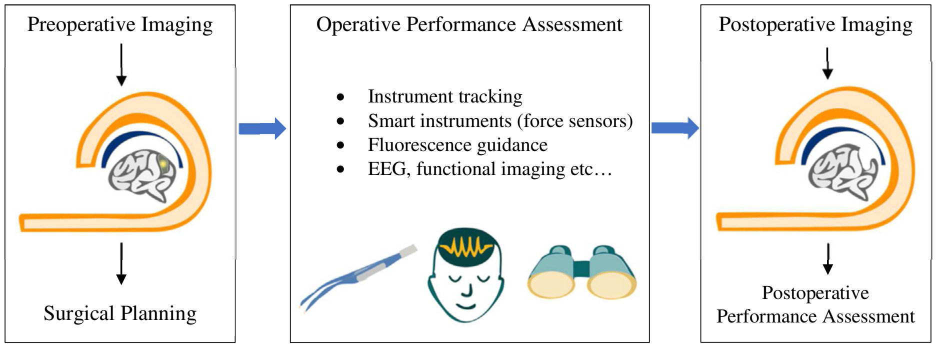

Figure 1. Creation of a comprehensive research platform for ex vivo brain simulation. EEG, electroencephalography.

skills needed to perform the procedure safely rather than the procedure length seems prudent in simulations.

Sample Size. The number of groups and participants per group should be clearly stated. In ex vivo brain simulation studies, it is often easier to recruit "less skilled" or "novices" (medical students) than to recruit "experts" (practicing neurosurgeons with a defined level of expertise). Finally, the number in each group should be sufficient to determine whether the results are significant.

Statistical Methods. Several statistical tests are available to assess metrics data. Care should be taken to explain why a particular statistical method was chosen and which values would be considered significant. Certain statistical tests and artificial intelligent algorithms will function poorly if the input data are few. Thus, the sample size must be appropriate for the statistical test or algorithm used.

Sensitivity and Specificity. The simulation literature differs on the reporting of test success.[18,19] The engineering community reports test success in terms of accuracy and equal error rates. However, these might be less intuitive to medical readers, who might be more familiar with the use of sensitivity and specificity. It is important to discuss sensitivity and specificity when reporting studies.

Discussion Quality

The discussion quality section contains 5 elements: educational application of ex vivo brain simulation, educational rationale of simulation model, the use of the simulation model as an assessment tool, methodological limitations, and future directions.

Educational Application of Ex Vivo Brain Simulation. The investigators should clearly state the educational aim for the use of their model. Some simulation models, such as those involving brain tumor resection, have the capability of generating large datasets capable of use by artificial intelligence machine learning classifier algorithms, artificial neural networks, and deep learning methods.[19,25-27] This ability to classify participants allows for the use of these technologies in summative assessments. The potential of these systems to be used in formative assessments is increasing with the application of intelligent tutoring systems.[26,27] Ex vivo simulation models that do not generate large datasets will have limited utility in artificial intelligence technologies. Summative assessments can have a significant impact on learner success, hence, they require high accuracy and reproducibility. Formative assessments require the learner to understand the specific metrics that must be mastered in order to improve technical skills.

Educational Rationale of the Simulation Model in the Context of Surgical Simulation. The investigators should outline why specific metrics were chosen and how they relate to the educational goals of the simulation. The questions include whether the metrics used demonstrate construct validity; and whether the metrics assessed can be used to classify the level of participant's performance. The use of metrics that do not demonstrate construct validity will lessen the ability of other investigators to reproduce the results.

Use of the Ex Vivo Simulation as an Assessment Tool. A clear description should be provided by the investigators of why the ex vivo simulation model is important in an educational context. Some metrics of technical skills performance such as instrument movement could be teachable and, thus, useful in both formative and summative evaluation. However, other metrics (e.g., electroencephalography [EEG] and eye movement) can be difficult to teach and, thus, limited in these evaluations.

Methodological Limitations. The limitations of the study should be addressed. Specifically, the shortcomings of the use of the ex vivo simulation in surgical skill assessment should be outlined.

Future Directions. The future directions should be outlined. This information will help the reader and the medical education community by providing a better understanding of how the field of ex vivo simulation in the assessment of surgical expertise might continue to evolve.

Literature Review

We structured our review in accordance with the PRISMA (preferred reporting items for systematic reviews and meta-analyses) guidelines, and the results were imported to EndNote X9 (The EndNote Team, Philadelphia, Pa; Clarivate Analytics, London, United Kingdom; 2013).[28] We performed a literature search using the PubMed Medline database to analyze ex vivo brain models used for neurosurgical simulation with an appointed duration from January 2000 to January 2021. The search terms used included a combination of "neurosurgery" and "training" and "simulation" and "animal" and "brain." Only English-language reports were included in our review. The eligible studies included ex vivo brain models derived from animal

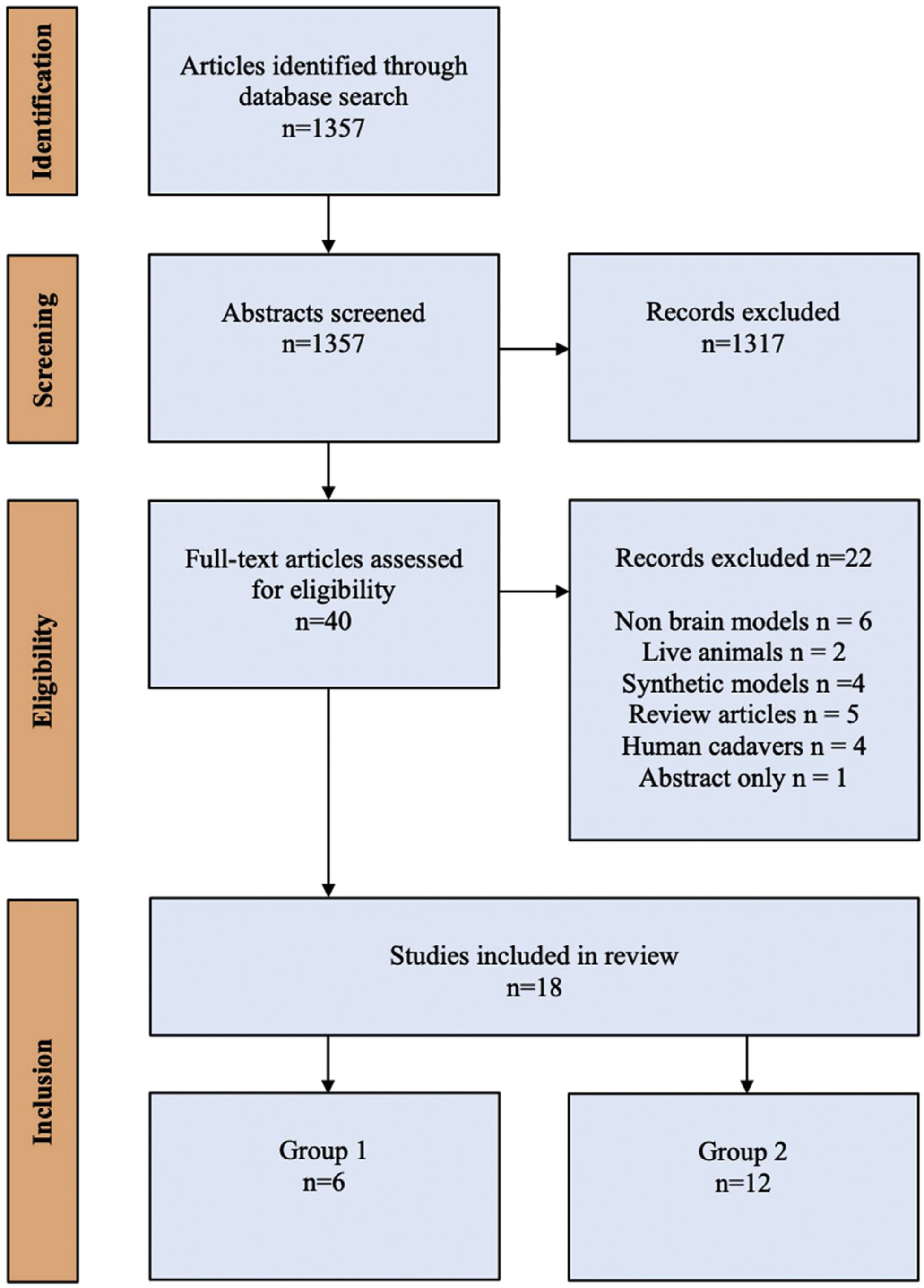

Figure 2. PRISMA (preferred reporting items for systematic reviews and meta-analyses) flowchart for study selection.

cadavers. Live animals, non−brain models, human cadavers, synthetic models, virtual simulators, publication of an abstract only, or models that were used for purposes other than neurosurgical training were excluded.

Two of us (A.A., M.B.) individually reviewed and scored each report using the EVBMASE checklist. One point was awarded for each element of the checklist for every report outlined in our review. If the scores were not consistent, the 2 reviewers attempted to reach a consensus. If no consensus was achieved for a specific score, the remaining author's input was used to provide a final score.

Statistical Analysis

The R software, version 4.0.2 (The R Foundation for Statistical Computing, Vienna, Austria; available at: http://www.r-project.org/), was used for descriptive and inferential statistical analysis. The Welch t test was used for intergroup comparisons. The interrater agreement was evaluated using Cohen's kappa. The level of statistical significance was determined at $P < 0.05$.

RESULTS

The initial search resulted in 1357 reports. The abstracts of these reports were screened, and 40 fit the criteria previously outlined. After assessing the full text and applying the review criteria, 18 reports were included in our review (Figure 2). The reasons for excluding the reports were the use of tissues other than brain or brain coverings ($n = 6$), synthetic models ($n = 4$), human cadavers ($n = 4$), live animals ($n = 2$), review articles ($n = 5$), and only having a published abstract ($n = 1$). When only 1 ex vivo brain model was used, bovine models were used in 8 of 17 studies (47%), ovine models in 5 of 17 (29%), and porcine models in 4 of 17 studies (24%). In 1 study, both ovine and porcine models were used (6%).

Classification

The ex vivo animal brain models identified were classified into 2 groups. In group 1, ex vivo animal brain simulation models were modified to mimic a human brain pathology with the goal of producing a re-creation of human disease on which the learner surgical technical skills could be assessed (Table 3).[29-34] In group 2, ex vivo animal brain simulation models were used without modification for microsurgical dissection training (Table 4),[23,35-41] and studies that provided learners with an anatomical replica of the human brain and coverings were used to assess the technical skills involved in defined human neurosurgical procedures (Table 5).[42-45]

Group 1 included 6 reports (33%). In all these investigations, the human pathologies re-created were a wide variety of human intra- and extra-axial tumors.[29-34] Several advantages exist for using animal models that replicate human tumors. These include the ability to create comprehensive research platforms for the assessment of surgical techniques and operative outcomes.[33] These platforms will further enhance learner face and content validity if these procedures are performed on ex vivo brains encased in human skull replicas in realistic operating room environments involving tumors with human characteristics. Although multiple different types of materials have been used to develop these artificial tumors, recent studies have focused on creating tumors with human tumor stiffness and color properties to create a more realistic visual and tactile experience for the learner.[31,33,34,46] The stiffness of alginate hydrogel-based tumors can be modified using different concentrations of calcium sulfate, a cross-linking agent, until similar biomechanical properties to human primary tumors are achieved.[33,34] Other investigators have used haptic tuning, which involves having expert neurosurgeons to use their haptic expertise to adjust the stiffness of artificial tumors to approximate those they experience when operating on human tumors.[31,46] A number of technologies that can quantitate tumor removal have been used. The presence of gadolinium in the artificial tumor allows for the use of pre- and postoperative 7T MRI investigations, not only to quantitate the grams of tumor resected, but also the grams of normal gray and white matter removed.[34] These artificial tumors can also incorporate a variety of fluorescein solutions, allowing

Table 3. Modified Ex Vivo Animal Brain Models

Investigator	Animal	Task	Artificial Tumor
Kamp et al.,[29] 2015	Ovine	Solid tumor resection; infiltrative tumor resection	Agar-agar and ink
Altun et al.,[30] 2019	Bovine	Intra-axial tumor resection; extra-axial tumor resection; cerebellopontine angle tumor resection; fourth ventricle tumor resection	Polyurethane foam
Valli et al.,[31] 2019	Ovine and porcine	Brain tumor resection	Agar, gelatin, and 5 various fluorescent dyes
Grosch et al.,[32] 2020	Bovine	Solid tumor resection; infiltrative tumor resection; stereotactic navigated biopsy; cerebellopontine tumor resection	Aspic powder and fluorescein
Winkler-Schwartz et al.,[33] 2020	Bovine	Brain tumor resection	Alginate, calcium sulfate, gadolinium, and fluorescein
Tran et al.,[34] 2021	Bovine	Brain tumor resection	Alginate, calcium sulfate, gadolinium, and fluorescein

Table 4. Nonmodified Ex Vivo Animal Brain Models

Investigator	Animal	Task
Hicdonmez et al.,[23] 2006	Bovine	Optic nerve and circle of Willis dissection; aneurysm clip application
Hicdonmez et al.,[35] 2006	Bovine	Interhemispheric transcallosal approach
Hamamcioglu et al.,[36] 2008	Ovine	Cranial nerve dissection; microvascular decompression of nerve V
Turan Suslu et al.,[37] 2013	Bovine	Cranial nerve dissection
Altunrende et al.,[38] 2014	Ovine	Intraorbital dissection; optic nerve dissection; optic canal exposure
Aurich et al.,[39] 2014	Porcine	Interhemispheric transcallosal approach; Sylvian fissure dissection; cerebellopontine angle dissection; internal auditory canal exposure
Gokyar et al.,[40] 2018	Bovine	Interhemispheric dissection; Sylvian fissure dissection; sulcal dissection
Elsayed et al.,[41] 2019	Porcine	Cerebellopontine angle dissection; cranial nerve dissection; internal auditory canal exposure

for the use of ultraviolet fluorescence to monitor surgical performance.[31,33,34] The potential exists for the use of a wide variety of wearables during resection procedures to monitor expert performance. These systems include EEG, eye motion, tremor, and instrument tracing technologies.[33,47,48] Functional neuroimaging, a portable, low-cost technology that detects changes in the trainee's brain activity and cerebral blood flow by measuring changes in oxyhemoglobin and deoxyhemoglobin concentration, can be used.[49] The results of studies incorporating artificial tumors in ex vivo models have suggested that more rigorous and comprehensive research platforms for assessment and quantitation of technical skills performance using ex vivo models have the potential to significantly advance the formative and summative assessments of learners.

Group 2 included 12 studies (66%), which were divided into 2 subgroups. In the first subgroup, we assessed their usefulness in developing microsurgical techniques.[23,35-41] In the second, we focused on the defined neurosurgical operative procedures performed.[42-45] In 8 studies, ovine, bovine, and porcine ex vivo models were used to train learners in microsurgical techniques involving interhemispheric transcallosal approaches, Sylvian fissure, cerebellopontine angle dissection, and dissection of cranial nerves and their foramen (Table 4). The other 4 studies included using ex vivo brain models to train cranioplasty and fronto-orbital advancement techniques, placement of intracranial pressure monitors, suturing of the dura, and the use of intraoperative ultrasound (Table 5). In these studies, only a small part of the investigation had focused on model validation and quantitation, the use of wearable systems, and other methods to monitor performance. These issues have limited the reproducibility and usefulness of these models for trainee formative and summative assessment in competency-based training programs.

Table 5. Nonmodified Procedure Using Ex Vivo Animal Brain Models

Investigator	Animal	Task
Hicdonmez et al.,[42] 2006	Ovine	Cranioplasty and fronto-orbital advancement
Vavruska et al.,[43] 2014	Ovine	Intraoperative ultrasound
Hanrahan et al.,[44] 2018	Porcine	Intracranial pressure monitor insertion
Hanrahan et al.,[45] 2018	Porcine	Dura mater suturing

The goal of the development of a best practice guidelines checklist that includes defined subsections and a total score is to provide investigators with suggestions to help improve their studies and to provide readers with methods to assess the quality of the research. The studies identified in our review were scored using the EVBMASE checklist. The interrater reliability between reviewers was calculated and resulted in an observed agreement of 87.2% (Cohen's kappa, 0.71). The results of assessing the group 1 and 2 studies using the EVBMASE checklist are summarized in Table 6. The overall average score of a possible total of 20 for all 18 studies was 6.9 (34.4%). We further divided the scores according to each EVBMASE section. The score for the simulation study design, data structure, and discussion quality sections was 4.2 (42.2%), 1.1 (22.2%), and 1.6 (31.1%), respectively, for all 18 studies. The mean overall score for group 1 was 11 (55%) and 4.8 (24.2%) for group 2, a statistically significant difference ($P = 0.006$). The mean scores for each EVBMASE section were compared between the 2 groups. A statistically significant difference was found for the simulation study design section ($P = 0.003$). However, the mean scores were not significantly different for the data structure ($P = 0.075$) and discussion quality ($P = 0.066$) sections. The individual EVBMASE checklist scores for the simulation study design, data structure, and discussion quality section for each study are shown in Figure 3. The largest differences between groups 1 and 2 were found for element 3, the operative realism of the model (50% vs. 0%), element 5, a description of simulation properties (100% vs. 16.7%), and element 6, model validity (100% vs. 0%; Figure 4).

DISCUSSION

Numerous ex vivo animal brain simulation models have been reported. However, these have used inconsistent reporting methods, limiting the ability to accurately assess and replicate these

Table 6. Results of Study Assessment Using the EVBMASE Checklist

EVBMASE Section	All Studies (n = 18)	Group 1 (n = 6)	Group 2 (n = 12)	Mean Difference (95% CI; P Value)
Simulation study design	4.2 (2–8); 42.2 (20–80)	6.8 (4–8); 68.3 (40–80)	2.9 (2–4); 29.2 (20–40)	3.9 (2–5.8; 0.003)
Data structure	1.1 (0–3); 22.2 (0–60)	1.8 (1–3); 36.7 (20–60)	0.8 (0–3); 15 (0–60)	1 (−0.1 to 2.3; 0.075)
Discussion	1.6 (0–4); 31.1 (0–80)	2.3 (1–4); 46.7 (20–80)	1.2 (0–2); 23.3 (0–40)	1.1 (−0.1 to 2.4; 0.066)
Overall	6.9 (3–15); 34.4 (15–75)	11 (7–15); 55 (35–75)	4.8 (3–9); 24.2 (15–45)	6.2 (2.5–9.9; 0.006)

Data presented as mean score (range); percentage (range), unless noted otherwise.
EVBMASE, ex vivo brain model to assess surgical expertise; CI, confidence interval.

platforms. Our results have indicated that a comprehensive conceptual framework has the potential to improve future studies using ex vivo brain animal models for neurosurgical simulation and to serve as a guide for the neurosurgical research community to rigorously document the conduction of these studies. The EVBMASE checklist is also applicable to ex vivo animal simulation platforms used for training in other surgical disciplines.

For the present investigation, we divided the ex vivo animal brain models into 2 groups. In group 1, modifications to mimic human brain tumor pathology were evident in all the studies. In contrast, in group 2, no attempt had been made to modify the ex vivo models, and these systems were used to teach dissection or other neurosurgical techniques. The average score for all the studies in our review has indicated that these studies adhered poorly to the elements included in the EVBMASE checklist. The group 1 studies had a significantly higher EVBMASE average score compared with the group 2 studies, which was attributed to the simulation study design, although no statistically significant differences were found in the data structure and discussion quality sections. When we examined how each group performed regarding each of the EVBMASE checklist elements, the 3 highest differentiating elements between the 2 groups were the extent of operative realism, description of simulation properties, and model validation. The ability of the group 1 studies in their attempt to replicate the color, location, and haptic qualities of human brain tumors and model validation gave these simulation models an advantage over group 2 in these elements of the checklist. Despite the lack of a statistically significant difference between the 2 groups in the data structure and discussion quality sections, we found that group 2 studies were deficient in many elements of the checklist, including the ability to quantitate the performance metrics and provide participant data, which led to poor data

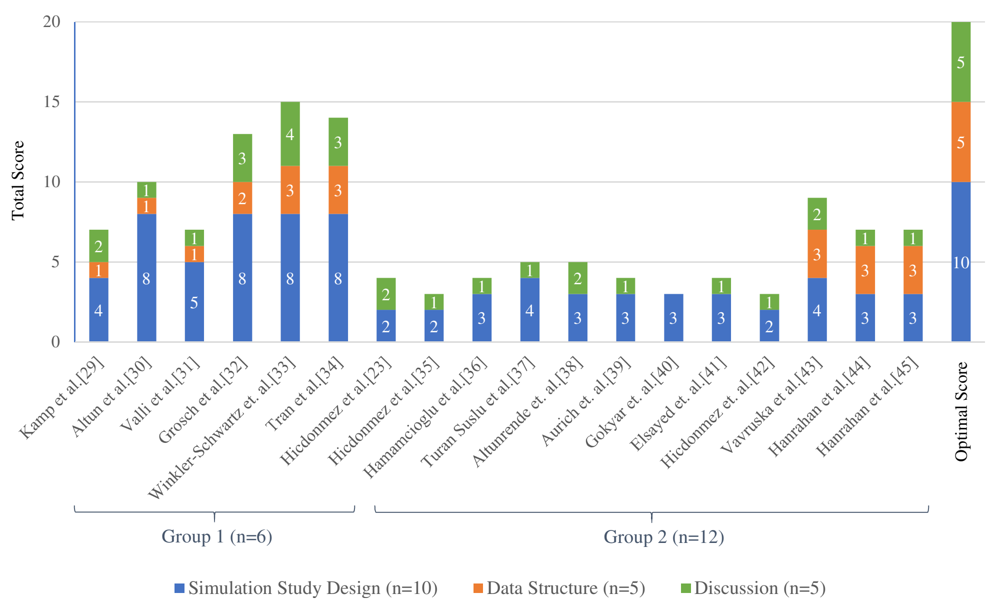

Figure 3. Individual ex vivo brain model to assess surgical expertise (EVBMASE) checklist scores for the simulation study design, data structure, and discussion quality for each study.

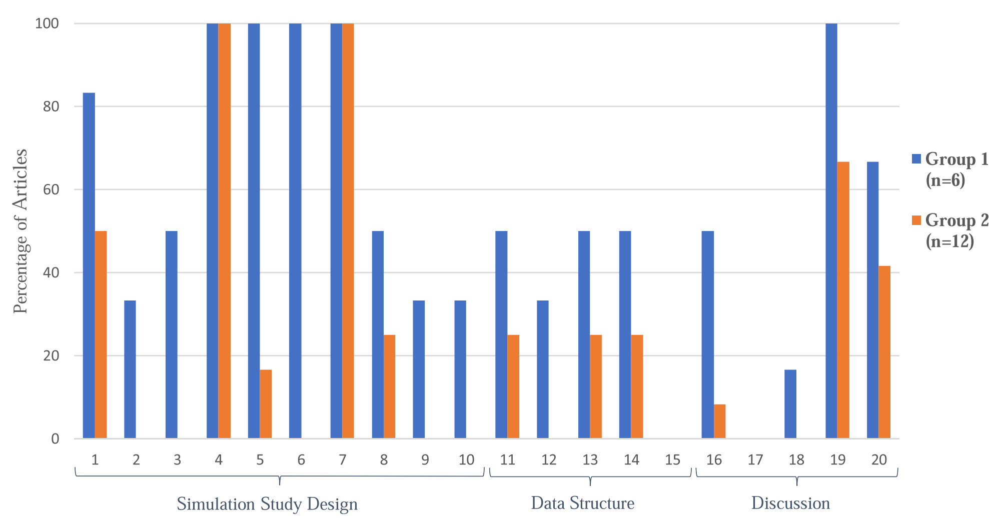

Figure 4. We applied the ex vivo brain model to assess surgical expertise (EVBMASE) checklist to the 18 studies included in the systematic review and calculated the percentage score for the 6 studies in group 1 and 12 studies in group 2 for each of the 20 elements assessed.

structures scores for this group (Figures 3 and 4). We found that the developing microsurgical techniques subgroup was particularly deficient in all sections of the checklist (Figure 3). These weaknesses limited the reproducibility and usefulness of the developing microsurgical technique subgroup models for trainee formative and summative assessments in competency-based training programs.

The members of the surgical education community have identified a need to increase the quality and standardization of the reporting of simulation studies.[19-21] Samaratunga et al.[20] focused their assessment on simulation validation studies in the orthopedic surgery literature and proposed a novel umbrella approach when reporting participant recruitment criteria. They also introduced an approach to classify simulation validity studies into 1 of 3 categories, depending on the skill level of the prioritized participants.[20] The first category encompassed face and content validity. These studies should aim to recruit experts rather than novices because the opinion of an expert surgeon is more useful in judging the realism and pedagogical value of the simulator. The second category included construct, concurrent, and predictive validity. These studies should recruit participants with differing skill levels to differentiate the performance levels and to accurately project the participant's skill level. The third category incorporated studies that aimed to evaluate simulator skill acquisition or skill transfer to an established training model. When recruiting participants for studies within the third category, it will be most useful to prioritize novice learners to effectively demonstrate high learning curves and improved performance.

The proposed best practices criteria in our report considered simulation validity as one of the elements under the simulation study design section. In the simulation study design section, we also identified other factors that improve reproducibility and replicate the operative room experience. These include simulation comprehensiveness and operative realism (Figure 1), along with the description of the simulation model and properties, task performed, assessment tools used, and expertise levels definitions. Our framework also considered the data structure an important component to help facilitate data interpretations when analyzing ex vivo animal brain study results. The discussion quality section is intended to assess ex vivo animal brain reports by considering their benefit to surgical educators and their applicability as assessment tools. Furthermore, the discussion quality section allows for the analysis of deficiencies and possible future evolution of the field. Using the EVBMASE checklist will lead to the scoring of reports, which will result in a quantitative numerical value. This provides an opportunity for investigators to perform quantitative data analysis and permits for the improved replication of results and the use of statistical methods to interpret the pedagogical strength of the studies.

The development of improved ex vivo animal brain models using the EVBMASE checklist will allow for randomized controlled trials to focus on demonstrating the predictive validity of these platforms. Ideally, the conduction of these studies will demonstrate that training on ex vivo animal brain models improves neurosurgical trainee performance and that these enhanced skills are transferable to the operating room, resulting in less operative error and improved patient outcomes.[50]

Study Limitations

The objective of the EVBMASE checklist is to provide a general framework when reporting or analyzing ex vivo animal brain studies. However, the present study had several limitations. First, the checklist was developed and tested using information from a limited number of studies. The EVBMASE checklist will need to be modified as studies using newer methods to assess surgical performance such as EEG, eye tracking, tremor assessment, force sensors, and instrument tracing technologies are performed.[47,48,51-53] Second, we awarded a point if any type of validation was considered in the study reviewed. A more robust point assessment of the validation methods used would seem appropriate as validation studies become more common. Third, artificial intelligence systems have been developed that not only classify participants according to surgical expertise but can also coach trainees to attain defined surgical standards.[26,54-56] These intelligent tutoring systems are being explored to outline the optimal approach to integrate these technologies in psychomotor skills teaching, and these data will need to be incorporated into any future checklists. Finally, as an increasingly holistic understanding of surgical expertise continues to evolve (i.e., one not determined solely by the number of procedures completed or years of practice), care will be required to incorporate this information into future checklists. A clear understanding of expert surgical performance, artificial intelligence methods, and educational best practices will be crucial to the ultimate success of these checklist systems.

Suggestions for Future Authors

We identified new elements that could further enhance the quality of future reports, because the quantitative assessment of these elements will allow investigators to improve and reviewers to better evaluate the quality of the study (Table 2). The EVBMASE checklist point assessment system provides a single point for each element present. However, as more quantitative data on the usefulness of this system in the evaluation of ex vivo animal models becomes available, a more accurate point allocation system will evolve and will be implemented and tested. This could be especially critical related to the further refinement of the validation methods used in ex vivo brain models. More effort is required to create comprehensive research platforms for ex vivo brain simulation (Figure 1). Simulation studies should strive to mimic the specific patient disease processes under investigation and the arc of patient preoperative, operative, and postoperative procedures involved. For some human disease states involving the control of bleeding, for example, ex vivo animal models might not be appropriate, unless modifications are used, which can limit the use of the checklist. It has been suggested that investigators give special attention to each of the 20 elements in the 3 sections of the checklist (simulation study design, data structure, discussion quality components) in their studies. Increasing the sample size in all studies involving these models will increase the ability to use statistical methods to improve the quality of the results. It is crucial to explain how each simulation trial has been used in the analysis. Often, the explanations of the simulation were vague and the exact purpose of the simulation was unclear. More research is needed on how to use the components of this EVBMASE checklist for other types of animal models and surgical education platforms.

CONCLUSIONS

The EVBMASE checklist is a framework to help researchers and surgical educators ensure quality when producing and reviewing reports involving the use of ex vivo animal brain models to assess surgical expertise. We believe the EVBMASE checklist will facilitate the field of simulation and improve surgical education.

CRediT AUTHORSHIP CONTRIBUTION STATEMENT

Ahmad Alsayegh: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – original draft, Writing – review & editing, Visualization, Project administration. **Mohamad Bakhaidar:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing – review & editing. **Alexander Winkler-Schwartz:** Conceptualization, Methodology. **Recai Yilmaz:** Conceptualization, Methodology, Writing – review & editing. **Rolando F. Del Maestro:** Conceptualization, Methodology, Validation, Resources, Writing – review & editing, Supervision.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Carlo Santaguida, Department of Neurology and Neurosurgery, Montreal Neurological Institute and Hospital, McGill University for his help with this project.

REFERENCES

1. Gelinas-Phaneuf N, Del Maestro RF. Surgical expertise in neurosurgery: integrating theory into practice. *Neurosurgery*. 2013;73(suppl 1):30-38.
2. Brightwell A, Grant J. Competency-based training: who benefits? *Postgrad Med J*. 2013;89:107-110.
3. Aboud E, Al-Mefty O, Yasargil MG. New laboratory model for neurosurgical training that simulates live surgery. *J Neurosurg*. 2002;97:1367-1372.
4. Colpan ME, Slavin KV, Amin-Hanjani S, Calderon-Arnuphi M, Charbel FT. Microvascular anastomosis training model based on a Turkey neck with perfused arteries. *Neurosurgery*. 2008; 62(suppl 2):ONS407-ONS410 [discussion: ONS410-ONS411].
5. Breimer GE, Haji FA, Bodani V, et al. Simulation-based education for endoscopic third ventriculostomy: a comparison between virtual and physical training models. *Oper Neurosurg (Hagerstown)*. 2017;13:89-95.
6. Kanazawa R, Teramoto A. The realization of preferable operative working space through the microsurgical training with rats—the importance of the process. *Surg Neurol*. 2009;71:380-387 [discussion: 387].
7. Tayebi Meybodi A, Borba Moreira L, Gandhi S, Preul MC, Lawton MT. Sylvian fissure splitting revisited: applied arachnoidal anatomy and proposition of a live practice model. *J Clin Neurosci*. 2019;61:235-242.
8. Mucke T, Scholz M, Kesting MR, Wolff KD, Schmieder K, Harders AG. Microsurgically induced aneurysm models in rats, part II: clipping, shrinking and micro-Doppler sonography. *Minim Invasive Neurosurg*. 2008;51:6-10.
9. Matsumura N, Hamada H, Yamatani K, Hayashi N, Hirashima Y, Endo S. Side-to-side arterial anastomosis model in the rat internal and external carotid arteries. *J Reconstr Microsurg*. 2001; 17:263-266.
10. Matsumura N, Endo S, Hamada H, Kurimoto M, Hirashima Y, Takaku A. An experimental model for side-to-side microvascular anastomosis. *J Reconstr Microsurg*. 1999;15:581-583.
11. Higurashi M, Qian Y, Zecca M, Park YK, Umezu M, Morgan MK. Surgical training technology for cerebrovascular anastomosis. *J Clin Neurosci*. 2014;21:554-558.
12. Olabe J, Olabe J. Microsurgical training on an in vitro chicken wing infusion model. *Surg Neurol*. 2009;72:695-699.

13. Schoffl H, Hager D, Hinterdorfer C, et al. Pulsatile perfused porcine coronary arteries for microvascular training. *Ann Plast Surg*. 2006;57:213-216.

14. Olijnyk LD, Patel K, Brandao MR, et al. The role of low-cost microsurgical training models and experience with exercises based on a bovine heart. *World Neurosurg*. 2019;130:59-64.

15. Oliveira MM, Araujo AB, Nicolato A, et al. Face, content, and construct validity of brain tumor microsurgery simulation using a human placenta model. *Oper Neurosurg (Hagerstown)*. 2016;12:61-67.

16. Badash I, Burtt K, Solorzano CA, Carey JN. Innovations in surgery simulation: a review of past, current and future techniques. *Ann Transl Med*. 2016;4:453.

17. Yiasemidou M, Gkaragkani E, Glassman D, Biyani CS. Cadaveric simulation: a review of reviews. *Ir J Med Sci*. 2018;187:827-833.

18. Cook DA, Beckman TJ, Bordage G. Quality of reporting of experimental studies in medical education: a systematic review. *Med Educ*. 2007;41: 737-745.

19. 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;76:1681-1690.

20. Samaratunga R, Johnson L, Gatzidis C, Swain I, Wainwright T, Middleton R. A review of participant recruitment transparency for sound validation of hip surgery simulators: a novel umbrella approach. *J Med Eng Technol*. 2021;45:434-456.

21. Van Nortwick SS, Lendvay TS, Jensen AR, Wright AS, Horvath KD, Kim S. Methodologies for establishing validity in surgical simulation studies. *Surgery*. 2010;147:622-630.

22. Gallagher AG, O'Sullivan GC. *Fundamentals of Surgical Simulation: Principles and Practice*. New York, NY: Springer; 2012.

23. Hicdonmez T, Hamamcioglu MK, Tiryaki M, Cukur Z, Cobanoglu S. Microneurosurgical training model in fresh cadaveric cow brain: a laboratory study simulating the approach to the circle of Willis. *Surg Neurol*. 2006;66:100-104 [discussion: 104].

24. Ledwos N, Mirchi N, Bissonnette V, Winkler-Schwartz A, Yilmaz R, Del Maestro RF. Virtual reality anterior cervical discectomy and fusion simulation on the novel sim-ortho platform: validation studies. *Oper Neurosurg (Hagerstown)*. 2020; 20:74-82.

25. Sawaya R, Alsideiri G, Bugdadi A, et al. Development of a performance model for virtual reality tumor resections. *J Neurosurg*. 2018;131:192-200.

26. Mirchi N, Bissonnette V, Yilmaz R, Ledwos N, Winkler-Schwartz A, Del Maestro RF. The virtual operative assistant: an explainable artificial intelligence tool for simulation-based training in surgery and medicine. *PLoS One*. 2020;15:e0229596.

27. Mirchi N, Ledwos N, Del Maestro RF. Intelligent tutoring systems: re-envisioning surgical education in response to COVID-19. *Can J Neurol Sci*. 2020;48:198-200.

28. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6:e1000097.

29. Kamp MA, Knipps J, Steiger HJ, et al. Training for brain tumour resection: a realistic model with easy accessibility. *Acta Neurochir (Wien)*. 2015;157: 1975-1981 [discussion: 1981].

30. Altun A, Cokluk C. The microneurosurgical training model for intrinsic and extrinsic brain tumor surgery using polyurethane foam and fresh cadaveric cow brain: an experimental study. *World Neurosurg*. 2019;4:100039.

31. Valli D, Belykh E, Zhao X, et al. Development of a simulation model for fluorescence-guided brain tumor surgery. *Front Oncol*. 2019;9:748.

32. Grosch AS, Schroder T, Schroder T, Onken J, Picht T. Development and initial evaluation of a novel simulation model for comprehensive brain tumor surgery training. *Acta Neurochir (Wien)*. 2020; 162:1957-1965.

33. Winkler-Schwartz A, Yilmaz R, Tran DH, et al. Creating a comprehensive research platform for surgical technique and operative outcome in primary brain tumor neurosurgery. *World Neurosurg*. 2020;144:e62-e71.

34. Tran DH, Winkler-Schwartz A, Tuznik M, et al. Quantitation of tissue resection using a brain tumor model and 7-T magnetic resonance imaging technology. *World Neurosurg*. 2021;148:e326-e339.

35. Hicdonmez T, Hamamcioglu MK, Parsak T, Cukur Z, Cobanoglu S. A laboratory training model for interhemispheric-transcallosal approach to the lateral ventricle. *Neurosurg Rev*. 2006;29:159-162.

36. Hamamcioglu MK, Hicdonmez T, Tiryaki M, Cobanoglu S. A laboratory training model in fresh cadaveric sheep brain for microneurosurgical dissection of cranial nerves in posterior fossa. *Br J Neurosurg*. 2008;22:769-771.

37. Turan Suslu H, Ceylan D, Tatarli N, et al. Laboratory training in the retrosigmoid approach using cadaveric silicone injected cow brain. *Br J Neurosurg*. 2013;27:812-814.

38. Altunrende ME, Hamamcioglu MK, Hicdonmez T, Akcakaya MO, Birgili B, Cobanoglu S. Microsurgical training model for residents to approach to the orbit and the optic nerve in fresh cadaveric sheep cranium. *J Neurosci Rural Pract*. 2014;5: 151-154.

39. Aurich LA, Silva Junior LF, Monteiro FM, Ottoni AN, Jung GS, Ramina R. Microsurgical training model with nonliving swine head: alternative for neurosurgical education. *Acta Cir Bras*. 2014;29:405-409.

40. Gokyar A, Cokluk C. Using of fresh cadaveric cow brain in the microsurgical training model for sulcal-cisternal and fissural dissection. *J Neurosci Rural Pract*. 2018;9:26-29.

41. Elsayed M, Torres R, Sterkers O, Bernardeschi D, Nguyen Y. Pig as a large animal model for posterior fossa surgery in oto-neurosurgery: a cadaveric study. *PLoS One*. 2019;14:e0212855.

42. Hicdonmez T, Parsak T, Cobanoglu S. Simulation of surgery for craniosynostosis: a training model in a fresh cadaveric sheep cranium. Technical note. *J Neurosurg*. 2006;105(suppl):150-152.

43. Vavruska J, Buhl R, Petridis AK, Maslehaty H, Scholz M. Evaluation of an intraoperative ultrasound training model based on a cadaveric sheep brain. *Surg Neurol Int*. 2014;5:46.

44. Hanrahan J, Sideris M, Tsitsopoulos PP, et al. Increasing motivation and engagement in neurosurgery for medical students through practical simulation-based learning. *Ann Med Surg (Lond)*. 2018;34:75-79.

45. Hanrahan J, Sideris M, Pasha T, et al. Hands train the brain—what is the role of hand tremor and anxiety in undergraduate microsurgical skills? *Acta Neurochir (Wien)*. 2018;160:1673-1679.

46. Sabbagh AJ, Bajunaid KM, Alarifi N, et al. Roadmap for developing complex virtual reality simulation scenarios: subpial neurosurgical tumor resection model. *World Neurosurg*. 2020;139: e220-e229.

47. Pugh CM, Ghazi A, Stefanidis D, Schwaitzberg SD, Martino MA, Levy JS. How wearable technology can facilitate AI analysis of surgical videos. *Ann Surg Open*. 2020;1:e011.

48. Siyar S, Azarnoush H, Rashidi S, Del Maestro RF. Tremor assessment during virtual reality brain tumor resection. *J Surg Educ*. 2020;77:643-651.

49. Gao Y, Yan P, Kruger U, et al. Functional brain imaging reliably predicts bimanual motor skill performance in a standardized surgical task. *IEEE Trans Biomed Eng*. 2020;68:2058-2066.

50. Rolston JD, Zygourakis CC, Han SJ, Lau CY, Berger MS, Parsa AT. Medical errors in neurosurgery. *Surg Neurol Int*. 2014;5(suppl 10): S435-S440.

51. Lu S, Sanchez Perdomo YP, Jiang X, Zheng B. Integrating eye-tracking to augmented reality system for surgical training. *J Med Syst*. 2020;44:192.

52. Li T, King NKK, Ren H. Disposable FBG-based tridirectional force/torque sensor for aspiration instruments in neurosurgery. *IEEE Trans Ind Electron*. 2020;67:3236-3247.

53. Marcus HJ, Payne CJ, Kailaya-Vasa A, et al. A "smart" force-limiting instrument for microsurgery: laboratory and in vivo validation. *PLoS One*. 2016;11:e0162232.

54. 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.

55. 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.

56. Bissonnette V, Mirchi N, Ledwos N, et al. Artificial intelligence distinguishes surgical training levels in a virtual reality spinal task. *J Bone Joint Surg Am*. 2019;101:e127.

Conflict of interest statement: The present study was supported by the Franco Di Giovani Foundation and the Montreal Neurological Institute and Hospital. Ahmad Alsayegh and Mohamad Bakhaidar were supported by King Abdulaziz University, Jeddah, Saudi Arabia. Alexander Winkler-Schwartz was supported by the Fonds de Recherche du Québec - Santé and a Robert Maudsley Fellowship for Studies in Medical Education from the Royal College of Physicians and Surgeons of Canada.

Received 13 June 2021; accepted 12 August 2021

Citation: World Neurosurg. (2021) 155:e369-e381.
https://doi.org/10.1016/j.wneu.2021.08.061

Journal homepage: ***www.journals.elsevier.com/world-neurosurgery***

Available online: ***www.sciencedirect.com***