Original Article

Assessment of Otoscopy: How Does Observation Compare to a Review of Clinical Evidence

Simon Jonathan Davis, MD[1*]; Jon Viljar Norvik, MD[2]; Kristin Elisa Ruud Hansen[3]; Ingrid Vognild, MD[4]; Eirik Reierth, PhD[5]

[1] MBChB, OSCE director & general practitioner, UiT, Norway
[2] Department of Medical Biology, UiT, Norway
[3] OSCE administrator & civil engineer (biotechnology), UiT, Norway
[4] Radiologist in training, Department of Radiology, UNN hospital, Tromsø, Norway
[5] Senior Academic Librarian, University Library, UiT, Norway

Abstract

Background and Purpose: To investigate how much the method of observation agrees with a standardised review of evidence of clinical examination, for the assessment of clinical otoscopic competence.
Methods: 65 medical students took part in an Objective Structured Clinical Examination (OSCE) station using patients with real pathology. Examiners assessed otoscopic competency in tympanic membrane examination solely by distant observation. An external examiner later reviewed candidates' documented findings on a schematic drawing of the tympanic membranes. Observed agreement of the two methods and Cohen's kappa coefficient were calculated.
Results: Mean otoscopy scores for examiner 1 and examiner 2 were 67.7% and 29.4% respectively. There was a significant difference using the Mann-Whitney U-test. OSCE observation declared 47.7% of candidates (31/65) to be clinically competent. Drawing-based analysis however deemed only 4.6% (3/65) to have achieved this competency. This represented more than a ten-fold overestimation of clinical competency by OSCE assessment. Observed agreement between assessment methods was 59.6%. Cohen's kappa coefficient was 0.1.
Conclusions: OSCE observational assessment of otoscopic clinical competency correlates poorly with review of evidence from clinical examination. If evidence review is acceptable as a better marker for competency, observation should not to be used alone in OSCE assessment. Evidence review itself is vulnerable to candidate guesswork. OSCE could possibly explore candidate demonstration with explanation of findings, by use of digital otoscopy offering a shared view of the tympanic membranes, as an improved standard of clinical competency assessment.

Keywords: Assessment, OSCE, Competency, Otoscopy, Otology, Otoscopic, Tympanic

Journal of Medical Education Summer 2016; 15(3):171-7

Introduction

Research estimates that 1 in 20 patients in general practice need otoscopy (1). It is therefore an important clinical competency for doctors to obtain. However very little has been published about the assessment of this competency for medical students and which methodology is most reliable in the OSCE setting. The few studies regarding competency assessment for otoscopy of the tympanic membrane have focused on general practitioners and pediatricians, whose diagnostic accuracy has been reported to be less than 50% (2, 3).

Otoscopy of the tympanic membrane presents a special challenge in assessment of clinical competency. To assess competency properly, the examiner needs to see what the candidate is viewing. This is important in order to

∗**Corresponding author:** Simon Jonathan Davis, Department of General Practice, Institute of Community Medicine, UiT The Arctic University of Norway, N 9019 Tromsø, Norway. Email: simon.davis@uit.no

assess systematic examination and clinical interpretation of the tympanic membrane. Otoscopy does feature in OSCE (4, 5) assessment where examiners judge clinical competence via observation. Unless standardised pathology is involved and the otoscopic view shared, this assessment can only assess technical handling of the otoscope and the candidate professionalism. It is difficult to be certain how much assessment without evidence of the candidate's clinical examination can determine clinical competence.

It may seem self-evident that assessment of unseen examination cannot be a valid test. However we decided to quantify the gap between assessments that involve unseen and seen evidence of clinical examination. We investigated how much observation of candidates agrees with a review of drawn evidence of clinical examination. Drawn evidence was adopted as the best available standard for assessing clinical competency. In addition, we aimed to stimulate discussion about the inclusion of evidence of clinical examination in the assessment of clinical competency.

Methods

At UiT the Arctic University of Norway (UiT) basic otological clinical examination is taught in the 2nd year of the medical curriculum. Hand-held otoscopic technique and the practical application of the tympanic membrane quadrant system (6) is taught in an otology lecture. Following this a rotation of small groups of 4-5 students attend the otology clinic for practical teaching with junior otologists, using hand held otoscopes on each other. Approximately, 14 months after this training, an OSCE otology station was introduced as part of a 12 station OSCE exam. The tasks included tympanic membrane otoscopy and drawing of findings. It was decided to use this opportunity to test two methods of assessment in parallel: OSCE observation and the review of drawn evidence of clinical examination. Drawing-based analysis was perceived to be a better assessment of clinical competency as it involved evidence of clinical interpretation.

In 2015 eighty-eight 3rd year medical students participated in an eight minute OSCE Ear station. Hand-held otoscopy on one of three simulated patients with pathology was assessed as 'performed with purpose', or not. 'Performed with purpose' equated to otoscopic clinical competence. This judgement relied on distant observation of 1.5-2 metres (fig.1), without a shared view of the tympanic membrane and no access to the candidates schematic drawing (Fig. 2). Pathological findings were marked in relation to the pre drawn malleus as a reference point. These drawings were used as evidence of the candidate's clinical competency, assuming that the candidates had not guessed.

Drawing-based analysis was done later by the OSCE medical director (a general practitioner with 4 years of otology experience), who was not involved as an OSCE examiner. The quadrant system was used to systematically assess the drawings (fig. 3). Candidates were deemed clinically competent if these drawings were 'correct', ie. markings on both membranes were confined to the correct quadrants. 10 drawings were misplaced and 13 unmarked. Available results of both the OSCE and drawing assessments, from the remaining 65 candidates were compared. In addition, actual otoscopic clinical competency was attributed if a candidate passed both assessments, which in combination tested all elements of clinical competency (technical skill, clinical interpretation and professionalism).

The OSCE examiners were two non-otology doctors, with more than three years of medical experience. Prior to the OSCE they were trained by the OSCE medical director in how to teach otoscopy and also how to assess this as an OSCE examiner. This training consisted of a practical demonstration of otoscopy, followed by analysis of an otology OSCE station video and thereafter 'mock OSCE' station practice with the actual simulated patients used for this station. In addition the examiners completed their preparation by teaching on an 'introduction to otoscopy' course for 2nd year students, which had learning objectives aligned with the OSCE objectives. Teaching otoscopy gave the examiners a comprehensive perspective of otoscopy learning which they reported to help focus and enhance their ability to assess OSCE candidates.

The three simulated patients with pathology were recruited to make the clinical task realistic and meaningful for the candidate. Patients with tympanosclerosis were chosen because is pathology did not alter its appearance and helped standardise the test. All tympanic membranes were photographed using a rigid otoscope one week before the OSCE. Two patients (fig. 4A and 4B) had one lesion per membrane and the third patient (fig. 4C) had two lesions per membrane. All pathology existed in individual quadrants and was identifiable to the examiners using the hand-held otoscope. Patients were re-examined throughout the OSCE period to ensure that the pathology was still visible. Patient participation in the OSCE was not randomized, yet it was not predictable.

We developed a search strategy and a systematic literature search was performed in the following databases: Ovid MEDLINE(R) (In Process & Other Non-Indexed Citations, Ovid MEDLINE(R), Daily, Ovid MEDLINE(R) and Ovid OLDMEDLINE(R) 1946 to Present), Embase Classic (Embase 1947 to 2015 November 15) (Ovid). The search was performed in October 2015. The controlled vocabulary of Medical Subject Headings (MeSH) from Medline and the Emtree thesaurus from Embase, including sub headings, were used. The search fields; Title, abstract and keywords, were searched when applicable. Languages included Scandinavian and English. There were no restrictions regarding publication year for the searches

Statistical methodology

SPSS (7) was used to perform all statistical analysis. Comparison of means between OSCE examiner scores used the Mann-Whitney U-test. Agreement of OSCE and drawing-based analysis results used Cohen's kappa coefficient.

Ethical standards

Local university ethical approval and written consent for all photographs was obtained.

Results

OSCE categorised 47.7% (n=31/65) of candidates to have 'performed with purpose' and hence passed them as clinically competent. However drawing-based analysis deemed that only 4.6% (3/65) of all candidates were clinically competent, having failed 90.3% (n=28/31) of OSCE 'competent' candidates. OSCE had a 9% positive predictive value for drawing-based analysis competency, with a specificity of 54.8% and a 100% negative predictive value.

There was a 59.6% (n=37/65) observed agreement between OSCE and drawing-based analysis (table I). Cohen's kappa coefficient (table II) was calculated to be 0.10, indicating a very low level of agreement.

The two examiners were compared as individuals. Examiner 1 passed 67.7% (n=21/31) and examiner 29.4% (n=10/34) which represented a statistically significant difference in their assessments (p=0.00). The

Table 1. Agreement between OSCE & drawing-based analysis

Method of assessment (n=65)	OSCE competent (n=31)	OSCE not competent (n=34)	
Drawing based analysis competent (n=3)	3	0	Sensitivity = 100%
Drawing based analysis non competent (n=62)	28	34	Specificity = 54.8%
	Positive predictive value (OSCE) = 0.09	Negative predictive value (OSCE) = 1.0	Observed agreement = 0.596

Table 2. Cohen's kappa coefficient for OSCE vs. drawing-based analysis

	Observed agreement	Probability of random agreement	Cohen's kappa coefficient, K
OSCE vs. drawing-based analysis (n=65)	0.596	0.521	0.10 (p=0.06)

Table 3. Agreement between drawing-based analysis & actual competency

Method of assessment (n=65)	Drawing-based analysis competent (n=3)	Drawing-based analysis not competent (n=62)	
Actual competent (n=3)	3	0	Sensitivity = 100%
Not actually competent (n=62)	0	62	Specificity = 100%
	Positive predictive value (OSCE) = 1.0	Negative predictive value (OSCE) = 1.0	Observed agreement = 100%

observed agreement between OSCE and drawing-based analysis for examiner 1 was 38.7% (n=12/31) with a kappa value of 0.06. Examiner 2 observed agreement was 73.5% (n=25/34), kappa 0.14. Both kappa results were not significant.

Discussion

This study showed poor agreement between observation and drawing-based analysis as methods of assessment of clinical otoscopic competency. OSCE observation overestimated competency by ten-fold. OSCE was also a poor predictor of actual clinical competency (both assessments passed), with a 9% positive predictive value, compared to 100% by using drawing-based analysis alone (Table III). It is possible that drawing-based analysis was inaccurate. Candidates could have correctly guessed the quadrant pattern pathology, which would have overestimated competency. Yet the chance of this was 0.51%. It is more plausible that drawing-based analysis underestimated clinical competency. Of the candidates who failed evidence review, 56.5% (35/62) did so because their drawings were incomplete. This may have happened due to of a lack of time, difficulty transferring the mental images onto paper, unfamiliarity with the quadrant system or candidates forgot how to apply it. If these 35 drawings had been completed properly, then OSCE observation 'from a distance' may then have correlated better with evidence review. However we are not currently convinced that such correlation would have been adequate enough to support observation

Figure 1. OSCE examiner view.

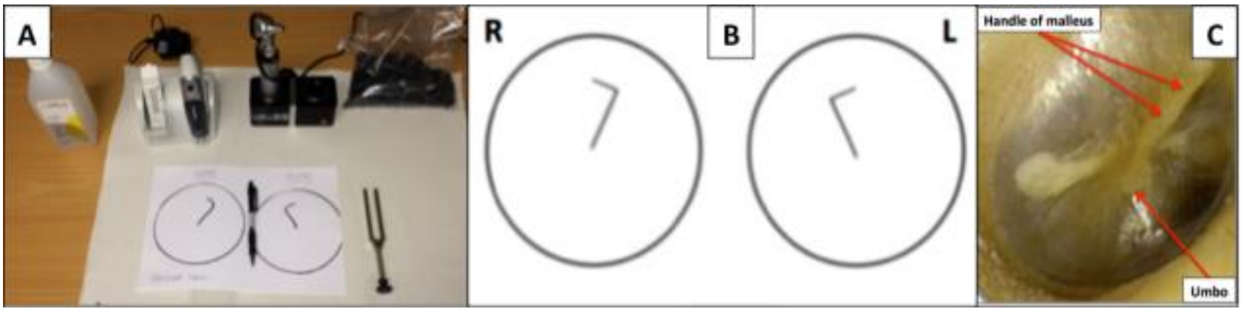

Figure 2. (A) Equipment, (B) Schematic drawing of tympanic membrane & (C) Anatomy.

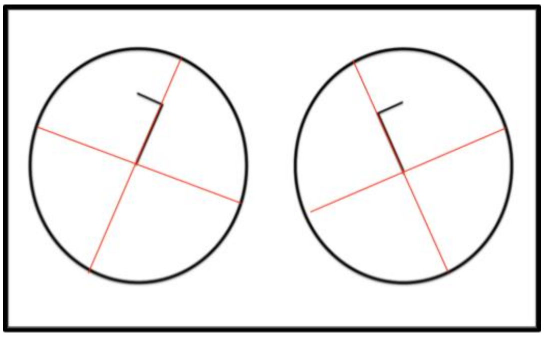

Figure 3. Drawing based analysis quadrant review.

as a single method of assessment. OSCE observation overestimated competency because examiners cannot assess what they cannot see.

We acknowledge that the correlation may be complicated by the fact that the assessment methods are actually measuring different competencies. OSCE measures technical and professional examination. Drawing-based analysis only assesses the correct documentation of findings, assumed to have been achieved by adequate technical skill and correct interpretation of pathology. This assumption was supported by the fact that all candidates who passed evidence review also passed the OSCE. Further research with a larger cohort could explore to what extent evidence review alone is a reliable marker for actual clinical competency. There was a significant difference between OSCE examiners. A comparison of the examiners showed that OSCE examiner 2's judgement correlated better with drawing-based analysis. Training and further experience should contribute to reducing this difference.

In this era of OSCE where otoscopy is a key skill (8, 9), very little has been published on the assessment of otoscopic clinical competency using evidence of examination. Smith (10) reported an OSCE station where candidates were assessed on their description of clinical findings. A Fischer and Pfleiderer (11) study introduced the practice of drawing of clinical findings, yet it is unclear how systematically the tympanic examination was performed and the drawing-based analysis was not standardised. Blomgren and Pitkäranta (12) studied the diagnostic agreement between a general practitioner and an otorhinolaryngologist, with the aid of photographic evidence of the tympanic membranes. Kaleida (13) used online video as a tool to assess otoscopic competency. From reviewing these studies we decided to incorporate drawing-based analysis and the quadrant system in order to standardise this process and adopt photographic evidence as proof of pathology.

We were not satisfied with how some authors defined competency without including the assessment of the technical skill of otoscopy or candidate professionalism. Competency is defined by the Oxford English dictionary as 'suitable, fit, appropriate, proper' (14). We elected to define clinical competency as the 'appropriate physical handling of the otoscope, interpretation of findings and professionalism'. To our knowledge, no one has attempted to assess all of these elements of otoscopic clinical competency in the OSCE setting.

Figure 4. Patients pathologies (A-C).

Recent articles have promoted technologies for otoscopy teaching. These are worth brief comment, as they might be considered for use in assessment. Wickens et al (15) describe a web-based simulator giving an artificial perception of the otoscope, but with little physical realism that is so vital in order to learn to use the otoscope effectively on real patients. Davies (16) reports large group simulation, which only technically prepared students for clinical otoscopy. We believe that training with otologists involving real patients with pathology still offers the best opportunity to teach and consolidate clinical competency realistically. Learning from clinical experts also boosts confidence amongst the medical students.

In the future to improve the assessment of competency we will include drawing-based analysis as part of the OSCE examiner's role and reassess how well this correlates well with external drawing-based analysis. We will introduce an additional task of demonstrating clinical findings to the examiner, via a digital probe connected to a display screen. This should best inform the examiner about purposeful examination and interpretation of the tympanic membrane. Lengthening the station time to 12 minutes should facilitate these changes. We plan to continue to use doctors who are teaching otoscopy as our OSCE examiners, as they can best recognise subtle aspects of clinical skill performance and as teachers they benefit most from the immediate feedback about their teaching. Assessment in learning has been described in the medical literature (17). More could be said about whether common forms of assessment, such as OSCE, truly test the competency in the role of the doctor. It is possible for students to appear competent using an otoscope and guess the findings. How assessment can do this is beyond the scope of this article, but in principle it should involve the candidate teaching the expert examiner. In this way candidates explain how and why pathology is present, which is difficult to fake in an exam situation.

We continue to challenge our doctor examiners to make clear what it is that they are actually aiming to assess and how the test does this beyond reasonable doubt. We remind the OSCE team that assessment means more than measuring technical ability. Assessment of clinical competency involves interpretation of clinical findings in the role of the doctor. The appropriate setting for assessment comprises of interpretation of clinical findings in a patient with pathology to a doctor examiner, who has years of otological experience and teaching of this skill. In this way can our 'test talk to the teaching' in a positive feedback loop. When we tell students that they have passed this design of skill test, it is then appropriate that they consider themselves to be clinically competent to use this skill as doctors. This should be the goal of all forms of skills assessment.

Conclusion

Our study shows that observational assessment of otoscopic clinical competency in OSCE correlates poorly with drawing-based analysis of clinical findings. In order to create a valid assessment, competency must be comprehensively assessed and evidence used to support examiner judgement. This could be achieved by an assessment which combines observation of professionalism and otoscopic technique with candidate-led demonstration and explanation of findings via digital otoscopy.

Conflict of Interest

I can declare that all authors have no known conflict of interest regarding this paper and its content. There are no known beneficiaries of this work that we know of. No one we work with is known to be a beneficiary of this study. The study was not funded.

References

1. Griffiths E. ENT problems in general practice. J R Soc. Med. 1979;10:740-2.
2. Pichichero ME. Diagnostic accuracy, tympanosclerosis training performance, and antibiotic selection by pediatric residents in management of otitis media. Pediatrics. 2002;110(2).
3. Pichichero & Poole. Comparison of performance by otolaryngologists, pediatricians, and general practitioners on an endoscopic diagnostic video examination. International Journal of Pediatric Otorhinolaryngology. 2005;69:361-6.
4. Harden R, Stevenson M, Downie W, Wilson G. Assessment of clinical competence using objective structured clinical examination. BMJ. 1975;1:447-51.
5. Drake Lee A B, Skinner D, Hawthorne M, Clarke R. A summative, Objective, Clinical Examination in ENT used to assess postgraduate doctors after one year of ENT training, as part of the Diploma of Otorhinolaryngology, Head and Neck Surgery. The Journal of Laryngology & Otology. 2009;123:1155-9.
6. Gray H. Anatomy of the Human Body. 20th Ed. Philadelphia: Lea & Febiger Publishing; 1918.
7. IBM Corp. Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.
8. Simpson JG. The Scottish doctor – learning outcomes for the medical undergraduate in Scotland: a foundation for competent and reflective practitioners. Med Teach. 2002;24:136-43.
9. General Medical Council. Tomorrow's Doctors: Recommendations on Undergraduate Medical Education. London: General Medical Council; 1993.
10. Smith M, Burton W, Mackay M. Development, impact, and measurement of enhanced physical diagnosis skills. Adv. in Health Sci Educ. 2009;14:547-56.
11. Fisher E, Pfleiderer A. Assessment of the otoscopic skills of general practitioners and medical students: is there room for improvement? British Journal of General Practice. 1992;42:65-7.
12. Blomgren K and Pitkäranta A. Is it possible to diagnose acute otitis media accurately in primary health care? Family Practice. 2003;20:524-7.
13. Kaleida P, Ploof DL, Kurs Lasky M, Shaikh N, Colborn DK, Haralam MA, et al. Mastering diagnostic skills: Enhancing Proficiency in Otitis Media, a model for diagnostic skills training. Pediatrics 2009;124(4):714-20.
14. “competent, adv.” The Oxford English Dictionary. 2nd ed. 1989. OED Online. Oxford University Press. Accessed 15 June 2016.
15. Wickens B, Lewis J, Morris D, Husain M, Ladak H, Agrawal S. Face and content validity of a novel, web based otoscopy simulator for medical education. Journal of Otolaryngology – Head & Neck Surgery. 2015;44:7.
16. 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):2594-7.
17. Wass V, Van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. Lancet 2001;357(9260):945-9.