

# Investigation of Objectivity in Scoring and Evaluating Microvascular Anastomosis Simulation Training

Yasuo MURAI,<sup>1</sup> Shun SATO,<sup>1</sup> Atsushi TSUKIYAMA,<sup>1</sup> Asami KUBOTA<sup>1</sup>,  
and Akio MORITA<sup>1</sup>

<sup>1</sup>Department of Neurological Surgery, Nippon Medical School Hospital, Tokyo, Japan

## Abstract

The increase in minimally invasive surgery has led to a decrease in surgical experience. To date, there is only limited research examining whether skills are evaluated objectively and equally in simulation training, especially in microsurgery. The purpose of this study was to analyze the objectivity and equality of simulation evaluation results conducted in a contest format. A nationwide recruitment process was conducted to select study participants. Participants were recruited from a pool of qualified physicians with less than 10 years of experience. In this study, the simulation procedure consisted of incising a 1 mm thick blood vessel and suturing it with a 10-0 thread using a microscope. Initially, we planned to have the neurosurgical supervisors score the simulation procedure by direct observation. However, due to COVID-19, some study participants were unable to attend. Thus requiring some simulation procedures to be scored by video review. A total of 14 trainees participated in the study. The Cronbach's alpha coefficient among the scorers was 0.99, indicating a strong correlation. There was no statistically significant difference between the scores from the video review and direct observation judgments. There was a statistically significant difference ( $p < 0.001$ ) between the scores for some criteria. For the eight criteria, individual scorers assigned scores in a consistent pattern. However, this pattern differed between scorers indicating that some scorers were more lenient than others. The results indicate that both video review and direct observation methods are highly objective techniques evaluate simulation procedures.

Keywords: microsurgery, objective assessment, skills, techniques

## Introduction

The Coronavirus Disease 2019 (COVID-19) pandemic has forced restructuring of surgical training programs around the world.<sup>1–3</sup> Simulation training for surgical procedures is a proven method to improve surgeon skills.<sup>4–6</sup> Some countries have adopted virtual reality simulation for training and recertification for surgical specialties, including obstetrics and gynecology<sup>6</sup> and abdominal surgery.<sup>7–9</sup> However, there is only limited reported research on the use of simulation training in the field of microsurgery.<sup>10–14</sup> This research<sup>11,13,15</sup> only discusses the development of simulation equipment and does not extend to

analysis of the objectivity of the evaluation results from the simulation training. A survey by the American Society for Reconstructive Microsurgery reported that only about 6% of microsurgeons had experience with high-precision simulation training.<sup>10</sup> Additionally, 24% of the respondents thought that simulation training was a useful indicator of clinical performance, despite having no actual experience with simulation training.<sup>10</sup>

In Japan, the Organization for the Certification of Cardiovascular Surgeons (which consists of three academic societies including the Japanese Society for Vascular Surgery) mandates 30 hours of off-the-job training under a new medical specialist system. The Japanese Society of Endoscopic Surgeons requires an educational seminar of more than 10 hours and at least three practical training sessions for technical certification. However, there is no mention of off-the-job training for board certification by The Japan Neurosurgical Society. Japanese Society on

Received June 16, 2021; Accepted August 18, 2021

Copyright© 2021 The Japan Neurosurgical Society  
This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives International License.

Surgery for Cerebral Stroke has a technical certification medical education seminar, but the content is limited to only a single 3-hour session with content that corresponds to off-the-job training. The Japanese Society of Plastic and Reconstructive Surgery (which conducts microsurgery) have a similar system for medical specialists but with no mention of off-the-job training. This is because the models that can be used in microsurgery for the fields of plastic surgery<sup>10,14)</sup> and neurosurgery<sup>11,13)</sup> are microscopic in nature. Consequently, research on the objectivity of evaluation may be difficult to conduct.

The evaluation of skills from simulation practice should not just confirm that a trainee has practiced but also identify the strengths and weaknesses of individual skills and report these back to the trainee. Thus, it is necessary to analyze how differences in evaluator scoring patterns may affect trainee results. There is only limited research examining the objectivity of evaluations of microsurgical techniques by scorers from different backgrounds or differences in scoring patterns between video review and actual face-to-face review of simulation procedures. This paper investigates the effect of scoring patterns on simulation training procedure evaluations even for situations where it is difficult for trainees to participate in-person due to the COVID-19 restrictions. Additionally, to objectively evaluate the improvement of skills over time due to simulation training, it needs to be confirmed that similar evaluations are possible even if there is a change of evaluator (e.g., due to retirement or transfer) or the evaluation method (video or face-to-face) changes.

The purpose of this study is to clarify the effectiveness of simulation-based evaluation in microsurgery education by analyzing the results of simulation training evaluations conducted in a contest format with many experts participating as judges. Additionally, this study assesses the objectivity and equality of these results.

## Materials and Methods

### Materials

We invited neurosurgeons from all over Japan who became physicians in 2010 or later to participate in a surgical technique contest with the task of incising and suturing artificial blood vessels. The contest was open to the public, and the judges observed the participants directly.

The task contents were uploaded to YouTube and made public. Each participant was given 5 minutes to incise and suture a 1 mm diameter artificial blood vessel. The suture thread was 10-0 nylon, and the participants could select, the micro forceps, scissors,

and needle holder for the procedure. An identical artificial blood vessel was prepared by the organizers to ensure equality and fairness. A scoring chart was prepared based on the Objective Structured Assessment of Technical Skill scoring method.<sup>16)</sup>

The scores were assigned based on a 6-point scale where 0 = Failing, 1 = Bad, 2 = Poor or Below Average, 3 = Fair or Average, 4 = Good or Fairly Good, and 5 = Excellent or Almost Excellent. There were eight evaluated criteria, resulting in a maximum possible score 40 points assigned per scorer. Each participant was evaluated by 5 scorers, so the maximum total score for the contest was 200 points. Details about of the eight evaluation criteria follow:

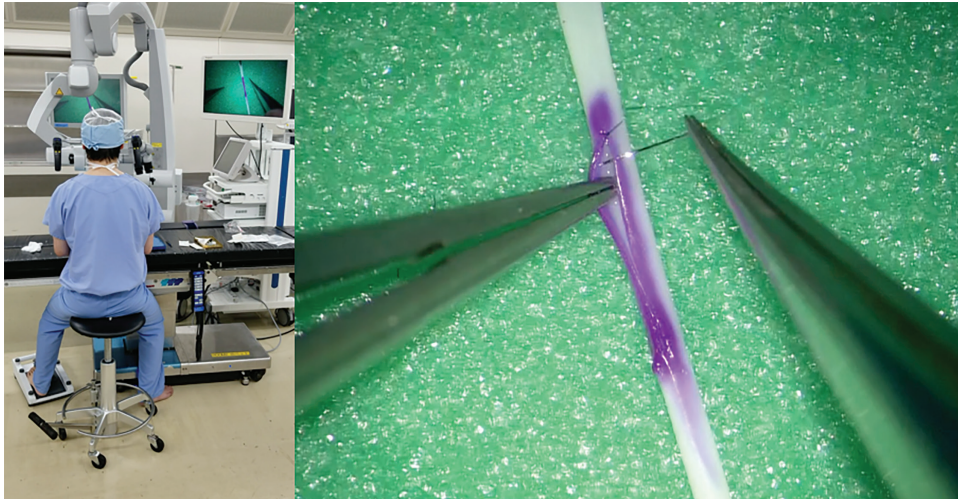
1. POSTURE: Surgeons Positioning/Instrument layout (preparation).
2. MICROSCOPE: microscope operation/Focus, Positioning, Hand movement (Knowledge of microscope),
3. TREMOR: Hand tremor (Motion),
4. CUT: Line drawing, linear incision, Use of scissors (Knowledge of micro instrument),
5. NEEDLE: Handling of Needle/Insertion Angle/Vessel Treatment (Handling of needle and thread),
6. SUTURE: Pull and Stop of thread (Respect for vessel wall),
7. STITCH: Ligation/making a loop pull direction (Flow of suture),
8. KNOT: final ligature/Slack, direction, length (Knowledge of completion).

### Methods

This study was conducted with the approval of the Ethical Review Committee of the Medical University Hospital.

We asked neurosurgeons who are qualified as supervisors by The Japan Neurosurgical Society and Japanese Society on Surgery for Cerebral Stroke from different universities and affiliated institutions to be the scorers for this experiment. Selection of scorers from these societies ensured that differences in surgical techniques and philosophical approaches were represented in the study. After the contest, the specialist judges discussed technique strengths and weaknesses with each participant due to the impact of COVID-19, the contest was rescheduled from March 2020 to September 2020, and the contest format was changed from requiring an in-person format for all participants to using a web conference for some participants. Video judging was provided for those who wanted to participate but was unable to attend the contest in-person due to restrictions on domestic travel.

An example of a video review was prepared and sent to participants not able to attend in-person.



**Fig. 1** The review video consisted of a video of the microscope screen and a video of the surgeon's entire body taken from the side and back of the surgeon.

The review video consisted of a video of the microscope screen and a video of the surgeon's entire body taken from the side and back perspectives of the surgeon (Fig. 1). The contest organizer mailed a urethane stand to fix the artificial blood vessel used for suturing, a 1-mm artificial blood vessel, and 10-0 thread to each of these participants. Additionally, this video review and the scoring sheet were sent to the judges. Judges were asked to write down the advantages and disadvantages of each participant's technique as a comment. Results of these evaluations were sent to the participants so that they could assess their skills. The same scoring chart was used for both the video review and on-site judging.

A venue with a 200 person capacity was chosen as the final judging site. However, due to COVID-19 restrictions, only about 40 people attended. A KINEVO 900 (Carl Zeiss medic, Germany) surgical microscope was placed in the front center of the room, and with five judges positioned around this microscope (Fig. 2). The final ranking score for each participant was determined based on the sum of the scores from all judges. Final contest ranking was determined by comparing average scores from the judges. For fairness, a judge's score was excluded from the final average if he/she were from the same institution as the participant. For instances where participants became doctors in the same year and had identical final average scores, the younger doctor was ranked higher.

### Statistical analysis of evaluation results

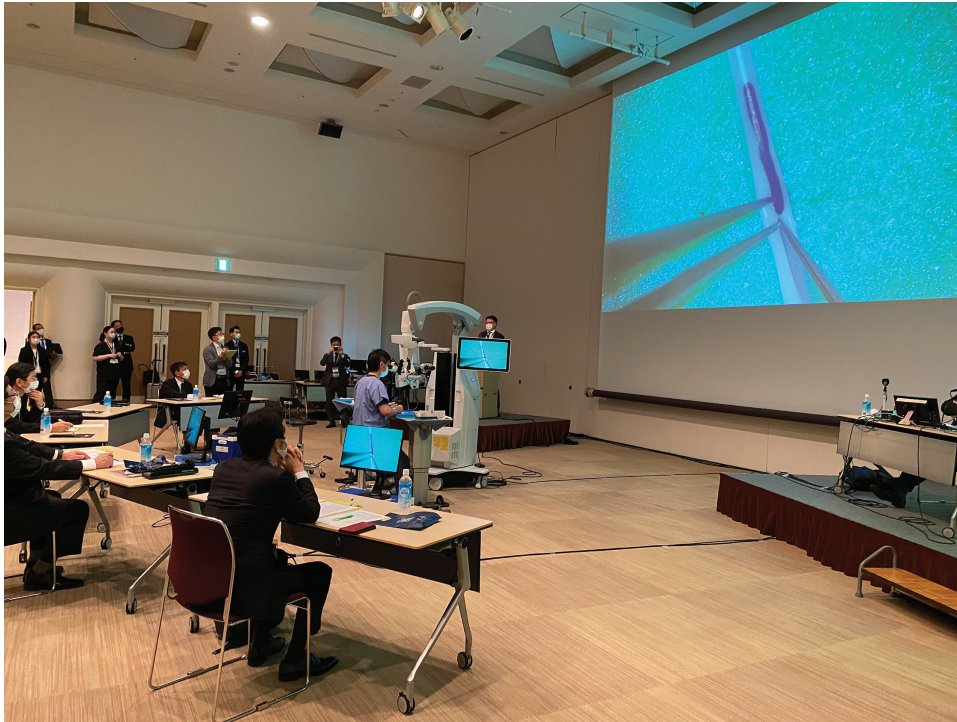
The evaluation scores from all the judges for both the video and on-site judging were averaged to form

final rankings. We examined whether there were any differences in the rankings patterns among the judges. Statistical analysis was performed to assess the following: (1) differences in overall scores and years of medical experience between the video and on-site judging; (2) differences in individual judging scores between video review and on-site judging; and (3) correlation analysis of all judges in the video and on-site judging. A commercially available software package (JMP Pro 14, SAS Institute, Cary, NC, USA) was used for the statistical analyses. All statistical results are presented as mean and standard error of mean or median and standard deviation. Statistical significance was evaluated using the nonparametric Mann-Whitney U test and Fisher's exact test. These tests were chosen due to heterogeneity of variance and the small sample size. The DUNN test with merged ranks was used for all paired mean testing between multiple groups. The statistical significance level was set to  $p < 0.05$ . Consistency of ratings among scorers was estimated using the Cronbach's alpha coefficient.

## Results

Initially, the March 2020 conference had an enrollment of 18 participants. But this conference had to be postponed due to the COVID-19 pandemic. When the rescheduled contest was held in September 2020, only six participants were able to attend the in-person contest. Reduction in attendance was due to work relocation, study abroad, and domestic travel restrictions. There were eight participants in the video review contest and four participants who attended both the in-person and video review contests. One





**Fig. 2** A view of the venue at the actual contest.

of the video review participants was currently studying abroad during the contest. Five judges participated in both the video review and on-site contests two of which participated in both contests.

The average number of years of graduation for the eight video review participants was  $13.8 \pm 0.86$  years, and the average number of years of graduation for the six face-to-face participants was  $14.3 \pm 0.66$  years. The average contest rating score of the eight participants in the video review was  $148.3 \pm 4.73$ , and the average rating score of the six participants in the face-to-face session was  $137.0 \pm 7.30$ .

There was no significant difference in overall score or years of medical experience between the participants in either of the contests. Of the four participants who participated in both the video review and in-person contests, two had high scores in the video review contest and two had high scores in the in-person contest. To examine whether there was a difference in the scoring results between the video review contest and in-person review contest, the scoring results of the two scorers and the four participants who participated in both were compared using the Wilcoxon rank-sum test for each scoring item. There was no statistically significant difference between the scoring results of the video review and in-person contests (mean  $\pm$  standard error of video review;  $3.906 \pm 0.1214$ , and in-person contest;  $4.016 \pm 0.1102$ ,  $p = 0.55$ ). Only participants and scorers

who participated in both video review and in-person contest were selected, and Cronbach's alpha coefficients were calculated for the consistency between the video review and in-person contests. The Cronbach's alpha coefficient was 0.9520, indicating high consistency.

Table 1 shows the participant evaluation scores by judge and overall total for the video review contest (Table 1 upper) and in-person contest. (Table 1 lower). Examination of these tables indicates that there between differences judge scoring patterns but final participant rankings tended to be similar.

Table 2 shows the statistical results for each evaluation criteria broken down by judge and overall total. Judges for each evaluation criteria for the video review contest (Table 2 upper, A–E) and the in-person contest (Table 2 lower, F–J).

Review of the video review contest results shows (Table 2 upper, A–E) that there was no statistically significant difference between judges in scoring the cutting and needle criteria. However, the other six criteria showed a statistically significant difference in scores between judges. Of the six items, Judge C scored statistically significantly lower than the other judges on five of the six evaluate criteria. For four of the six evaluated criteria, Judge D gave statistically significant higher scores than the other judges. Judge D also scored statistically significantly higher than Judge C in the total score. ( $p = 0.0118$ ).

**Table 1 Results of video review (upper) and in-person (lower) contest**

Grader	Judge A	Rank	Judge B	Rank	Judge C	Rank	Judge D	Rank	Judge E	Rank	Final Total score
1	39	1	32	1	27	3	39	2	30	3	167
2	37	2	30	4	26	4	37	3	32	1	162
3	27	7	32	1	33	1	40	1	25	6	157
4	31	4	25	6	29	2	35	6	29	4	149
5	29	6	25	6	25	5	32	7	32	1	143
5	31	4	28	5	21	6	36	5	27	5	143
7	37	2	24	8	18	8	37	3	23	7	139
8	22	8	31	3	21	6	32	7	20	8	126
Final Rank	Judge F	Rank	Judge G	Rank	Judge H	Rank	Judge I	Rank	Judge J	Rank	Final Total score
1	33	1	32	1	33	1	33	1	39	1	170
2	31	2	27	2	22	3	26	4	36	3	142
3	24	6	27	2	21	5	27	2	37	2	136
4	26	3	21	4	23	2	27	2	33	5	130
5	25	4	20	5	18	6	24	5	35	4	122
6	26	3	20	5	22	3	21	6	33	5	122

For the in-person contest (Table 2 lower, F–J), there was no statistically significant difference in scoring between the judges for the posture, ligation, and knots evaluation criteria. The remaining five evaluation criteria had a statistically significant difference in scores between judges. In particular, Judge I statistically significantly scored higher than Judge F ( $p = 0.0083$ ) and Judge H ( $p = 0.001$ ) for the in-person contest.

Note: Judge C in the video review contest and Judge H in the in-person contest are not the same person. Similarly, Judge D in the video review contest and Judge J in the in-person contest are not the same person.

Although there was no statistically significant difference between the scoring results of the video review and in-person contests, Cronbach's alpha coefficient was calculated to verify the correlation of the scoring among the scorers in each of the video review and in-person contests. The alpha coefficient of Cronbach's among the scorers of the in-person contest was 0.9914, indicating high consistency. In the video review contest, the Cronbach's alpha coefficient was 0.9902, which also shows a high degree of consistency. The partial correlation coefficients extracted for individual judges of the video review and in-person contests were also highly consistent, above 0.95 for all combinations (Table 3).

## Discussion

Using the current evaluation criteria, there was homogeneity in the evaluation results among the five judges in both the video review and in-person contests. In both contests, judges tended to give higher or lower evaluations scores depending on the individual evaluation criteria. Some judges tended to give "severe" evaluation scores while some others tended to give "gentle" scores evaluations. Two judges participated in both the video review and the in-person review contests, but neither of these judges who were analyzed as assigning "harsh" or "gentle" scoring patterns. There was no statistically significant difference in the scoring results when comparing only the data of the scorers and participants who participated in both the video review and in-person contests.

Additionally, from the results in Table 2, the rankings were reconfirmed on the basis of the total scores when scorers C, D, I, or G, who had statistically significant differences in total scores, were excluded. The ranking was exactly the same as the final result when scorers D or G were excluded. When only scorer C was removed, the top three and the last place were exactly the same. The top four places were exactly the same even when only scorer I was removed. However, further research is needed to determine whether the contestants would have the same rankings if they were scored by five

**Table 2 Results of video review broken down by evaluation criteria (Judges A-E), and results of in-person contest between broken down by evaluation criteria (Judges F-J)**

Judge	Posture		Microscope		Tremor		Cut		Needle		Suture		Stitch		Knot		Total			
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE		
A	4.38	0.18	4.13	0.35	3.38	0.53	4	0.33	4.25	0.37	4.13	0.35	3.38	0.26	**	4	0.33	31.63	2.04	
B	3.88	0.13	3.38	0.26	3.75	0.16	3	0.42	3.75	0.16	3.75	0.16	*	3.5	0.19	*	3.38	0.26	28.38	1.18
C	3.38	0.18	**	3.25	0.25	*	3.13	0.3	3.25	0.31	3.13	0.3	**	3	0.27	**	3.13	0.23	25	1.72
D	4.13	0.23	4.5	0.19	4.62	0.18	4.38	0.18	3.88	0.23	4.88	0.13	5	0			4.63	0.26	36	1.04
E	3.88	0.13	3.75	0.16	3.13	0.3	**	3	3.25	0.31	3.5	0.27	**	3.38	0.32	**	3.38	0.26	27.25	1.53
F	3.67	0.21	*	3.5	0.22	*	3.17	0.4	2.67	0.33	4.17	0.31	4.17	0.17			3.17	0.31	27.5	1.48
G	3.57	0.2	*	3.29	0.18	**	3.29	0.29	3	0.38	2.71	0.29	**	3	0.38		3.14	0.26	24.86	1.61
H	3.29	0.29	**	3.29	0.36	**	2.71	0.29	*	2.43	0.3	**	3	0.38		3	0.53	23.57	1.81	
I	3.43	0.3	**	3.57	0.3	*	3.14	0.26	3	0.22	3.43	0.3	3.43	0.3		3.43	0.2	26.57	1.39	
J	5	0	5	0	4	0.38	4	0.22	4.29	0.29	4.71	0.18	4.43	0.3		4.43	0.3	35.86	0.88	

\*p<0.05, \*\*p<0.01. SE: standard error

**Table 3 The results of the partial correlation coefficients extracted for the individual judges (A-J) for the video review (upper) and in-person contest (lower)**

	B	C	D	E
A	0.9579	0.9433	0.9791	0.9754
B	(-)	0.9731	0.9877	0.9656
C		(-)	0.9757	0.97
D			(-)	0.9757
E				(-)
	G	H	I	J
F	0.9812	0.9799	0.9851	0.988
G	(-)	0.9808	0.9906	0.9848
H		(-)	0.9847	0.9753
I			(-)	0.9907
J				(-)

completely different scorers. The results of this study indicate that video review and in-person contests can produce statistically homologous results if the same expert is in charge of the review and if the task content and scoring tables are standardized.

The impact of COVID-19 has created a need for social distancing. This has affected the simulation educational environment. In other words, it has become difficult for many trainees to assemble in-person, practice a simulation, and subsequently be evaluated on their performance. Of course, the purpose of simulation training is not competition in surgical procedure contests, but to motivate trainees to improve their skills and to appropriately deliver performance feedback back to them.<sup>17)</sup> To achieve this, we conducted this study to confirm that objective and equal evaluation are not affected by differences in judges, video observation, or in-person observation. Results from this study show that it is possible to objectively evaluate the technique of microsurgery using a web system because the details of the technique can be projected on a monitor. This is an advantage of microsurgical education, since it is easy to record detailed techniques.<sup>14)</sup> After both contests conducted for this study, contestants were given the opportunity to discuss the strengths and weaknesses of their techniques with the judges and commented to the study organizers that their participation was very meaningful.<sup>17)</sup>

**Simulation materials needed for microsurgical technique evaluation**

In this contest, we used artificial blood vessels. Artificial blood vessels were selected based on the

following factors: availability of the same simulation material for all participants, cost, safety, animal welfare, ease of access, and similarity to actual surgery. There are many reports on the practice of vascular anastomosis using arteries from chicken wings,<sup>13,14)</sup> and the similarity to actual human blood vessels has been discussed. However, for this study, artificial blood vessels were selected because they could serve as a better guarantee of identity than a chicken wing artery. However, if artificial blood vessels are not available, chicken wing arteries are a feasible alternative. A 5-minute time limit for each participant was set to allow time for scoring after the observation of the actual procedure, as well as provide sufficient time for participant change and preparation. The in-person contest with six participants took more than 3 hours from the start to the end of the contest, including 2 hours of participant discussion with judges; obviously, more participants will require more time, accordingly. If chicken wing arteries are used, drying of the wing blood vessels is required in the preparation, which will require more time for the contest. In addition to artificial blood vessels and chicken wing arteries, other simulation training that reported using placenta<sup>18)</sup> and vascular models created by 3D printers.<sup>19,20)</sup> These models have higher similarity to the actual surgical field than use of artificial blood vessels.<sup>17,21,22)</sup> It has been suggested that these alternatives may contribute to the improvement of surgical techniques, but identical items are difficult to prepare, thus limiting their use in objective evaluation programs. Furthermore, they are not readily available.

To investigate and compare the educational effects of simulation-based training models, McGaghie, et al.<sup>23)</sup> proposed a translational outcome effect classification for simulation-based training models. This<sup>23)</sup> is a five-level classification of effectiveness, ranging from trainee satisfaction (level of effectiveness 1) to patient outcomes (level of effectiveness 4), cost reduction, and skill improvement (level of effectiveness 5). Skill assessment with simulator tools is categorized per their level of effectiveness. A recent review of 108 articles on neurosurgical simulation training reported<sup>17)</sup> that there were 15 models at level 2 and only six models above level 2. In other words, most of the papers on simulation training are about useful methods for improving skills, but not about objective evaluation of skills. The objective evaluation method for microsurgical skills used in this study is McGaghie's classification<sup>23)</sup> level 2, but previous reports that fall into this classification have used 3D model<sup>19,20)</sup> creations and cadavers.<sup>24)</sup> Both of these preparations require

the time and costs for model creation. Patel et al.<sup>17)</sup> pointed out the need for a cost-effective training simulator to continue simulation training. Our method is a model that can be practiced with an artificial blood vessel or a winged blood vessel and has the potential to be a cost effective method of technical evaluation when used in conjunction with a web-based screening method.

## Limitations

Due to the small number of participants, multifactorial statistical analysis could not be performed. We requested expert reviewers from all over Japan, but not all facilities have experts, and some facilities do not have people who can conduct evaluations. The use of an artificial vessel model allows for a more basic setup at a lower cost. However, it limits the ability to evaluate completion due to the lack of active blood flow to evaluate anastomotic patency.

Ultimately, it is necessary to be able to objectively evaluate the relationship between actual improvement in procedure skill and subsequently with patient prognosis. In other words, we would like to investigate the correlation between evaluation results from simulation training and patient prognosis.

## Acknowledgments

The authors would like to thank professor Miki Fujimura of the Department of Neurosurgery, Graduate School of Medicine, Hokkaido University, professor Ken-ichiro Kikuta of the Department of Neurosurgery, Faculty of Medical Sciences, University of Fukui, professor Satoshi Kuroda of the Department of Neurosurgery, Graduate School of Medicine and Pharmaceutical Sciences, University of Toyama, professor Takatoshi Sorimachi of the Department of Neurosurgery, Tokai University School of Medicine, Katsumi Takizawa of the Department of Neurosurgery, Asahikawa Red Cross hospital, Rokuya Tanikawa of the Department of Neurosurgery, Stroke Center, Teishinkai Hospital, and professor Shinichi Yoshimura; Department of Neurosurgery, Hyogo College of Medicine for their kind support and advice.

## Conflicts of Interest Disclosure

The authors declare that they have no conflicting or competing interests. All authors who are members of The Japan Neurosurgical Society have already registered online self-reported COI Disclosure Statement Forms through the website.



## References

- 1) Wittayanakorn N, Nga VDW, Sobana M, Bahuri NFA, Baticulon RE: Impact of COVID-19 on neurosurgical training in Southeast Asia. *World Neurosurg* 144: e164–e177, 2020
- 2) Hau HM, Weitz J, Bork U: Impact of the COVID-19 pandemic on student and resident teaching and training in surgical oncology. *J Clin Med* 9: 3431, 2020
- 3) Al-Ahmari AN, Ajlan AM, Bajunaid K, et al.: Perception of neurosurgery residents and attendings on online webinars during COVID-19 pandemic and implications on future education. *World Neurosurg* 146: e811–e816, 2021
- 4) Thomson JE, Poudrier G, Stranix JT, Motosko CC, Hazen A: Current status of simulation training in plastic surgery residency programs: A review 45: 395–402, 2018
- 5) Agrawal N, Turner A, Grome L, et al.: Use of simulation in plastic surgery training. *Plast Reconstr Surg Glob Open* 8: e2896, 2020
- 6) Nguyen N, Watson WD, Dominguez E: Simulation-based communication training for general surgery and obstetrics and gynecology residents. *J Surg Educ* 76: 856–863, 2019
- 7) Robinson WP, Baril DT, Taha O, et al.: Simulation-based training to teach open abdominal aortic aneurysm repair to surgical residents requires dedicated faculty instruction. *J Vasc Surg* 58: 247–253, 2013
- 8) Rao R, Dumon KR, Neylan CJ, et al.: Can simulated team tasks be used to improve nontechnical skills in the operating room? *J Surg Educ* 73: e42–e47, 2016
- 9) Shimizu T, Taniai N, Yoshioka M, et al.: Preoperative three-dimensional virtual simulation for safe liver surgery. *J Nippon Med Sch* 81: 354–355, 2014
- 10) Theman TA, Labow BI: Is there bias against simulation in microsurgery training? *J Reconstruct Microsurg* 32: 540–545, 2016
- 11) Sakamoto Y, Okamoto S, Shimizu K, Araki Y, Hirakawa A, Wakabayashi T: Hands-on simulation versus traditional video-learning in teaching microsurgery technique. *Neurologia Medico-Chirurgica* 57: 238–245, 2017
- 12) Cikla U, Sahin B, Hanalioglu S, Ahmed AS, Niemann D, Baskaya MK: A novel, low-cost, reusable, high-fidelity neurosurgical training simulator for cerebrovascular bypass surgery. *J Neurosurg* 130: 1663–1671, 2019
- 13) Abila AA, Uschold T, Preul MC, Zabramski JM: Comparative use of turkey and chicken wing brachial artery models for microvascular anastomosis training. *Journal of Neurosurgery* 115: 1231–1235, 2011
- 14) Hayashi K, Hattori Y, Yii Chia DS, Sakamoto S, Marei A, Doi K: A supermicrosurgery training model using the chicken mid and lower wing. *J Plast Reconstr Aesthet Surg* 71: 943–945, 2018
- 15) Hayashi K, Ozawa E: Vital labelling of somite-derived myogenic cells in the chicken limb bud. *Roux Arch Dev Biol* 200: 188–192, 1991
- 16) VanBlaricom AL, Goff BA, Chinn M, Icasiano MM, Nielsen P, Mandel L: A new curriculum for hysteroscopy training as demonstrated by an objective structured assessment of technical skills (OSATS). *Am J Obst Gynecol* 193: 1856–1865, 2005
- 17) Patel EA, Aydin A, Cearns M, Dasgupta P, Ahmed K: A systematic review of simulation-based training in neurosurgery, part 1: cranial neurosurgery. *World Neurosurg* 133: e850–e873, 2020
- 18) Ribeiro De Oliveira MM, Ferrarez CE, Ramos TM, et al.: Learning brain aneurysm microsurgical skills in a human placenta model: predictive validity. *J Neurosurg* 128: 846–852, 2018
- 19) Mashiko T, Otani K, Kawano R, et al.: Development of three-dimensional hollow elastic model for cerebral aneurysm clipping simulation enabling rapid and low cost prototyping. *World Neurosurg* 83: 351–361, 2015
- 20) Wang L, Ye X, Hao Q, et al.: Comparison of two three-dimensional printed models of complex intracranial aneurysms for surgical simulation. *World Neurosurg* 103: 671–679, 2017
- 21) Patel EA, Aydin A, Cearns M, Dasgupta P, Ahmed K: A systematic review of simulation-based training in neurosurgery, part 2: spinal and pediatric surgery, neurointerventional radiology, and nontechnical skills. *World Neurosurgery* 133: e874–e892, 2020
- 22) Palter VN, Orzech N, Reznick RK, Grantcharov TP: Validation of a structured training and assessment curriculum for technical skill acquisition in minimally invasive surgery. *Ann Surg* 257: 224–230, 2013
- 23) McGaghie WC, Issenberg SB, Barsuk JH, Wayne DB: A critical review of simulation-based mastery learning with translational outcomes. *Med Educ* 48: 375–385, 2014
- 24) Baidya NB, Berhouma M, Ammirati M: Endoscope-assisted retrosigmoid resection of a medium size vestibular schwannoma tumor model: a cadaveric study. *Clin Neurol Neurosurg* 119: 35–38, 2014

---

Corresponding author: Yasuo Murai, MD, PhD  
 Department of Neurological Surgery, Nippon Medical School Hospital, 1-1-5 Sendagi, Bunkyo-ku, Tokyo 113-8603, Japan.  
*e-mail:* ymurai@nms.ac.jp