Face, Content, and Construct Validity of a Virtual Reality Otoscopy Simulator and Applicability to Medical Training

Tobias Albrecht, MD[1], Christoph Nikendei, MD[2], and Mark Praetorius, MD[3]

Abstract

Objective. Otologic diseases are common in all age groups and can significantly impair the function of this important sensory organ. To make a correct diagnosis, the correct handling of the otoscope and a correctly performed examination are essential. A virtual reality simulator could make it easier to teach this difficult-to-teach skill. The aim of this study was to assess the face, content, and construct validity of the novel virtual reality otoscopy simulator and the applicability to otologic training.

Study Design. Face and content validity was assessed with a questionnaire. Construct validity was assessed in a prospectively designed controlled trial.

Setting. Training for medical students at a tertiary referral center.

Method. The questionnaire used a 6-point Likert scale. The otoscopy was rated with a modified Objective Structured Assessment of Technical Skills. Time to complete the task and the percentage of the assessed eardrum surface were recorded.

Results. The realism of the simulator and the applicability to medical training were assessed across several items. The ratings suggested good face and content validity as well as usefulness and functionality of the simulator. The otolaryngologists significantly outperformed the student group in all categories measured ($P < .0001$), suggesting construct validity of the simulator.

Conclusion. In this study, we could demonstrate face, content, and construct validity for a novel high-fidelity virtual reality otoscopy simulator. The results encourage the use of the otoscopy simulator as a complementary tool to traditional teaching methods in a curriculum for medical students.

Keywords

simulation, medical education, high-fidelity simulator, virtual reality, otoscopy, validation study

Received April 30, 2021; accepted June 28, 2021.

Otoscopy is the medical term for the visual examination of the ear canal and eardrum. Proper handling of the otoscope and a correctly performed examination are essential skills for the diagnosis and treatment of otologic diseases, which are widespread in all age groups and can significantly impair the function of this important sensory organ. Acute otitis media is the most common reason why children in the United States are prescribed antibiotics.[1] Often a general practitioner or pediatrician is consulted for ear pain, who then has to assess the eardrum. Due to the anatomy, only part of the entire surface can be seen through the otoscope such that a systematic scanning is needed to assess the entire tympanic membrane. Although otoscopy thus represents a key competence for various medical specialties, pediatricians and general practitioners typically have a training deficit in ear assessment. In 1992, Fisher and Pfleiderer demonstrated that the otoscopic skills of general practitioners are comparable to those of fourth-year medical students. Moreover, Pichichero and Poole revealed that when assessing acute otitis media or otitis media with effusion, general practitioners and pediatricians had a diagnostic accuracy of 36% to 51%.[2]

To make the correct diagnosis, examiners need to interpret the findings correctly and, more important, be able to spot the pathologic findings. To achieve this, correct usage of the otoscope is key. Simulation-based learning provides a fruitful setting to enable important learning and offers a variety of opportunities to practice complex skills.[3,4] Otoscopy simulators to date use interchangeable ear pieces with different diseases or a small 2-dimensional screen, serving as the tympanic membrane, where physiologic and pathologic findings can be displayed.[5,6] Since most simulators do not allow

[1]Department of Otorhinolaryngology–Head and Neck Surgery, Medical Center–University of Heidelberg, Heidelberg, Germany
[2]Department for General Internal Medicine and Psychosomatics, Medical Center–University of Heidelberg, Heidelberg, Germany
[3]Department of Otorhinolaryngology–Head and Neck Surgery, Medical Center–University of Hamburg, Hamburg, Germany

Corresponding Author:
Tobias Albrecht, MD, Department of Otorhinolaryngology–Head and Neck Surgery, University Hospital Heidelberg, University of Heidelberg, Im Neuenheimer Feld 400, Heidelberg, 69120, Germany.
Email: Tobias.albrecht@med.uni-heidelberg.de

the tutor to see what the student sees and do not provide feedback regarding the correct insertion depth, otoscopy is a skill that is still difficult to teach.[7]

This study investigated a virtual reality otoscopy simulator that is equipped with a lifelike model ear and an otoscope handpiece. The handpiece allows the user to see, in real time, a detailed 3-dimensional model of the outer and middle ear when looking through the otoscope. Not only can regular anatomy be selected but also common pathologies. Furthermore, the simulator has a scale indicating the insertion depth of the otoscope, and virtual patients moan when the otoscope is inserted too deeply. A second screen allows the tutor to see the same image as the student. This feature and the feedback on the insertion depth of the otoscope could improve the way that this skill is taught. However, before implementation of the simulator into a training program, the subjective and objective validity has to be evaluated.

Therefore, the aim of this study was to assess the face, content, and construct validity of the otoscopy simulator for an otologic training.

Methods

Sample, Study Design, and Hardware Used

We divided this validation study into 2 parts. In part 1, 39 medical students (after attending their otolaryngology rotation at the University of Heidelberg) and 25 otolaryngologists were invited to participate in a cross-sectional study to assess the face and content validity as well as the applicability to medical training. For the second part of this study, 51 medical students (in their first week of their otolaryngology rotation) and 25 otolaryngologists were recruited to participate in a prospectively designed controlled trial to assess the construct validity.

In both parts of the study, the validity of a commercially available otoscopy simulator (Earsi Otoscope; VRMagic) was investigated.

Study Part 1: Face and Content Validity and Applicability to Medical Training

To assess the face and content validity, we developed a questionnaire based on the work of Wickens et al,[2] as a validated tool assessing otoscopy simulation was not available in the literature. The survey contained 2 assessment components. The first component consisted of 11 items and had a 6-point Likert scale (1, strongly disagree; 3, neutral; 6, strongly agree) to assess the face validity of the auricle, the appearance of the ear canal, and the eardrum in physiologic and pathologic conditions. Face validity describes the extent of a simulator's realism and appropriateness when compared with the actual task. The second component assessed the usability of the simulator and its applicability for training of medical students, otolaryngology residents, and nonotolaryngology residents and included questions to assess the content validity of the simulator. In this case, content validity is defined as the extent to which the content of a simulator represents the knowledge or skills that have to be acquired in the real environment based on detailed examination of the learning resources, tutorials, and tasks.[8]

To evaluate the content validity of the simulator, 5 items were scored with the same 6-point Likert scale regarding the assessment of the external auditory canal and tympanic membrane, the width of the pathologies, and the application as a training tool for novices.

The questionnaire on face validity was answered by 39 medical students (after finishing their training in otolaryngology), 12 otolaryngology residents, and 13 otolaryngology specialists. Only otolaryngologists were asked to determine the applicability to medical training and the content validity.

Study Part 2: Experimental Setup and Procedure for Construct Validity

Construct validity is defined as "a set of procedures for evaluating a testing instrument based on the degree to which the test items identify the quality, ability, or trait it was designed to measure." A common example is the ability of an assessment tool to differentiate between experts and novices performing a given task. To determine whether the training with the simulator actually captures aspects of the skills needed to perform a correct otoscopy, thus providing evidence for construct validity, 51 medical students in their first week of their otolaryngology rotation were recruited to form the novice group. The otolaryngology specialists and otolaryngology residents, forming the expert group, were asked individually whether they would take part in this study. The novice group received a standardized lesson on the anatomy of the ear and the otoscopy procedure by 1 selected instructor.

The lesson was based on the Peyton 4-step teaching method.[9] A standardized introduction to the simulator was presented to both groups, with short videos provided by the manufacturer. Both groups could ask questions after the technical introduction; after which, they were asked to perform an otoscopy to assess an eardrum. The eardrum examined during the otoscopy contained a large plaque as a pathology (case 1207: tympanosclerosis). After the otoscopy, participants marked the region and shape of the pathology. The procedure was rated by an independent rater using a modified OSATS checklist (Objective Structured Assessment of Technical Skills).[10,11] For each item, the OSATS checklist allowed the choice between *correctly done* and *not correctly done* and consisted of the following items: "Used the correct hand," "Otoscope was held correctly," "Otoscope was stabilized with a finger," "Pinna was pulled upwards and backwards," "Otoscope was inserted to the correct insertion depth," and "Otoscopy was performed atraumatically." Furthermore, the time to complete the task and the percentage of the assessed eardrum surface were recorded. The sketch of the eardrum pathology was rated with a 5-point Likert scale (1, no sketch of the pathologic finding possible; 5, correct region and size of the pathologic finding is marked). All sketches were rated by 2 blinded otolaryngologists. Both were not otherwise actively involved in this study and compared the sketches with the displayed pathology of the eardrum.

Ethical Approval

The study was conducted in accordance with the general terms and conditions and approval of the Heidelberg University Ethics Committee (reference S-514/2018). All participants were voluntarily recruited and informed about the aims of the study, and all provided informed consent prior to participation.

Statistics

For face and content validity data, mean values and standard deviations were calculated. Mean scores were calculated for each statement as rated by novices and experts. Face validity and general statement items were stratified by novice and expert groups, while content validity data were collected in the expert group only. Differences between novices and experts' ratings were analyzed with a nonparametric Mann-Whitney U test, with a P value $\leq.05$ to indicate significance. For the construct validity, the performances of the novice and expert groups were assessed and consisted of the following: the global scores of the modified OSATS, the time to complete the otoscopy, the percentage of the assessed eardrum, and the score achieved in the task of sketching the pathologic finding. Data were tested for normality with the D'Agostino-Pearson normality test. Since the data were not normally distributed, data were analyzed with a nonparametric Mann-Whitney U test with a P value $\leq.05$ to indicate significance. The interrater reliability for the rating of the sketches was assessed with the nonparametric Spearman correlation coefficient, since the data were not normally distributed. The assessment of the sketches was also analyzed with a Mann-Whitney U test.

Results

Demographic Characteristics

Study Part 1. The questionnaires on face validity were answered by 39 medical students: 19 men (49%) and 20 women (51%). Their average age was 25.3 years (range, 21-31). Nine students had completed vocational training in a medical profession (nurse) prior to their medical studies.

Twenty-five otolaryngologists, 10 women (40%) and 15 men (60%), answered the questionnaires on face validity and applicability to medical training and on the content validity. The average age was 36.5 years (range, 27-54).

Study Part 2. To determine construct validity, the expert group was formed by the same 25 otolaryngologists as in part 1 of the study. The nonexpert group consisted of 51 students who performed the otoscopy.

The average age of the nonexpert group was 24.94 years (range, 22-31). Twenty-one participants were men (41%) and 30 were women (59%). None of the participants had previous experience in otolaryngology.

Face Validity and Applicability to Medical Training

The realism of the simulator and the applicability to medical training were assessed across several items. The mean scores of the statements are shown in **Table 1**, depicted as total scores and subscores of the nonexpert and expert groups.

In summary, with the exception of the haptics of the pinna, the rating of nonexperts and experts for the face validity parameters was between 4.4 and 5.39. The most satisfying was the realism of the anatomic structures of the pinna (5.23), ear canal (5.02), and the eardrum (5.31), falling between the *agree* and *strongly agree* categories. This was considered an acceptable realistic representation of the relevant anatomy.

The realistic feeling of "the fabric of the pinna" was rated below average (mean score, 3.8), falling between *mostly agree* and *partly disagree*. The evaluation for the applicability to medical training parameters was between 4.51 and 5.68, which was considered a useful application of the simulator.

Content Validity

The degree to which the simulator addresses all subject material and curriculum requirements was specified with the same 6-point Likert scale across the following items: the model (1) provides a useful introduction to otoscopy, (2) provides adequate breadth of pathologies, (3) is useful for training the hand-eye coordination, (4) should be embedded into the curriculum of medical students, and (5) should be embedded into the training of otolaryngology residents. The detailed results are shown in **Table 2**. The mean scores across all 5 items measured >4, falling in the categories *mostly agree*, *agree*, and *strongly agree* and were thus considered acceptable across all items. In particular, the educational value of the simulator was highlighted by the experts when asked whether the model should be embedded in the curriculum of medical students (mean, 5.4).

Construct Validity

All participants completed the otoscopy and marked their pathologic finding on an illustration of a tympanic membrane. Otoscopy was rated with a modified OSATS. Time to complete the task and the percentage of the assessed eardrum were recorded. The expert group significantly outperformed the nonexpert group in all categories measured. The mean OSATS score was 2.31 (SD, 1.31) for the nonexpert group and 0.63 (SD, 0.83) for the expert group ($P < .0001$). In particular, the expert group scored significantly better in the following categories: "Otoscope was held correctly" ($P = .002$), "Otoscope was stabilized with a finger" ($P < .0001$), "Otoscope was inserted to the correct insertion depth" ($P < .0001$), and "Otoscopy was performed atraumatically" ($P = .0227$). Only 2 subcategories, "Used the correct hand" ($P > .99$) and "Pinna was pulled upwards and backwards" ($P = .093$), showed no significant difference between the groups.

The average time for the otoscopy was 56.01 seconds (range, 20-140) in the nonexpert group and 30.59 (range, 10.4-71.8) in the expert group ($P = .0001$). The Mann-Whitney U test showed that the task completion time in the nonexpert group was significantly higher than that of experts ($P < .0001$). The mean percentage of the examined eardrum surface was 66.43% (range, 0%-98%) in the nonexpert group and 83.84% (range, 53.77%-98.5%) in the expert group ($P = .0027$). In terms of the time required to complete the

Table 1. Items Used to Determine the Realism of the Simulator and the Applicability to Medical Training on a 6-Point Likert Scale.

Validation parameters: face validity	Total		Nonexperts		Experts		P value[a]
	Mean	SD	Mean	SD	Mean	SD	
Pinna							
The anatomic structures of the pinna appear realistic.	5.23	0.81	5.21	0.73	5.28	0.94	.441
The fabric of the pinna feels realistic.	3.89	1.32	3.95	1.34	3.8	1.32	.703
Ear canal							
The anatomic structures of the ear canal appear realistic.	5.02	0.77	5.13	0.70	4.84	0.85	.209
The difference between a healthy ear canal and an ear canal with pathology can be clearly identified.	5.08	0.9	5.10	0.94	5.04	0.84	.664
The ear canal with pathology looks realistic.	4.5	1.11	4.56	1.07	4.4	1.19	.515
The depth perception is realistic.	4.67	1.20	4.74	1.14	4.56	1.29	.643
Eardrum							
The anatomic structures of the eardrum appear realistic.	5.31	0.64	5.39	0.63	5.20	0.65	.258
The difference between a healthy eardrum and an eardrum with pathology can be clearly identified.	5.31	0.79	5.31	0.77	5.32	0.85	.830
The eardrum with pathology looks realistic.	4.66	1.12	4.51	1.30	4.88	0.73	.274
Applicability to medical training							
The model is useful . . .							
for teaching anatomy.					5.28	0.68	
for the training of medical students.					5.68	0.56	
for training of otolaryngology residents.					4.72	1.28	
for training of nonotolaryngology residents.					5.0	1.08	
the model is user-friendly.					5.32	0.63	

[a]Mann-Whitney *U* test.

Table 2. Items to Determine the Degree to Which the Simulator Addresses All Subject Material and Curriculum Requirements on a 6-Point Likert Scale.

Validation parameters: content validity	Experts	
	Mean	SD
The model . . .		
provides a useful introduction to otoscopy.	4.74	0.96
provides adequate breadth of pathologies.	5.16	0.89
is useful for training the hand-eye coordination.	5.0	0.82
should be embedded into the curriculum of medical students.	5.4	0.87
should be embedded into the training of otolaryngology residents.	4.48	1.41

otoscopy in relation to the examined area of the eardrum, the nonexpert group examined an eardrum area of 1.38% per second (SD, 0.75%) and the expert group, 3.42% per second (SD, 1.63%; $P < .0001$). The items assessed during the OSTATS are summarized in **Table 3**. Interrater reliability for the rating of the sketches was assessed with the nonparametric Spearman correlation coefficient and had an excellent correlation (Spearman $r = 0.9162$; 95% CI, 0.8689-0.9469). The mean score of the nonexpert group was 2.775, whereas the expert group achieved a mean score of 3.84 ($P = .0012$).

Discussion

The primary objective of this study was to establish face, content, and construct validity as well as to evaluate the applicability to medical training of a novel virtual reality otoscopy simulator. The ability to perform an accurate otoscopy is the crucial step to make an accurate diagnosis. Several publications reported alarmingly low diagnostic accuracy for otologic diseases among primary care physicians and pediatricians, suggesting the need to improve otoscopy education for medical students.[12-15]

Table 3. Items Assessed During the OSATS.

Item assessed during the OSATS	Nonexperts		Experts		P value [a]
	Mean	95% CI	Mean	95% CI	
Used the correct hand	0.20	0.08 to 0.31	0.22	0.04 to 0.40	>.99
Otoscope was held correctly	0.51	0.37 to 0.65	0.13	−0.02 to 0.28	.002
Otoscope was stabilized with a finger	0.57	0.43 to 0.71	0	0	<.0001
Pinna was pulled upwards and backwards	0	0	0.09	−0.03 to 0.22	.093
Otoscope was inserted to the correct insertion depth	0.69	0.55 to 0.82	0.087	−0.03 to 0.21	<.0001
Otoscopy was performed atraumatically	0.35	0.22 to 0.49	0.09	−0.03 to 0.21	.0227
Time to complete the task (in seconds)	56.01	47.84 to 63.13	30.59	24.11 to 37.07	<.0001
Percentage of the assessed eardrum surface	66.43	58.98 to 73.88	83.84	77.83 to 89.86	.0027

Abbreviation: OSATS, Objective Structured Assessment of Technical Skills.
[a] Mann-Whitney U test.

While high-fidelity and virtual reality simulation has a long history in other training programs in domains such as military and aviation, its appearance in the medical profession is more recent and often limited to surgical procedures.[16] Simulators offer several important advantages over didactic teaching and learning by performing procedures on patients. Simulator-based training provides a risk-free and error-forgiving environment. It allows the procedure to be performed many times in a compressed period without the trainee having to consider the well-being of patients. Feedback mechanisms allow error detection and testing of corrective strategies.[17,18] Simulators have been shown to prevent harm and discomfort to patients and allow for shortened learning curves.[19]

The commercially available virtual reality simulator offers trainees a variety of physiologic and pathologic ear findings. Moreover, trainees receive immediate and objective feedback after each examination, such as pain caused, area of tympanic membrane examined, and correct and missed findings. However, data on face, content, and construct validity were missing. The results of the face and content validity questionnaires and the responses on applicability to medical training proof the validity, usefulness, and functionality of the simulator. Most important, all otolaryngologists agreed that the trainer is useful for the training of medical students.

The perceived degree of realism of the feeling of the pinna was reduced in both groups. The entire outer ear is made of the same silicone, which feels somewhat rigid and offers no noticeable difference between the earlobe and the cartilaginous auricle. Since the difference between these subunits of the outer ear is of marginal importance for otoscopy, more attention was paid to the anatomy of the pinna. Here, the degree of realism was found to be very satisfactory. However, since the pinna is an exchangeable part, the realistic feeling could be a possible target for improvement. Apart from that, the realism of a simulator and its impact on learning are discussed controversially in the literature. Although the educational value of simulator-based training was found to be valuable, the influence of realism is discussed controversially, with studies reporting better, worse, or similar outcomes after training on a low-fidelity simulator as compared with a high-fidelity simulator.[20-25]

In addition to face validity, construct validity is regarded as one of the most important aspects of simulator evaluation. Construct validity determines whether the device simulates the given task as in real life and thus can differentiate between a novice and an expert.[26] By rating all steps required to correctly perform an otoscopy, we were able to convincingly demonstrate construct validity for the simulator.

Limitations

With providing evidence for face, content, and construct validity of the simulator, there is room for improvement. The simulator was able to differentiate between experts and non-experts. Whether it can also differentiate residents with different training levels was not tested and will be examined in future studies. Even though a simulator proves to be valid, conclusions cannot be drawn about knowledge and skill transfer and thus improved performance on real patients. The simulator comes with a variety of cases representing physiologic and pathologic findings. The expert group rated the item "The model provides adequate breadth of pathologies" very high (>5). Von Buchwald et al were also able to demonstrate content validity evidence for the pathologic cases provided by the simulator.[27] With a validated simulator to learn the skill and a validated selection of cases, it would be interesting to implement the simulation-based training of otoscopy in future curricula for undergraduate and postgraduate training.

Conclusion

In this study, we demonstrated face, content, and construct validity for a virtual reality otoscopy simulator. Our findings are encouraging to implement the otoscopy simulator into a curriculum for medical students as a complementary tool to traditional teaching methods.

Author Contributions

Tobias Albrecht, designed the study and acquired the data; did the data analysis, drafted the manuscript; acknowledged the drafting of the article and approved the version to be published; agrees to be

References

1. Rovers MM, Schilder AG, Zielhuis GA, Rosenfeld RM. Otitis media. *Lancet*. 2004;363(9407):465-473. doi:10.1016/S0140-6736(04)15495-0

2. Wickens B, Lewis J, Morris DP, Husein M, Ladak HM, Agrawal SK. Face and content validity of a novel, web-based otoscopy simulator for medical education. *J Otolaryngol Head Neck Surg*. 2015;44:7. doi:10.1186/s40463-015-0060-z

3. Barry Issenberg S, Scalese RJ. Best evidence on high-fidelity simulation: what clinical teachers need to know. *The Clinical Teacher*. 2007;4(2):73-77. doi:10.1111/j.1743-498X.2007.001 61.x

4. Chernikova O, Heitzmann N, Stadler M, Holzberger D, Seidel T, Fischer F. Simulation-based learning in higher education: a meta-analysis. *Review of Educational Research*. 2020;90(4): 499-541. doi:10.3102/0034654320933544

5. Davies J, Djelic L, Campisi P, Forte V, Chiodo A. Otoscopy simulation training in a classroom setting: a novel approach to teaching otoscopy to medical students. *Laryngoscope*. 2014; 124(11):2594-2597. doi:10.1002/lary.24682

6. Morris E, Kesser BW, Peirce-Cottler S, Keeley M. Development and validation of a novel ear simulator to teach pneumatic otoscopy. *Simul Healthc*. 2012;7(1):22-26. doi:10.1097/SIH.0b0 13e31822eac39

7. Polk ML, Lailach S, Kemper M, Bendas A, Zahnert T, Neudert M. Lernkurve der HNO-Spiegeluntersuchung: Zielgerichtete Lehrveranstaltungsplanung zu einer psychomotorischen Fertigkeit. Learning curve of the ENT mirror examination: goal-oriented curriculum planning for psychomotor learning. *HNO*. 2020;68(4):238-247. doi:10.1007/s00106-020-00829-w

8. Alsalamah A, Campo R, Tanos V, et al. Face and content validity of the virtual reality simulator ScanTrainer. *Gynecol Surg*. 2017;14(1):18. doi:10.1186/s10397-017-1020-6

9. Krautter M, Dittrich R, Safi A, et al. Peyton's four-step approach: differential effects of single instructional steps on procedural and memory performance—a clarification study. *Adv Med Educ Pract*. 2015;6:399-406. doi:10.2147/AMEP.S81923

10. Chipman JG, Schmitz CC. Using objective structured assessment of technical skills to evaluate a basic skills simulation curriculum for first-year surgical residents. *J Am Coll Surg*. 2009; 209(3):364-370, e2. doi:10.1016/j.jamcollsurg.2009.05.005

11. van Hove PD, Tuijthof GJ, Verdaasdonk EG, Stassen LP, Dankelman J. Objective assessment of technical surgical skills. *Br J Surg*. 2010;97(7):972-987. doi:10.1002/bjs.7115

12. Pichichero ME. Diagnostic accuracy, tympanocentesis training performance, and antibiotic selection by pediatric residents in management of otitis media. *Pediatrics*. 2002;110(6):1064-1070. doi:10.1542/peds.110.6.1064

13. Pichichero ME. Diagnostic accuracy of otitis media and tympanocentesis skills assessment among pediatricians. *Eur J Clin Microbiol Infect Dis*. 2003;22(9):519-524. doi:10.1007/s10096-003-0981-8

14. Pichichero ME, Poole MD. Comparison of performance by otolaryngologists, pediatricians, and general practioners on an otoendoscopic diagnostic video examination. *Int J Pediatr Otorhinolaryngol*. 2005;69(3):361-366. doi:10.1016/j.ijporl.2004.10 .013

15. Pichichero ME, Poole MD. Assessing diagnostic accuracy and tympanocentesis skills in the management of otitis media. *Arch Pediatr Adolesc Med*. 2001;155(10):1137-1142. doi:10.1001/ archpedi.155.10.1137

16. Satava RM. Historical review of surgical simulation—a personal perspective. *World J Surg*. 2008;32(2):141-148. doi:10.1007/ s00268-007-9374-y

17. Fried MP, Uribe JI, Sadoughi B. The role of virtual reality in surgical training in otorhinolaryngology. *Curr Opin Otolaryngol Head Neck Surg*. 2007;15(3):163-169. doi:10.1097/MOO.0b 013e32814b0802

18. Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. *Med Teach*. 2005;27(1):10-28. doi:10.1080/01421590500 046924

19. Schout BM, Hendrikx AJ, Scheele F, Bemelmans BL, Scherpbier AJ. Validation and implementation of surgical simulators: a critical review of present, past, and future. *Surg Endosc*. 2010; 24(3):536-546. doi:10.1007/s00464-009-0634-9

20. Massoth C, Roder H, Ohlenburg H, et al. High-fidelity is not superior to low-fidelity simulation but leads to overconfidence in medical students. *BMC Med Educ*. 2019;19(1):29. doi:10. 1186/s12909-019-1464-7

21. Roberts F, Cooper K. Effectiveness of high fidelity simulation versus low fidelity simulation on practical/clinical skill development in pre-registration physiotherapy students: a systematic review. *JBI Database System Rev Implement Rep*. 2019;17(6): 1229-1255. doi:10.11124/JBISRIR-2017-003931

22. Lefor AK, Harada K, Kawahira H, Mitsuishi M. The effect of simulator fidelity on procedure skill training: a literature review. *Int J Med Educ*. 2020;11:97-106. doi:10.5116/ijme.5ea6.ae73

23. Scerbo MW, Schmidt EA, Bliss JP. Comparison of a virtual reality simulator and simulated limbs for phlebotomy training. *J Infus Nurs*. 2006;29(4):214-224. doi:10.1097/00129804-2006 07000-00008

24. Rodgers DL, Securro S Jr, Pauley RD. The effect of high-fidelity simulation on educational outcomes in an advanced cardiovascular life support course. *Simul Healthc*. 2009;4(4):200-206. doi:10.1097/SIH.0b013e3181b1b877

25. Crofts JF, Bartlett C, Ellis D, Hunt LP, Fox R, Draycott TJ. Training for shoulder dystocia: a trial of simulation using low-fidelity and high-fidelity mannequins. *Obstet Gynecol*. 2006; 108(6):1477-1485. doi:10.1097/01.AOG.0000246801.45977.c8

26. McDougall EM. Validation of surgical simulators. *J Endourol*. 2007;21(3):244-247. doi:10.1089/end.2007.9985

27. von Buchwald JH, Frendo M, Guldager MJ, Melchiors J, Andersen SAW. Content validity evidence for a simulation-based test of handheld otoscopy skills. *Eur Arch Otorhinolaryngol*. Published online September 9, 2020. doi:10.1007/s00405-020-06336-6