Face, content, and construct validity of an aneurysm clipping model using human placenta

Evgenii Belykh, MD, Eric J. Miller, BS, Ting Lei, MD, Kristina Chapple, PhD, Vadim A. Byvaltsev, MD, PhD, Robert F. Spetzler, MD, Peter Nakaji, MD, Mark C. Preul, MD

PII: S1878-8750(17)30937-3
DOI: 10.1016/j.wneu.2017.06.045
Reference: WNEU 5916

To appear in: World Neurosurgery

Received Date: 20 March 2017
Revised Date: 7 June 2017
Accepted Date: 8 June 2017


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Face, content, and construct validity of an aneurysm clipping model using human placenta

Evgenii Belykh, MD\textsuperscript{1,2,3}
Eric J. Miller, BS\textsuperscript{4}
Ting Lei, MD\textsuperscript{1}
Kristina Chapple, PhD\textsuperscript{1}
Vadim A. Byvaltsev, MD, PhD\textsuperscript{3}
Robert F. Spetzler, MD\textsuperscript{1}
Peter Nakaji, MD\textsuperscript{1}
Mark C. Preul, MD\textsuperscript{1}

\textsuperscript{1}Department of Neurosurgery
Barrow Neurological Institute
St. Joseph’s Hospital and Medical Center
Phoenix, Arizona

\textsuperscript{2}Arizona State University
School of Life Sciences
Tempe, Arizona

\textsuperscript{3}Irkutsk State Medical University
Irkutsk, Russia
University of Arizona College of Medicine – Phoenix
Phoenix, Arizona

**Correspondence:** Mark C. Preul, MD
c/o Neuroscience Publications; Barrow Neurological Institute
St. Joseph’s Hospital and Medical Center
350 W. Thomas Rd.; Phoenix, AZ 85013
Tel: 602.406.3593; Fax: 602.406.4104
E-mail: Neuropub@dignityhealth.org

**DISCLOSURES:** None

**FINANCIAL SUPPORT:** This research was supported with funds from the Barrow Neurological Foundation, the Women’s Board of the Barrow Neurological Institute, and in part by the Newsome Family Endowment in Neurosurgery to Dr. Preul. Drs. Belykh and Lei are supported by funds from the Barrow Neurological Foundation. Dr. Byvaltsev is supported by funds from the Russian Science Foundation (Project 15-15-30037).

**ACKNOWLEDGMENTS:** The authors would like to thank the staff of the Neuroscience Publications office at Barrow Neurological Institute for support with manuscript preparation.

**SUBMISSION CATEGORY:** Original Article
ABSTRACT

Objective: Proficient open surgical treatment of cerebral aneurysms requires extensive training and practice. The goal of this study was to test the validity of a human placenta aneurysm model for surgical simulation training of aneurysm clipping.

Methods: Thirty participants were divided into 3 groups (n=10 per group) according to their neurosurgical experience: low experience, intermediate experience, and attending. Subjective measures were collected using the Aneurysm Clipping Participant Survey (n=27). Objective measures were collected by observing the participants (n=30) work through a series of aneurysm clipping tasks, while 2 independent raters evaluated them using a newly developed standardized tool, the Objective Structured Assessment of Aneurysm Clipping Skills (OSAACS).

Results: In terms of the subjective measures of face validity, most of the attending group (7/10, 70%) rated the models as “somewhat” to “very well” replicating real surgery. The content validity assessment of the model showed that it could improve the following skills: microdissection technique (27/27, 100%), use of clip applicers for aneurysm clipping (27/27, 100%), and surgical technique when applied to patients (25/27, 93%). The objective measure for construct validity demonstrated that mean OSAACS scores between the 3 groups (low-experience group, 22.9±5.4; intermediate-experience group, 32.8±4.0; attending group, 43.3±1.3) differed significantly (P<0.001).

Conclusions: The human placenta aneurysm clipping model is a useful training tool for teaching residents, with evidence of internal consistency, and face, content, and construct validities. The OSAACS scale is a feasible tool to assess aneurysm clipping skills quantitatively.

RUNNING TITLE: Validity of an aneurysm clipping model
**KEYWORDS:** Cerebral aneurysm; clipping; model; placenta; scale; training; validation

**ABBREVIATIONS:** OSAACS, Objective Structured Assessment of Aneurysm Clipping Skills; PGY, postgraduate year
HIGHLIGHTS

- Human placenta with modeled aneurysms is a relevant simulation model.
- The study showed high face, content, and construct validities of the model.
- The OSAACS is a new tool for quantification of aneurysm clipping skills.
INTRODUCTION

To effectively train new vascular neurosurgeons in the era of restricted duty hours and decreased exposure to open vascular cases, neurosurgery training programs and residents seek the best possible training tools. The importance of simulation as a training tool to advance technical surgical skills has been extensively emphasized.\(^2\)\(^-\)\(^4\) By practicing on simulation models, residents can train without putting real patients at risk, and the advantages of such models can be leveraged for both initial and continuing neurosurgical training. An added benefit of simulation is that it allows practice after hours, which has been shown to be beneficial to the training of surgical residents.\(^2\)\(^,\)\(^3\)

Aneurysm clipping surgery is an extremely complex surgical procedure with a long learning curve. Several aneurysm models have been described for training and rehearsal of open aneurysm surgery.\(^5\)\(^-\)\(^10\) Our study aims to add to this body of work by providing a validation of a new model for surgical training purposes.

In a previous study, we described construction of an aneurysm model using a human placenta and various possible vascular procedures that can be performed on the model.\(^11\) The goal of this study was to assess the usefulness of a human placenta aneurysm model for training in aneurysm clipping skills by evaluating the internal consistency and face, content, and construct validity of the model. Additionally, the current study demonstrates the validity of a newly developed assessment scale, the Objective Structured Assessment of Aneurysm Clipping Skills (OSAACS).
METHODS

Study Approval and Placenta Collection

Approval for this study was obtained from the Institutional Review Board (IRB) for Human Subjects at Barrow Neurological Institute (St. Joseph’s Hospital and Medical Center, Phoenix, Arizona), and from the ethics committee of the Irkutsk Scientific Center of Surgery and Traumatology (Irkutsk, Russia). The research coordinator obtained informed consent from appropriate obstetrics patients to use their donated placentas for surgical practice. Prior to delivery and placenta donation, patients were screened for infectious diseases. Consent forms were approved by the IRB. Donated placentas were refrigerated before being used in practice surgery, as previously described.11

Simulation Model of Brain Aneurysm in Human Placentas

Simulated brain aneurysms were created on appropriate placental vasculature by inflating intravascular balloons and using suture ligation where necessary, according to a previously described protocol.11 Descriptions of the placental vasculature have been published previously.11-13 The models were set up in a training room in the neurosurgery laboratory on a flat, rigid surface under the operative microscope with the neurosurgery aneurysm clipping instrument set. Arterial blood flow was simulated by using pressurized infusion lines filled with a red-colored solution.

Procedures Performed

Each trainee was instructed to perform two exercises: 1) sylvian fissure–like dissection of the placental surface and 2) direct clipping of the simulated aneurysm (Fig. 1).
tasks for these procedures involved various skills similar to those used during cerebral aneurysm clipping procedures in human patients. Other tasks that were also tested included microsurgical sharp and blunt dissection of thin membranes; obtainment of proximal and distal vascular control; dissection of the aneurysmal neck and dome; and clip application, adjustment, and removal. In the event of an injured vessel or aneurysm, we also assessed bleeding control skills with suction, coagulation, and/or temporary clip application.

Participants were observed and scored on the OSAACS scale (Supplement 1) during training with the placenta aneurysm model by 2 independent raters. Upon completion of their training on the model and the collection of OSAACS scores, participants completed the Aneurysm Clipping Participant Survey (Supplement 2).

**Study Participants**

Medical students, residents, and staff neurosurgeons participated in this study. Three groups of 10 participants each were classified by skill level: group 1 included those with “low experience” (medical students and postgraduate year [PGY] 1-3), group 2 included those with “intermediate experience” (PGY 4-7 and fellows), and group 3 included attending neurosurgeons with vascular neurosurgery practice. Raters were fully trained neurosurgical fellows acquainted with the aneurysm clipping techniques.

**Assessment Tools**

Two assessment tools were used in this study. The OSAACS tool was used to objectively assess construct validity (Supplement 1). The Aneurysm Clipping Participant Survey (Supplement 2) was used to subjectively assess face and content validities of the model.
The OSAACS, which is an adaptation of the Objective Structured Assessment of Technical Skills tool, aims to measure the specific operative nuances of aneurysm clipping surgery. The tool is completed by a proctor who assesses the performance of the participants as they complete training tasks. The OSAACS score chart includes 9 dimensions that relate to specific aspects of surgical performance within an aneurysm surgery. Each dimension is scored on a scale from 1 to 5 (lowest to highest value), with a total possible score of 45 points.

The Aneurysm Clipping Participant Survey is a 10-item measure completed by participants after the training session. It consists of demographic information related to age, training, and neurosurgical experience, and 6 items regarding the placenta aneurysm model. Specifically, it asks questions about how well the model replicates actual brain surgery, how difficult it is to complete surgery tasks with the model compared to actual brain surgery, and how the participant judges the model in terms of its potential for effective training for actual open neurovascular procedures. Each question was scored on a 20-point scale with ordinal descriptors at both extremes and at the middle of the scale.

**Statistical Analysis**

Statistical analysis was performed using Statistica (Dell Inc.; Round Rock, Texas) and SPSS Statistics for Windows (IBM Corp.; Armonk, New York). P<0.05 was chosen as the lower limit of statistical significance.

**OSAACS Instrument Validation**

Cronbach’s alpha was used as a measure of internal consistency of the OSAACS scale. Alphas greater than 0.70 were considered acceptable, and those greater than 0.80 were
considered good. The Spearman’s ρ correlation coefficient between the 2 raters for each of the 9 items and the total score was used to assess interrater reliability. Strength of correlation (and the kappa statistic, see below) was determined by the following scale: <0 = poor agreement; 0–0.20 = slight agreement; 0.21–0.40 = fair agreement; 0.41–0.60 = moderate agreement; 0.61–0.80 = substantial agreement; and 0.81–1.00 = almost perfect agreement.15

The interrater agreement analysis between the standard and the observations was used to determine the reliability of the OSAACS tool. The known standard was the “group” variable. The observation variable was a new ordinal grouping measure derived from the total OSAACS score to assign participants to new groups (labeled as “1,” “2,” and “3”) on the basis of ranges of OSAACS scores. It was created by using a logistic regression based on total OSAACS scores and by using the Wilk’s lambda test in the discriminant analysis module of Statistica software (Table 1). Interrater agreement analysis comparing these groups was then completed using the attribute agreement analysis module of Statistica, on the basis of calculation of the Cohen kappa coefficient. The ultimate goal of the OSAACS instrument validation was to assess its predictive value in discriminating different levels of manual surgical skills.

**Face and Content Validity Calculations**

Face and content validities were obtained from the Aneurysm Clipping Participant Survey. Kruskal-Wallis and Mann-Whitney U tests were used to assess significance of differences among the 3 groups.
Construct Validity

Construct validity was obtained from the OSAACS scores. Analysis of each participant’s final OSAACS score was completed using the mean of their 2 OSAACS scores from the 2 raters. The Kruskal-Wallis test and the Mann-Whitney U test were used to assess significance of differences among the 3 groups.

RESULTS

General Group Demographics

Our cohort consisted of 30 medical professionals of varying experience, ranging from medical students to practicing neurosurgeons. The low-experience group was composed of 4 medical students and 6 neurosurgery residents (1 in PGY 1 and 5 in PGY 2). The mean age of the group was 27.4±2.7 years. The group averaged 1.2±1.1 years of microsurgical experience and 7.3±11.6 cases of aneurysm surgery assistance. The intermediate-experience group was composed of 4 neurosurgery fellows and 6 neurosurgery residents (1 in PGY 4, 4 in PGY 5, and 1 in PGY 6). The mean age of the group was 31.1±1.2 years. This group had a mean of 3.6±1.7 years of microsurgical experience and a mean of 27.6±54.0 cases of aneurysm surgery assistance. The attending group was composed of 10 attending neurosurgeons with a mean age of 39.6±6.2 years. This group had a mean of 12.1±5.5 years of microsurgical experience. The members of this group had an average number of 378±677 (range 10-1,800) cases as the primary surgeon.
**Internal Consistency**

Cronbach’s alpha is a measure of the extent to which a scale measures one underlying construct. It was used as a measurement of internal consistency of the 9-item OSAACS scale. Cronbach’s alpha was 0.963 for rater 1 and 0.961 for rater 2; these scores were well within the excellent range for both raters. Cronbach’s alpha, assuming the deletion of each item, is shown in Table 2. The deletion of any single item did not result in a higher alpha.

**Interrater Reliability**

Interrater reliability was used to assess the extent to which the 2 raters agreed. Interrater reliability of the OSAACS scale was extremely high for the total score, $\rho=0.98$ (95% CI 0.95-0.99). Interrater reliability at the item level for the 9 items ranged from 0.80 to 0.97 (Table 2).

**Discriminant Function Analysis**

Discriminant function analysis, as detailed in the methodology section, was used to create new observational groupings based on OSAACS score intervals. The total score interval from 0 to 28 was labeled as “1,” the interval from 29 to 39 was labeled as “2,” and the interval from 40 to 45 was labeled as “3.” The prediction of this model was found to be 85% accurate, indicating that the grouping into intervals was statistically and conceptually valid. The interrater agreement coefficient between the standard and the discriminant groupings was 90% (95% CI, 73-97%), and the Cohen kappa coefficient was $0.85\pm0.12$ (P<0.01). This interrater agreement could be interpreted as “almost perfect agreement.”
Face Validity (Aneurysm Clipping Participant Survey)

Of the 30 participants, 27 completed the Aneurysm Clipping Participant Survey: all 10 participants in the low-experience group, 7 participants in the intermediate-experience group, and all 10 participants in the attending group. Analysis for replication and difficulty ratings considered only the intermediate-experience and attending groups because the low-experience group lacked adequate experience with practical neurosurgery. Data from the low-experience group that were related to these questions were excluded from analysis.

In terms of how well the model replicated actual brain aneurysm surgery, the participants’ ratings for the model ranged from 10 to 20, using a scale where 0 = not at all and 20 = very well. More specifically, a large proportion of both groups (5 of 7 intermediate experience, 71%; 7 of 10 attending, 70%) rated replication between 13 and 20, meaning the model replicated real surgery more than “somewhat” and up to “very well.” The remaining 5 participants (29%) gave a replication score between 10 and 12, indicating that the model “somewhat” replicated real surgery. The difference between the intermediate-experience and attending groups was not statistically significant (P=0.54; Fig. 2).

In terms of difficulty of the tasks compared to real surgery, a large proportion of participants gave the model a rating between 5 and 15, on a scale of 0-20 where 10 was “same.” Specifically, 7 of 7 intermediate-experience (100%) and 7 of 10 attending (70%) participants indicated that the tasks with the model were roughly the same difficulty as tasks in real surgery. The difference in this rating between the intermediate experience and attending group was not statistically significant (P>0.99).

Analysis of the replication and difficulty ratings can be further restricted by only looking at the scores provided by the attending neurosurgeons. When only looking at these ratings for
both questions, the fact that a majority (7/10, 70% for each question) rated the models as “somewhat” to “very well” replicating surgery, with the difficulty being roughly the “same” as that of real surgery, indicated reasonably high face validity for the model.

As a final measure of face validity, participants were asked how successful they were in completing the surgical tasks of the model. For the low- and intermediate-experience groups, approximately half of each group (5 of 10 with low experience, 50%; 4 of 7 with intermediate experience, 57%) reported that they were more than “somewhat” to “completely” successful, which was indicated by a rating between 13 and 20, using a scale of 0 = complete failure and 20 = complete success. This proportion was higher for the attending neurosurgeons, with 9 of 10 (90%) rating themselves as more than “somewhat” to “completely” successful. However, Kruskal-Wallis testing showed an absence of differences among the 3 groups (P=0.56). Additional Mann-Whitney U test results also failed to show any differences when the groups were compared pairwise (low experience vs. intermediate experience, P=0.89; intermediate experience vs. attending, P=0.42; attending vs. low experience, P=0.39).

Content Validity (Aneurysm Clipping Participant Survey)

Content validity was assessed by asking participants if practice using the placenta model could help improve 1) microdissection technique, 2) clip applier handling and aneurysm clipping skills, and 3) surgical technique when working with real patients (Fig. 3).

The participants’ responses in regard to whether the model could improve microdissection technique and clip applier and aneurysm clipping skills were unanimously positive, with 27 of 27 participants (100%) answering both questions with scores between 15 and
20, on a scale with 0 = absolutely no and 20 = absolutely yes. There were no differences among the 3 groups for these first 2 questions (P=0.27 and P=0.35, respectively).

When participants were asked if the model would improve surgical technique when working with real patients, 25 of 27 participants (93%) gave a score between 15 and 20, which indicated a positive response. There were no differences among the 3 groups (P=0.58).

The results on these 3 measures signify high content validity of the model.

**Construct Validity (OSAACS Tool)**

Mean OSAACS scores were compared to assess for construct validity. The low-experience group (n=10) had a mean OSAACS score of 22.9±5.4, the mean of the intermediate-experience group (n=10) was 32.8±4.0, and the mean for the attending group (n=10) was 43.3±1.3. Mean differences between these groups were statistically significant (P<0.001) (Fig. 4). Additionally, pairwise comparisons with the Bonferroni correction revealed significant differences (P<0.001 for all 3 pairs).

**DISCUSSION**

Microsurgical training has been augmented with human placenta models since 1979.\(^1^6\) We recently reported on the use of a placenta model for aneurysm clipping training\(^1^1,\)\(^1^2\) as an important addition to the teaching tools for neurosurgeons in training. Human placenta models can have numerous advantages over other training models, including the absence of the ethical and logistic limitations associated with live biological models, such as maintaining and sacrificing laboratory animals for training purposes. In another study, we demonstrated face, content, and construct validity of a human placenta model for microvascular anastomosis
training. Our findings demonstrated that human placenta arteries and human cerebral arteries have similar segmental lengths, diameters, and wall thicknesses. One notable difference was that placental arteries contain fewer elastic fibers than non-diseased cerebral arteries. The current study builds on these past studies by taking the placenta aneurysm model we previously described and evaluating its effectiveness as a training tool using quantitative validity testing.

The placenta model was validated according to 3 dimensions: face, content, and construct validities. Face validity is a subjective measurement used to determine whether the model provides a realistic demonstration of real aneurysm clipping, in terms of replication of real conditions, similarity of difficulty compared to real conditions, and ability to be successful in completing training tasks. Content validity is a subjective measurement that relates to whether the model is able to teach participants about microdissection and the handling of clip appliers, while also improving the ability of participants to perform aneurysm clipping and to actually improve their microsurgical technique when working with real patients. Finally, construct validity is used to determine whether the model can distinguish between the proficiency levels of participants, as defined by their year in a neurosurgical residency or status as an attending neurosurgeon. Altogether, these 3 types of validity measurements (face, content, and construct) have been used in previous studies to determine the effectiveness of a placenta training model as a substrate for building bypass and brain tumor removal skills.

It was previously noted that assessment of the training system by untrained participants could not be a reliable measure of face validity. This is an intuitive point because less experienced participants would not be familiar enough with actual operative neurosurgical conditions to determine whether models were accurately replicating real conditions. Therefore, face validity of the current aneurysm model was determined by assessing the responses of the
attending neurosurgeon participants. These responses were generally high for the 2 questions about whether the model replicated actual neurosurgery and whether the neurosurgeons believed that they were successful with the model. At the same time, the attending neurosurgeon’s responses for the question about the difficulty of using the model compared to the difficulty of real neurosurgery indicated that the difficulties of the 2 procedures were similar. On the basis of the responses to these 3 questions, the placenta model in this study showed high face validity.

The human placenta aneurysm clipping model also showed high content validity due to a high frequency of positive responses to the 3 questions related to whether the participants believed the model was helping them improve certain neurosurgical skills.

Finally, the model showed a high construct validity with statistically significant differences in OSAACS scores between the 3 groups, which indicates that increased surgical proficiency in dissection and clipping of an aneurysm translates to better performance on the model. Furthermore, the construct validity findings were supported by high interrater reliability of the OSAACS scale.

The determination of whether a training program should have its residents practice aneurysm clip application skills on a placenta model instead of on other models requires consideration of multiple aspects of the models. In addition to comparing the human placenta model to human cerebral arteries, it is important to compare the model to alternative vasculature models, including virtual, silicone, and animal models. Compared to these models, the placenta aneurysm model avoids many obstacles, such as the artificial and expensive nature of the virtual models, the lack of physical similarity and tactile response of non-biological materials, and the ethical issues associated with distress in animals used for models. More in-depth discussions of other training models have been previously published elsewhere.
The current study demonstrated that a placenta aneurysm clipping model can be used as a realistic, appropriate, and convenient teaching tool. It also provides users with the ability to objectively differentiate between different proficiency levels of microsurgical aneurysm clipping skills.

**Limitations**

It should be noted that aside from evaluating measures of face, content, and construct validity, concurrent validity measures could also have been used to validate that this model is a useful training tool for new neurosurgeons. Concurrent validity would show that skills gained from training are directly related to improved performance in the operating room. A limitation of this study is that this variable could not be assessed because of patient safety concerns related to letting untrained participants perform neurosurgical tasks on actual patients.

Our findings may also be limited by our assumption that training with the model results in improved surgical performance. When the validity of a training model is formally assessed, one particular metric—the construct validity—should be measured to test whether the model provides a variety of tasks with a wide range of difficulty that allows performance to be measured and differentiated in novices and in experts. One might ask whether training on the model results in improved dexterity over time. Such improvement in performance on a model is usually assumed to be obvious and thus is not measured in this type of validation study. Although training on a model does not prove that it improves actual performance, this assumption can be drawn from the construct validity.

The responses of participants to the questions about the similarity of the model to the technique of aneurysm surgery (face and content validity), and about whether the model
presented relevant challenging tasks to the novices, also lead us to assume that novices who practice on the model would improve their manual performance and their OSAACS scores. Although the threshold number of training sessions needed to improve a score and the length of the learning curve have not been determined, a good surgeon would likely ultimately achieve an expert OSAACS score. The effects of training with this model should be confirmed by the participant’s performance in a real aneurysm surgery (a concurrent predictive validity measurement study). Such validity testing has been attempted by us and by other groups, and some results that have recently been published supplement the validity metrics of the placenta aneurysm model.24

Another limitation of the present study is that participants were not blinded to the raters, which potentially introduced some bias in the OSAACS scoring. Future studies might use video recording to overcome this limitation.

Limitations of this type of study include the subjective nature of face and content validities and the absence of a control group, which could potentially be used with a silicone aneurysm model. Therefore, it is difficult to be absolutely certain that the highly positive responses to the questions about the usefulness of the aneurysm model are actually related to the model or simply a reflection of the usefulness of any microsurgical activity. Furthermore, recorded responses on the questionnaire were not different among the study groups. Potential limitations of the placenta aneurysm model have been mentioned previously.11 These include the limited time to use a placenta before decay and increased risk of infection, although such risk can be mitigated by screening donors for infectious diseases. Additionally, a studied exercise of aneurysm dissection from the adherent membranes and aneurysm clip application provides comparable, but not perfect, real anatomical replication of the actual neurosurgical tasks. The
model described in this study most closely replicates the clipping of middle cerebral artery aneurysms located within the sylvian fissure, rather than internal carotid or posterior circulation aneurysms. Giant, rigid, thrombosed aneurysms and parent cerebral vessels with thick calcified walls are not well replicated by the human placenta model. At the same time, skull base approaches, clinoid process removal, and surgical vascular anatomy also cannot be simulated with the current model. The differences with actual aneurysm surgery were reduced by overlaying placentas to add extra soft tissue dissection, placing the model in a skull box model to simulate an actual craniotomy, and using pressurized flow to simulate actual blood pressure.

CONCLUSION

This study demonstrated face, content, and construct validity for an aneurysm clipping model constructed from human placenta, and it indicated that the model is a useful surgical skill training tool for residents and neurosurgeons in training. This model can supplement current neurosurgical training curricula. The developed OSAACS instrument can also be used to assess surgical skills related to aneurysm clipping.
REFERENCES


FIGURE LEGENDS

Figure 1. Intraoperative photos of (A) sharp dissection and (B) clip application exercises performed on the placenta aneurysm model. Used with permission from Barrow Neurological Institute, Phoenix, Arizona.

Figure 2. Responses regarding face validity measures from the 3 study groups. For the 3 questions, responses between groups did not differ significantly: \( P=0.59 \), \( P=0.85 \), and \( P=0.56 \), respectively. Used with permission from Barrow Neurological Institute, Phoenix, Arizona.

Figure 3. Responses regarding content validity measures from the 3 study groups. For the 3 questions, responses between groups did not differ significantly: \( P=0.27 \), \( P=0.35 \), and \( P=0.58 \), respectively. Asterisks and circles indicate outliers. Used with permission from Barrow Neurological Institute, Phoenix, Arizona.

Figure 4. Mean total OSAACS scores from the 3 study groups. The differences in mean scores between all 3 groups were statistically significant (\( P<0.001 \)). Used with permission from Barrow Neurological Institute, Phoenix, Arizona.
Table 1. Discriminant analysis: conversion of continuous values of total OSAACS scores into interval values.*

<table>
<thead>
<tr>
<th>Group</th>
<th>Total OSAACS score (mean ± SD)</th>
<th>Interval in the model</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low experience</td>
<td>22.9±5.3</td>
<td>0–28</td>
<td>Novice and beginning skills</td>
</tr>
<tr>
<td>Intermediate experience</td>
<td>32.8±4.4</td>
<td>29–39</td>
<td>Intermediate skills level</td>
</tr>
<tr>
<td>Attending</td>
<td>43.25±1.3</td>
<td>40–45</td>
<td>Advanced skills level</td>
</tr>
</tbody>
</table>

*Wilks’ lambda = 0.18. For the model, P<0.01.
Table 2. Interrater agreement and Cohen kappa coefficients for the scoring dimensions of the OSAACS scale.

<table>
<thead>
<tr>
<th>Scoring dimension</th>
<th>Cronbach’s alpha scale, if item deleted</th>
<th>Interrater reliability (95% CI)</th>
<th>Agreement n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1: Posture</td>
<td>0.963</td>
<td>0.92 (0.83-0.96)</td>
<td>25 (83.3)</td>
</tr>
<tr>
<td>Item 2: Microscope</td>
<td>0.961</td>
<td>0.90 (0.80-0.95)</td>
<td>27 (90.0)</td>
</tr>
<tr>
<td>Item 3: Knowledge of instruments</td>
<td>0.960</td>
<td>0.83 (0.68-0.92)</td>
<td>30 (100)</td>
</tr>
<tr>
<td>Item 4: Handling of instruments</td>
<td>0.959</td>
<td>0.80 (0.62-0.90)</td>
<td>17 (56.7)</td>
</tr>
<tr>
<td>Item 5: Time and motion</td>
<td>0.956</td>
<td>0.96 (0.92-0.98)</td>
<td>25 (83.3)</td>
</tr>
<tr>
<td>Item 6: Forward planning</td>
<td>0.957</td>
<td>0.91 (0.82-0.96)</td>
<td>23 (76.7)</td>
</tr>
<tr>
<td>Item 7: Clipping quality</td>
<td>0.958</td>
<td>0.97 (0.95-0.99)</td>
<td>27 (90.0)</td>
</tr>
<tr>
<td>Item 8: Respect for tissue</td>
<td>0.958</td>
<td>0.93 (0.86-0.97)</td>
<td>24 (80.0)</td>
</tr>
<tr>
<td>Item 9: Quality of dissection</td>
<td>0.956</td>
<td>0.95 (0.90-0.98)</td>
<td>23 (76.7)</td>
</tr>
</tbody>
</table>
Do you think practice on this model could help microdissection technique?

- **Attendings**
  - Thick Line = Median
  - P = 0.81

- **Intermediate Experience**
  - = 25% - 75%
  - P = 0.19

- **Low Experience**
  - P = 0.32

Do you think practice on this model could help to improve skills in handling clip applier and aneurysm clipping?

- **Attendings**
  - P = 0.32

- **Intermediate Experience**
  - P = 0.36

- **Low Experience**
  - P = 0.74

Do you think that practicing on this model will improve surgical technique when applied to patients?
How well do you think the model replicates actual brain aneurysm surgery?

How difficult was the task compared to real surgery?

How successful were you in accomplishing the task?