OTOLOGY

Content validity evidence for a simulation-based test of handheld otoscopy skills

Josefine Hastrup von Buchwald[1,2] · Martin Frendø[1,2] · Mads J. Guldager[1] · Jacob Melchiors[1,2] ·
Steven Arild Wuyts Andersen[1,2,3]

Abstract

Purpose At graduation from medical school, competency in otoscopy is often insufficient. Simulation-based training can be used to improve technical skills, but the suitability of the training model and assessment must be supported by validity evidence. The purpose of this study was to collect content validity evidence for a simulation-based test of handheld otoscopy skills.

Methods First, a three-round Delphi study was conducted with a panel of nine clinical teachers in otorhinolaryngology (ORL) to determine the content requirements in our educational context. Next, the authenticity of relevant cases in a commercially available technology-enhanced simulator (Earsi, VR Magic, Germany) was evaluated by specialists in ORL. Finally, an integrated course was developed for the simulator based on these results.

Results The Delphi study resulted in nine essential diagnoses of normal variations and pathologies that all junior doctors should be able to diagnose with a handheld otoscope. Twelve out of 15 tested simulator cases were correctly recognized by at least one ORL specialist. Fifteen cases from the simulator case library matched the essential diagnoses determined by the Delphi study and were integrated into the course.

Conclusion Content validity evidence for a simulation-based test of handheld otoscopy skills was collected. This informed a simulation-based course that can be used for undergraduate training. The course needs to be further investigated in relation to other aspects of validity and for future self-directed training.

Keywords Handheld otoscopy · Simulation-based training · Technical skills training · Evidence-based medical education · Otology

Introduction

Handheld otoscopy is performed on a daily basis by a wide range of healthcare professionals including general practitioners, nurse practitioners, paediatricians, otorhinolaryngologists, and others. Otoscopy skills are important for accurate diagnosis and treatment and have been identified as key to preventing inappropriate use of antibiotics [1]. Even though otoscopy is a very common diagnostic procedure, competency in otoscopy after pre-graduate training is often insufficient: In one study, otoscopy skills were ranked by junior doctors as the second lowest of skills they perceived to be competent in after finishing medical school [2]. Additionally, other studies have highlighted the perceived need for improvement in under- and post-graduate otoscopy training [3, 4]. If not performed properly, the correct diagnosis can be missed and the examination can be painful to the patient. For the junior doctor, the procedure can be further challenging because of variation in patient anatomy as well as the loss of depth perception with the single ocular of the handheld otoscope. Finally, the procedure is difficult for the clinical teacher to supervise because of the limited field of view. For example, it is hard to guarantee that the student has the

✉ Josefine Hastrup von Buchwald
  Fine_buch@hotmail.com

1   Department of Otorhinolaryngology-Head
    and Neck Surgery, Rigshospitalet, Blegdamsvej 9,
    2100 Copenhagen Ø, Denmark

2   Copenhagen Academy for Medical Education and Simulation
    (CAMES), Center for HR and Education, The Capital Region
    of Denmark, Copenhagen, Denmark

3   Department of Otolaryngology, Nationwide Children's
    Hospital, The Ohio State University, Columbus, OH, USA

proper visualization and uses a systematic approach to the examination unless a video otoscope is used. In addition, especially in a clinical/pre-graduate teaching environment, it can be a challenge to ensure that a wide range of normal variations and pathologies is offered to all students, in particular when the exposure is of short duration which is usually the case in pre-graduate teaching [3].

It is thought that simulation-based training can provide a standardized learning experience and introduce a range of normal and abnormal anatomies and findings. Current simulation-based models for handheld otoscopy training include mannequins/task trainers [5, 6], a web-based model [7, 8], a mobile otoscopy simulator [9], and sophisticated technology-enhanced models [3, 10–15]. In some setups, the human instructor can monitor the procedure on a second screen and provide feedback. In other models, feedback is automated and integrated directly into the simulator [16]. This allows for self-directed training [17] where the learner can practice at their own convenience without the presence of an instructor. If a mastery learning approach is used, the learner can practice repeatedly until achieving a predefined level of proficiency [18]. Simulation-based training of handheld otoscopy has been found to improve confidence [10, 11, 13] and diagnostic accuracy [3, 8, 12, 14]. Regardless of the training approach, evidence-based medical education requires validity evidence for both the training model and the assessment.

Validity is a term to describe the extent to which a test measures what it is intended to measure [19]. The classical concepts of different validity types (content, criterion, and construct [19, 20]) has now been abandoned in favor of unitary approaches such as, for example, Messick's framework, where five different sources of evidence contribute to the validity argument [20, 21]: content evidence, response process, internal structure, relation to other variables, and consequences [19–21]. Validity frameworks can help structure research in validity evidence and identify missing evidence [20]. High-quality validity evidence is important because without it, the appropriateness of the training model and the performance assessment cannot be evaluated. However, only a few studies in surgical simulation systematically collect validity evidence using contemporary validity frameworks: A 2017 systematic review found that more than 9 out of 10 studies on surgical simulation published between 2008 and 2017 used either outdated or no validity frameworks [21].

Content evidence as described in Messick's framework is used to evaluate whether the test content reflects the construct it intends to measure [21] and aligns with the purpose of the assessment. A structured approach is needed so that all the items that represent the construct—for example handheld otoscopy skills—are considered in the test. In this study, we used the Delphi method to achieve consensus among content experts to establish content validity for the

cases in a simulation-based test of handheld otoscopy skills. To the best of our knowledge, no curriculum for simulation-based training of handheld otoscopy has yet been reported based on a systematic approach using a contemporary validity framework.

An essential first step is to collect content validity evidence, which we aim to do in this study. Our research questions were:

1. What is the content (i.e. normal and pathological cases) requirements for a simulation-based test of competency in handheld otoscopy?
2. Do the otoscopy simulator cases adequately represent the intended pathologies (i.e. can specialists in otorhinolaryngology correctly identify the cases)?
3. How can the content be integrated into a course in an otoscopy simulator?

Materials and methods

Part 1: Determining content

We conducted a Delphi study to identify which key normal variations and pathologies all medical doctors, regardless of specialty, should be able to recognize with a handheld otoscope at the time of graduation. The Delphi method is an iterative process to achieve consensus among content experts on a topic [22]. We recruited nine specialists in otorhinolaryngology (ORL) as content experts for our panel by e-mail invitation. Content experts were key opinion leaders teaching handheld otoscopy in under- and postgraduate medical training from all three postgraduate training regions of Denmark. For this research on content as well as technical skills related to handheld otoscopy, a three-round electronic survey using SurveyXact© (Rambøll, Aarhus, Denmark) was planned and conducted from March 2017 to March 2018. The latter survey resulted in the development of the Copenhagen Assessment Tool of Handheld Otoscopy skills (CATHOS) [23]. The panellists were allowed 4 weeks to complete the survey in each round before they were sent a reminder. Panellists' responses were blinded by (by MG) before being reviewed and aggregated by (by JM and SA) for the following round. Participants were asked to provide background information on age, sex, years of being a specialist, training region, and whether they worked in a private practice or hospital.

Round 1: Brainstorming phase

Panellists were asked to list (in free text) all the normal variations and pathologic conditions that, in their opinion, can be diagnosed with a handheld otoscope, irrespective of

training level or specialty. Panellists could also add comments on their suggestions. Duplicates and similar responses were merged and irrelevant responses (i.e. responses that were unrelated to the question) removed, resulting in a list of distinct diagnoses of normal variations and pathologies.

Round 2: Prioritization

Each panellist was presented with the list from Round 1 and asked to rank each item according to its relevance for a newly graduated junior doctor. Ranking was performed on a 1–5 Likert scale (1 = Irrelevant, 2 = Less relevant, 3 = Relevant, 4 = More relevant, 5 = Highly relevant). Panellists could also add comments on each pathology in free text. Items that received a ranking of > 3 by more than two-thirds of the panellists were selected by the study group for final consensus in Round 3. Further, tympanic membrane perforation was added to this list of diagnoses: the ability to identify perforation had made the consensus cut-off in the parallel Delphi study on technical skills, but was moved to the list of diagnoses since it represents a diagnosis rather than a technical skill. The study group also decided to include a specific follow-up question on myringosclerosis that had not made the cut-off because some of the panellists had considered it to belong under the "normal eardrum" diagnosis. Ultimately, myringosclerosis is a common variation that can cause a lot of referrals if not recognized as benign, and, therefore, the study group chose to ask the panel to consider adding it to the final content list.

Round 3: Consensus

The panellists were presented with the list of diagnoses from Round 2. To indicate if the list was comprehensive for the normal variations and pathologies that any junior doctor should be able to recognize, the panellists provided free-text responses. They also shared their thoughts on adding myringosclerosis to the content list.

Part 2: Pilot evaluation of simulator cases

To evaluate if the otoscopy simulator cases adequately represent the intended normal variations and pathologies, we recruited attendees at the annual meeting of the Danish Society of Otorhinolaryngology-Head and Neck Surgery (DSOHH) held the 12th and 13th of April 2018. The inclusion criteria were ORL specialization. Participants were asked to provide background information on their age, years of being a specialist, and whether they worked in private practice or at a hospital. Next, each participant performed three handheld otoscopies on The Earsi Otoscopy Simulator (VRmagic, Mannheim, Germany) with review of three random cases selected from 15 different cases from the simulator's case library. The fifteen cases represented both cases relevant to the nine diagnoses identified in the Delphi study (Part 1) and some more difficult pathologies relevant to the experienced participants, such as cholesteatoma and glomus tumor. Each case was presented to between one and six participants. After each case, participants were asked to provide a diagnosis in the free text without receiving any supplemental information such as patient history. The answers were anonymized and at a later point reviewed by the investigators. The diagnosis was considered correct if it matched the diagnosis which could be made without knowledge of the patient history. To further explore case difficulty, we used the background data to determine if experience or workplace were predictors of providing the correct diagnosis.

Part 3: Integrating content in an otoscopy simulator course

Informed by study part 1 and 2, we wanted to design the content of a simulation-based course in handheld otoscopy for medical students. The Earsi otoscopy simulator (VRmagic, Mannheim, Germany) was used as the simulation platform for the course. This simulator is a technology-enhanced simulator consisting of a rubber ear representation with an attached model of a handheld otoscope that tracks the position and projects the case into the otoscope view (Fig. 1). The learner can, therefore, examine the ear and see the pathology through the otoscope similarly to a traditional handheld otoscope. Simultaneously, a secondary touch-screen adjacent to the ear model displays the otoscope view allowing the instructor to follow the examination. The touch screen is also used to access the simulation software, select the user, choose the case, mode of feedback, present the case history etc. After each otoscopy examination, the software presents the learner with a structured questionnaire concerning findings and diagnosis. Together with data collected during the examination, the simulator provides automatic summative feedback including scores of instrument handling, which structures were observed, the examined area of the tympanic membrane, time, and whether the findings and diagnosis selected by the student in the post-case questionnaire were correct.

The simulator provides three types of built-in courses with increasing levels of difficulty: an introductory course concerning the healthy ear, a teaching course for self-directed learning, and an exam course concerning different pathologies. These courses cover almost all cases in the simulator case library. Based on the content requirements determined by our Delphi study along with the exploration of case difficulty, we designed a new simulation-based handheld otoscopy course and integrated this into the simulator in collaboration with the simulator developers.

Fig. 1 The Earsi simulator. **a** The learner examines the ear and sees the pathology through the otoscope similarly to a traditional handheld otoscope. Simultaneously, on a secondary touch-screen adjacent to the ear model, the otoscope view is displayed, allowing the instructor to follow the examination. The touch screen is also used to access the simulation software, select the user, choose the case, mode of feedback, present the case history, and answer follow-up questions concerning findings and diagnosis. The inserts exemplify the otoscope view as seen by the learner: **b** a case showing a tympanic membrane perforation and **c** a case showing acute otitis media
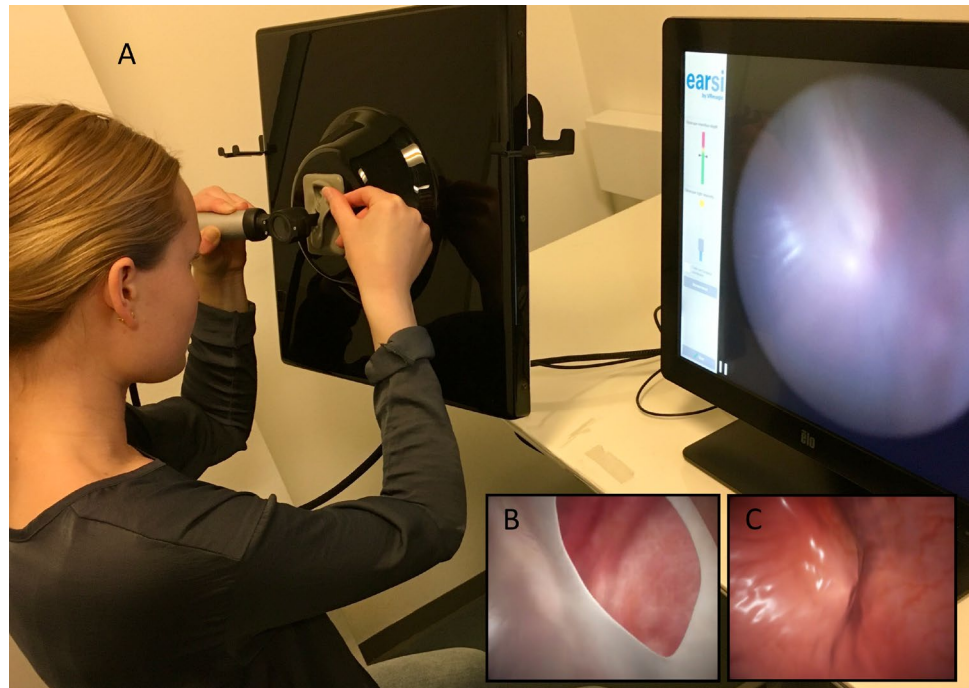
Statistics

Microsoft Excel® version 15.21.1 (Microsoft Excel, Microsoft, Redmond, Washington, U.S.) was used to organize data from study part 1 and 2. Statistical analyses for study part 2 were performed using Rstudio version 1.1.463 (Rstudio, Boston, U.S.). Chi-square test was used to compare diagnostic performance in relation to experience and workplace. To better illustrate simulation performance for ORL specialists with different experience levels, years of experience was dichotomized using 5 years as a cut-off (Table 3). *P*-values < 0.05 were considered statistically significant.

Results

Part 1: Determining content

All nine panellists completed the three planned rounds in the Delphi study. Their median age was 50 years (range 33–61); two were females and seven were males; three were in private practice and six were employed at teaching hospitals. Panellists were recruited from all three training regions of Denmark. Their median specialist experience was 9 years (range 1–25).

In Round 1, the panellists provided a total of 78 answers on normal variations and pathologies (flow chart, Fig. 2). After review by the investigators, seven unrelated answers were discarded and the remaining 71 answers aggregated resulting in 29 separate diagnoses. This list of diagnoses was sent out in Round 2 for ranking of relevance when training junior doctors. Seven made the pre-defined cut-off (assigned > 3 in relevance by more than two-thirds of the panellists). Additionally, the ability to "identify perforation" also made the cut-off in the parallel Delphi study on technical skills. The list of eight diagnoses was sent out for a confirmation in Round 3 along with the specific question on whether or not to include myringosclerosis. Myringosclerosis was added to the final list because a large majority (6 out of 9) agreed it was relevant. In the free text field, one of the panellists for example wrote "*I have had a patient referred suspected for cholesteatoma and it was myringosclerosis*" on the importance, whereas two panellists indicated that it should not be considered since the diagnosis is without clinical significance. The final list of the nine diagnoses of normal variations and pathologies is provided in Table 1.

Part 2: Pilot evaluation of simulator cases

Fourteen attendees at the annual meeting of the Danish Society for Otorhinolaryngology-Head & Neck Surgery met the inclusion criteria and volunteered for the case review. They represented a wide range of experience and both private practice and hospitals (Table 2). Of the 15 Earsi simulator cases that were included in the pilot, 12 represented the essential diagnoses determined by the Delphi study (Part 1) while three cases represented more challenging cases. Twelve out of the 15 cases were correctly identified by at least one specialist and considered passed. The specialists only made the correct diagnosis in 25 out of 42 attempts,
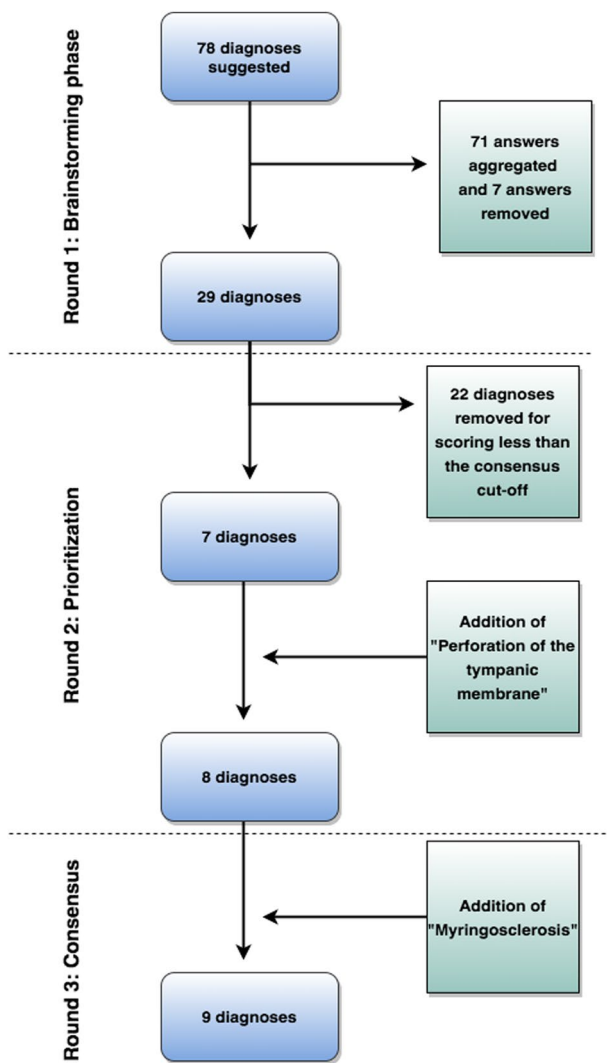
Fig. 2 Delphi study flowchart (study part 1). In blue boxes, the work of the Delphi panel is shown. In green boxes, the work of the study group is shown. Dotted lines mark the separate study rounds

Table 1 The final list of diagnoses of normal variations and pathologies that any junior doctor should be able to recognize using a hand-held otoscope (study part 1)

Normal ear canal and tympanic membrane
Cerumen
Foreign body
External otitis
Acute otitis media
Secretory otitis media
Myringitis bullosa
Myringosclerosis
Perforation of the tympanic membrane

Table 2 Participant demographics (study part 2)

	Median/count (range)
Age	45 (38–61)
Years as specialist	5.5 (1–25)
Experience	
<5 years	7
≥5 years	7
Workplace	
Private practice	6
Hospital	8

suggesting that in general, the cases in the simulator can be difficult to recognize without any context such as case history.

To explore case difficulty further, we analysed whether the distribution of correct/incorrect diagnoses differed based on experience (<5 vs. ≥ 5 years of experience) or workplace (private practice vs. hospital; Table 3). We found no statistically significant difference in the distribution of correct and incorrect diagnoses based on experience (χ^2 (1, $N=42$) $=0.099$, $p=0.75$) or workplace, (χ^2 (1, $N=42$) $=0.21$, $p=0.65$).

Part 3: Integrating content in an otoscopy simulator course

Fifteen cases from the Earsi simulator matched the essential diagnoses found through the Delphi study and were therefore eligible for the simulator course. In collaboration with the simulator developers, a course with all 15 cases was set up (Table 4). The course met the content requirement by including all essential diagnoses except for myringitis bullosa, which couldn't be included because there currently is no such case available in the simulator software. In the new course, the cases are repeated twice for training and listed in random order. The course is set up so that all cases, including the ones representing normal variations, have a case history as well as structured follow-up questions concerning both findings and diagnosis.

Twelve of the 15-course cases were tested in the pilot evaluation of the simulator cases (Part 2). In ten out of these 12 cases, the correct diagnosis was provided by at least one specialist. Despite not being correctly diagnosed by the specialists, the cases "otitis media with effusion" and "diffuse otitis externa" were included in the training program because they represent two of the essential diagnoses determined by the Delphi study (Part 1). Importantly, these cases will need to be further evaluated after implementation, but the cases might be less difficult when given the patient history and complaints provided in the simulator. Three additional cases were included in the course but were not pilot tested

Table 3 The distribution of correct/incorrect diagnoses based on experience and workplace (study part 2)

Experience	< 5 years of experience	≥ 5 years of experience	Total
Correct diagnoses	12 (29%)	13 (31%)	25 (60%)
Incorrect diagnoses	9 (21%)	8 (19%)	17 (40%)
Total	21 (50%)	21 (50%)	42 (100%)
Workplace	Private practice	Hospital	Total
Correct diagnoses	10 (24%)	15 (36%)	25 (60%)
Incorrect diagnoses	8 (19%)	9 (21%)	17 (40%)
Total	18 (43%)	24 (57%)	42 (100%)

Table 4 Content of the otoscopy simulator course (study part 3). For each case name, essential diagnosis/diagnoses, and result of the pilot evaluation are presented

Course cases	Representation of essential diagnoses	Pilot evaluation (study part 2)
Ear canal and tympanic membrane	Normal ear canal and tympanic membrane	Passed
Tympanic membrane in detail	Normal ear canal and tympanic membrane	Passed
Middle ear	Normal ear canal and tympanic membrane	Passed
Earache	Normal ear canal and tympanic membrane	Not tested
Cerumen impaction	Cerumen	Passed
Foreign body	Foreign body	Passed
Barotrauma (rupture)	Perforation of the tympanic membrane	Passed
Acute otitis media (initial stage)	Acute otitis media	Not tested
Acute otitis media (presuppuration)	Acute otitis media	Passed
Acute otitis media (suppuration)	Acute otitis media	Passed
Acute otitis media (resolution)	Acute otitis media and perforation of the tympanic membrane	Not tested
Otitis media with effusion	Secretory otitis media	Failed
Tympanosclerosis	Myringosclerosis	Passed
Chronic suppurative otitis media	Secretory otitis media and perforation of the tympanic membrane	Passed
Diffuse otitis externa	External otitis	Failed

because without case history they were very similar in findings to other cases (earache to normal anatomy, initial stage acute otitis media to normal anatomy with slightly reddening of the eardrum, and resolution of acute otitis media to perforation).

Discussion

In this study, content validity evidence for a simulation-based test of handheld otoscopy skills was collected. First, the content requirements were explored through a Delphi study (Part 1). The study resulted in nine essential diagnoses of normal variations and pathologies, which all junior doctors should be able to diagnose with a handheld otoscope. Second, the authenticity of a technology-enhanced otoscopy simulator's cases was tested in a pilot by specialists in ORL (Part 2). The ability to recognize the cases was surprisingly

low, and factors such as visual representation, lack of patient history, and technical differences between the simulator and real-life handheld otoscopy are potential explanations. Finally, the content requirements (i.e. relevant cases) were integrated in a course for basic training of handheld otoscopy (Part 3). The determined course content can be used in any type of training curriculum for handheld otoscopy, and we chose to integrate this into a commercially available technology-enhanced simulator.

Simulations-based training of handheld otoscopy has been investigated in numerous studies [3, 5, 7–14]. All these studies report favourable outcomes of simulation-based training of handheld otoscopy skills, however, only a few studies compare simulation-based training with other training modalities. Training on a technology-enhanced otoscopy simulator has been compared with both training using web-based modules as well as standard classroom instructions [12, 14]: even though all training modalities improved

diagnostic accuracy, clinical skills were most improved in the group that received simulation-based training. Although these results are encouraging for the use of simulation-based training of handheld otoscopy, validity evidence for the specific simulator has not yet been collected systematically using a contemporary validity framework. In general, structured evaluations of otoscopy simulators and simulation-based assessment of technical skills in handheld otoscopy are limited. In relation to content, one study investigated whether experts perceived a web-based otoscopy simulator to address all subject material and curriculum requirements [7]. In addition to some content validity, this study also investigated face validity—a concept that has been abandoned in modern validity frameworks because it is a subjective evaluation of appearance [21]. A systematic review found that > 40% of studies on simulation-based training of technical skills reported face validity as evidence of assessment validity [21]. This is problematic because it adds no actual evidence of the validity of a test, and consequently it is no longer considered relevant in modern medical education [19, 24].

A strength of our study is the systematic collection and evaluation of data contributing to content validity. The Delphi method used allows panellists to individually and anonymously contribute, eliminating the bias of following the majority of the most dominant/authoritarian panellist [25, 26]. However, there is also a risk that phrasing and selection by the researchers between rounds can influence the judgment of the respondents [22, 26]. We tried to minimize this by consensus on wording of questions in the research group as well as in the aggregation of responses. Next, the size of the panel needs consideration as there is no standard on how many experts to include in the panel [22, 26]. Recruiting too few panellists can make the results unrepresentative; conversely, too many can result in prolonged time between rounds. Therefore, the right balance between achieving saturation in responses and feasibility needs to be found: a panel size of 6–9 panellists has been recommended in medical educational research [26]. Although other healthcare professionals might be competent in handheld otoscopy, we chose to use only specialists in otorhinolaryngology because in our context they are in charge of teaching basic handheld otoscopy in the undergraduate medical curriculum and are very experienced with otoscopy. For other educational contexts, a more comprehensive Delphi panel including representation from multiple specialities that perform handheld otoscopy might be valuable. Specific Delphi studies should be performed to develop the curriculum for more advanced training, for different specialities, and other groups of practitioners.

There are several limitations to the pilot evaluation of the simulator cases with ORL specialists, such as the small sample size, uneven distribution of times each case was presented, and not presenting the case history for time reasons (during a national meeting). Consequently, it makes it difficult to firmly conclude whether the cases are adequate representations of the intended normal variations or pathologies, although it was surprisingly difficult for the ORL specialists to recognize the cases based on visual cues only. In the final course, which will need to be systematically evaluated, we chose to include all possible cases with diagnoses related to the list determined by the Delphi study (Part 1).

Further, we only used otorhinolaryngologists in our study because they teach otoscopy in our context. However, several other specialists including general practitioners and paediatricians see a lot of patients with ear complaints, which could warrant developing a curriculum relevant for these specialties (i.e. specific contexts). Given the high prevalence of acute otitis media and otitis media with effusion and the resulting high cost of prescribed antibiotics, improving otoscopy skills and increasing otoscopy diagnostic skills of health care practitioners could ultimately lower the use of antibiotics [5]. In addition to being better for the patients this would also be cost-beneficial for the society and help mitigate the global threat of antibiotic resistance. This further emphasises that high-quality and evidence-based training is imperative in health care professional education.

Conclusion

In this study, we collected content evidence for a simulation-based test of handheld otoscopy skills specific for undergraduate training using a commercially available technology-enhanced simulator. Content evidence is only one source of evidence in Messick's framework of validity. Therefore, an important next step is to systematically gather validity evidence for the response process, internal structure, relation to other variables, and consequences. This could, for example, be a collection of validity evidence for the simulator metrics (performance scores) and establishing a pass/fail standard for self-directed training in the simulator. In turn, this could inform the use of simulation-based training in future curricula for undergraduate and postgraduate training of handheld otoscopy skills.

Compliance with ethical standards

Conflicts of interest None.

Ethics approval The Regional Committee of Health Research Ethics of the Capital Region, Denmark, determined the study to be exempt (ref. 17006972).

Consent to participate Informed consent was obtained from all participants in the study.

Consent for publication Additional informed consent was obtained for publication of the article photograph.

References

1. Jones WS (2006) Video otoscopy: bringing otoscopy out of the "black box". Int J Pediatr Otorhinolaryngol 70:1875–1883. https://doi.org/10.1016/j.ijporl.2006.06.014
2. Ringsted CV, Pallisgaard J, Falck G (2002) Physicians' clinical skills after finishing internship. Ugeskr Læg 164:3211–3215
3. Oyewumi M, Brandt MG, Carrillo B et al (2016) Objective evaluation of otoscopy skills among family and community medicine, pediatric, and otolaryngology residents. J Surg Educ 73:129–135. https://doi.org/10.1016/j.jsurg.2015.07.011
4. Fisher EW, Pfleiderer AG (1992) Assessment of the otoscopic skills of general practitioners and medical students: is there room for improvement? Br J Gen Pract 42:65–67
5. Morris E, Kesser BW, Peirce-Cottler S, Keeley M (2012) Development and validation of a novel ear simulator to teach pneumatic otoscopy. Simul Healthc 7:22–26. https://doi.org/10.1097/SIH.0b013e31822eac39
6. Kyoto Kagakuku "Ear" Examination simulator 2. https://www.kyotokagaku.com/products/detail01/mw12.html. Accessed 19 May 2020
7. Wickens B, Lewis J, Morris DP et al (2015) Face and content validity of a novel, web-based otoscopy simulator for medical education. J Otolaryngol Head Neck Surg 44:7. https://doi.org/10.1186/s40463-015-0060-z
8. Stepniak C, Wickens B, Husein M et al (2017) Blinded randomized controlled study of a web-based otoscopy simulator in undergraduate medical education. Laryngoscope 127:1306–1311. https://doi.org/10.1002/lary.26246
9. Xu J, Campisi P, Forte V et al (2018) Effectiveness of discovery learning using a mobile otoscopy simulator on knowledge acquisition and retention in medical students: a randomized controlled trial. J Otolaryngol Head Neck Surg 47:1–8. https://doi.org/10.1186/s40463-018-0317-4
10. Lee DJ, Fu TS, Carrillo B et al (2015) Evaluation of an otoscopy simulator to teach otoscopy and normative anatomy to first year medical students. Laryngoscope 125:2159–2162. https://doi.org/10.1002/lary.25135
11. Davies J, Djelic L, Campisi P et al (2014) Otoscopy simulation training in a classroom setting: a novel approach to teaching otoscopy to medical students. Laryngoscope 124:2594–2597. https://doi.org/10.1002/lary.24682
12. Wu V, Sattar J, Cheon S, Beyea JA (2018) Ear disease knowledge and otoscopy skills transfer to real patients: a randomized controlled trial. J Surg Educ 75:1062–1069. https://doi.org/10.1016/j.jsurg.2017.12.011
13. Dell'Era V, Garzaro M, Carenzo L et al (2020) An innovative and safe way to train novice ear nose and throat residents through simulation: the simorl experience. Acta Otorhinolaryngol Ital 40:19–25. https://doi.org/10.14639/0392-100X-N0128
14. Wu V, Beyea JA (2017) Evaluation of a web-based module and an otoscopy simulator in teaching ear disease. Otolaryngol Head Neck Surg 156:272–277. https://doi.org/10.1177/0194599816677697
15. VRmagic Earsi Otoscope. https://www.vrmagic.com/medical-simulators/earsi-otoscope. Accessed 18 May 2020
16. Wiet GJ, Sørensen MS, Andersen SAW (2017) Otologic skills training. Otolaryngol Clin N Am 50:933–945. https://doi.org/10.1016/j.otc.2017.05.005
17. Saks K, Leijen Ä (2014) Distinguishing self-directed and self-regulated learning and measuring them in the E-learning context. Procedia Soc Behav Sci 112:190–198. https://doi.org/10.1016/j.sbspro.2014.01.1155
18. Yudkowsky R, Park YS, Lineberry M et al (2015) Setting mastery learning standards. Acad Med 90:1495–1500. https://doi.org/10.1097/ACM.0000000000000887
19. Cook DA, Beckman TJ (2006) Current concepts in validity and reliability for psychometric instruments: theory and application. Am J Med 119:166.e7–166.e16. https://doi.org/10.1016/j.amjmed.2005.10.036
20. Cook DA, Hatala R (2016) Validation of educational assessments: a primer for simulation and beyond. Adv Simul 1:31. https://doi.org/10.1186/s41077-016-0033-y
21. Borgersen NJ, Naur TMH, Sørensen SMD et al (2018) Gathering validity evidence for surgical simulation: a systematic review. Ann Surg 267:1063–1068. https://doi.org/10.1097/SLA.0000000000002652
22. Hsu CC, Sandford BA (2007) The Delphi technique: making sense of consensus. Pract Assess Res Eval 12:Article 10. https://doi.org/10.7275/pdz9-th90
23. Guldager MJ, Melchiors J, Andersen SAW (2020) Development and validation of an assessment tool for technical skills in hand-held otoscopy. Ann Otol Rhinol Laryngol 129:715–721. https://doi.org/10.1177/0003489420904734
24. Downing SM (2006) Face validity of assessments: Faith-based interpretations or evidence-based science? Med Educ 40:7–8. https://doi.org/10.1111/j.1365-2929.2005.02361.x
25. McMillan SS, King M, Tully MP (2016) How to use the nominal group and Delphi techniques. Int J Clin Pharm 38:655–662. https://doi.org/10.1007/s11096-016-0257-x
26. Waggoner J, Carline JD, Durning SJ (2016) Is there a consensus on consensus methodology? Descriptions and recommendations for future consensus research. Acad Med 91:663–668. https://doi.org/10.1097/ACM.0000000000001092