

BEME PROTOCOL

Title: Objective Structured Clinical Examination (OSCE) as a formative assessment tool to motivate learners and improve performance of medical and dental undergraduates – A BEME systematic review.

Abstract:

Background: Students frequently have no opportunity to practice OSCE other than in the high-stakes examination itself (Young et al., 2014). This led to the use of OSCE formatively, where OSCE is primarily a learning tool. Its primary objective is providing feedback on students' performance to promote learning and enhance performance (Chisnall et al., 2015). Different types of formative OSCE with different type of feedback mechanisms have been identified in the literature. Based on the process, individual or group formative OSCE have been reported. Whereas depending on the type of feedback, direct immediate or video based formative feedback have been reported. With respect to who gives the feedback, formative OSCEs with peers, tutors and specialists have been reported (Rush et al., 2014; Perron et al., 2016). Although, literature suggests that formative OSCE promotes learning and improves performance, it is not clear how it does so. In addition, it is important to see if the different types of formative OSCE and feedback have a role in student learning. The effect of formative OSCE on behavioral change in students' daily clinical practice needs to be explored. Finally, the perceptions of both faculty and students need to be identified. This systematic review will allow us to address these issues.

Methods: Using BEME (Best Evidence in Medical Education) Collaboration review guidelines, electronic databases including MEDLINE via Pubmed, MEDLINE via OVID CINAHL via EBSCO AMED via EBSCO, ERIC via EBSCO, British Education Index via EBSCO, Australian Education Index via EBSCO will be searched for studies on formative OSCEs. Controlled trials (randomized and non-randomized), quasi-experimental designs such as repeated measures or pretest/post-test, cohort studies including cross-sectional studies, case control studies, qualitative studies, mixed methods studies reporting on formative OSCE for undergraduate medical or dental students will be included. Studies reported in all languages will be included and non-English articles will be translated. Identified studies will be screened by four of the authors who will work in pairs and extract data for 20% of the retrieved articles each using a modified BEME Coding sheet. Methodological quality will be evaluated using either the BEME Quality Indicators (Buckley et al., 2009) and/or Medical Education Research Study Quality Instrument (MERSQI) (for quantitative studies) (Reed et al., 2007) and

BEME PROTOCOL

Standards for Reporting Qualitative Research items (for qualitative studies) (O'Brien et al., 2014).

Results: If the trials reporting how different types of formative OSCE promote learning, present qualitative data, thematic analysis will be done using Braun and Clarke's method (Braun and Clarke, 2014). If the results are presented as quantitative data, and the data is sufficiently homogenous (for example, interventions, comparisons, outcomes, and study designs are similar), we will follow standard methods for meta-analysis as per the Cochrane handbook (Higgins and Thomas, 2019). Similarly, for the outcome on performance in summative OSCE after formative OSCE, we expect mixed data and will follow the methods described above. (both qualitative and quantitative).

If quantitative data is expressed dichotomously, we will use risk ratios (RRs), and for continuous data, we will assess the mean difference (MD). We will present the results with 95% confidence intervals (CI). If ordinal scales are used, we will use proportional odds ratio. If the data obtained is too heterogenous to be combined for quantitative statistical meta-analysis, we will perform a qualitative review of the evidence.

Kirkpatrick's levels will be used to analyze behavioral changes after formative OSCE. (Steinert et al., 2006). Thematic analysis will be done for the perceptions obtained from faculty and students, using Braun and Clarke's method (Braun and Clarke, 2014).

Systematic review registration: The topic was successfully registered with BEME on 5th March with Reg No: 0122

Keywords:

1. Formative assessment
2. Objective structured clinical examination
3. OSCE
4. Formative OSCE
5. Medical education and dentistry
6. Undergraduate students
7. Formative feedback
8. Group OSCE (GOSCE)
9. Mini OSCE

BEME PROTOCOL

1. Introduction:

Objective structured clinical examination (OSCE) was posited by Harden in 1975 to minimize the disadvantages of the traditional clinical examination (Harden et al., 1975). OSCE gives a ‘snapshot’ of the candidate’s demonstrated performance in a particular area where the task, patients and examiners remain the same (Khan et al., 2013; Akhigbe, 2018). It allows many competencies to be assessed in a series of stations, which are the same for all candidates. In OSCE the variables and complexity of the examination are more easily controlled, aims can be more clearly defined, and varying aspects of student's knowledge can be tested. The examination is more objective, and a marking strategy can be decided in advance. The examination results in improved feedback to students and staff (Harden et al., 1975).

OSCE are generally used to assess different constructs than those that are targeted by written tests, such as clinical skills and the soft skills that include ability to establish rapport with a patient and professionalism (Zayyan, 2011). With reference to Miller's pyramid of assessment, the OSCE assesses whether a candidate can ‘show how’ one would perform in a simulated environment (Norcini, 2003). There is good evidence to support the use of OSCE as a reliable tool to assess clinical competence and determine progress (Chisnall et al., 2015; Pugh, Desjardins and Eva, 2017). Both these factors suggest that the learning that is required in OSCE is less likely to reflect simple declarative knowledge and more likely to reflect complex behaviors that will interact in variably interpreted ways (Chisnall et al., 2015; Pugh, Desjardins and Eva, 2017).

Over the years OSCE have been used for various purposes as an assessment tool. Primarily OSCE are used summatively to formally assess skills and knowledge and constitutes part of the high-stakes examinations meant for grading, ranking, declaring pass-fail or progression of student to next level (Chisnall et al., 2015). Students frequently have no opportunity to practice OSCE other than in the high-stakes examination itself (Young et al., 2014). This led to the use of OSCE formatively, where OSCE is primarily a learning tool; it does not contribute to a student's final assessment mark, and “passing” a formative OSCE is not an academic requirement. Its primary objective is providing feedback on their performance, thereby enabling improvement in learning and performance before summative OSCE (Chisnall et al., 2015). Formative OSCE has been used to provide feedback to students on their clinical skills, aiming to enhance learners’ behavior and help students recognize their weaknesses

BEME PROTOCOL

(Brazeau, Boyd and Crosson, 2002; Aeder, 2007; Larsen and Jeppe-Jensen, 2008; O'Sullivan et al., 2008; Shute, 2008; Bagri et al., 2009; Lele, 2011; Alsenany and Al Saif, 2012; Browne et al., 2012; Bennett and Furmedge, 2013; Pelgrim et al., 2013; Chisnall et al., 2015; Junod Perron et al., 2016; Pugh, Desjardins and Eva, 2017; Sulaiman et al., 2018)

1a. Background:

The scoping search we conducted identified different types of formative OSCE with different type of feedback mechanisms in literature.

Different types of formative OSCE:

Formative OSCE in the form of Group Objective Structured Clinical Examination (GOSCE) is a variation of the traditional OSCE format which has been reported in a few studies in the medical education literature (Biran, 1991; Sulaiman et al., 2018). In GOSCE, learners are assigned to groups rather than individually as they rotate around the OSCE stations. (Elliot et al., 1994; Meagher et al., 2009; Konopasek et al., 2014). GOSCE provides a chance for the learners to observe each other performing the clinical task at each station, and more importantly practice self-assessment and receive feedback (Biran, 1991). The format of the GOSCE has the added benefit of being resource-effective for assessing clinical skills, requiring fewer facilities per learner than the conventional formative OSCE (Biran, 1991).

Mini-OSCE with five stations has also been used in dentistry and was found to be a fairly valid and reliable tool for formative assessment (Lele, 2011).

Mock OSCE have been reported as OSCE arranged prior to the summative assessment to prepare the student for the examination. It is important to distinguish between mock OSCEs, and formative OSCE, as they serve different purposes and have different educational outcomes. A mock OSCE replicates the summative OSCE, allowing students to experience the timings, format, layout, length, and station content of a summative OSCE (Bennett and Furmedge, 2013). It is done just a few days prior to the summative OSCE. A formative OSCE, however, may take a different format or length to a summative OSCE, and is primarily designed to enhance learning of clinical and examination skills (Bagri et al., 2009; Young et al., 2014; Chisnall et al., 2015). A formative OSCE is used as a learning tool and may be helpful in identifying struggling students to enable additional support prior to any summative OSCE by effective feedback mechanism. It also contributes to reinforce student's self-assessment and self-confidence (Junod Perron et al., 2016).

BEME PROTOCOL

Different types of feedback:

External feedback in formative OSCE plays a crucial role in enhancing appropriate learning, correcting deficiencies, and monitoring students' self-learning (Eva and Regehr, 2005; Eva and Regehr, 2008) since self-assessment is not enough to ensure accurate identification of areas for improvement and to develop effective learning (Shute, 2008; Kogan et al., 2012). Formative feedback effectiveness is influenced by several factors, including tutor-student interactions, the feedback message, and the environment in which feedback takes place (Archer, 2010; Kogan et al., 2012). It seems to be more effective when learners are more oriented towards learning than performance demonstration (Shute, 2008; Teunissen et al., 2009; Archer, 2010; Pelgrim et al., 2013), and when the feedback content matches some of the learners' self-perceptions (Sargeant, Mann and Ferrier, 2005). Formative feedback is also more effective when it is based on observed facts, focuses on tasks, is specific, concise, and suggests areas of improvement (Norcini and Burch, 2007).

Feedback may be directive where the tutor tells the learner what to do or improve; or it may be facilitative and elaborative where the tutor makes suggestions to guide learners in self-reflection about the encounter and their skills (Shute, 2008; Kogan et al., 2012). Feedback also seems to be better accepted when the tutor is perceived as being credible (Kogan et al., 2012; Archer, 2010; Eva et al., 2011) and is comfortable with the topic. Literature reports that personal and professional characteristics of tutors who give feedback to students during formative OSCE may influence the way feedback is delivered and perceived (Perron et al., 2016).

Giving and receiving feedback from someone at the same level has been shown to be both effective and has the potential to raise the self-esteem of everyone involved. Formative OSCE driven by medical students, for medical students, or driven by residents for residents, has been reported in several studies (Thoresen, n.d.)

Feedback after formative OSCE can be given to students immediately after the clinical encounter or subsequently at a scheduled time (Rush et al., 2014). Some studies reported use of video-based feedback. (Perron et al., 2016). Video review with expert feedback was reported to be more effective in terms of students' satisfaction and performance and was superior to expert feedback alone (Perron et al., 2016).

BEME PROTOCOL

1b. Importance of this review:

There are several reasons why we want to undertake this review.

1. Although, literature suggests that formative OSCE promotes learning and improves performance, in our opinion, it is not clear how it does so. In addition, it is important to see if the different types of formative OSCE and feedback have a role in student learning. Formative OSCE and its effect on behavioral change in students daily clinical practice needs to be explored. Finally, the perceptions of both faculty and students need to be identified.
2. To the best of our knowledge, there are no systematic reviews reported till date exploring how formative OSCE promotes learning and improves performance.
3. While formative assessments have been institutionalized in some medical schools, such a practice is rather infrequent or even rare in most regions of the world particularly in Mediterranean and Asian countries. (Alkhateeb 2018). The results of this systematic review can be a reference for various institutions to bring about changes in their formative OSCE policies.

2. Review questions, Review objectives and Key words:

2.1 Review question:

1. How do different types of formative OSCE promote learning?
2. How do different types of feedback in formative OSCE promote learning?
3. How does formative OSCE improve performance in summative OSCE?
4. How does formative OSCE influences change in student behavior in adhering to protocols/guidelines in clinical practice
5. What are the views of educators on formative OSCE?
6. What are the views of students on formative OSCE?

2.1 Review objectives:

1. To explore how different types of formative OSCE promotes learning
2. To explore how different types of feedback in formative OSCE promotes learning
3. To assess in what ways formative OSCE improves performance in summative OSCE

BEME PROTOCOL

4. To analyze the extent to which formative OSCE influences change in student behaviour in adhering to protocols/guidelines in clinical practice
5. To summarize the perceptions of educators on formative OSCE.
6. To summarize the perceptions of students on formative OSCE.

3. Search sources and strategies:

Search terms will use subject headings where thesauri exist in conjunction with free text terms using truncation and appropriate Boolean operators. A combination of MeSH terms and keywords will be used to capture studies of interest under the following headings

1. Formative assessment
2. Objective structured clinical examination
3. OSCE
4. Formative OSCE
5. Medical education and dentistry
6. Undergraduate students
7. Formative feedback
8. Group OSCE (GOSCE)
9. Mini OSCE

3.1 Scoping search:

We performed an initial search of the six databases in November 2019. The search strings and numbers of records identified are as follows:

Medline via Pubmed: 14 November 2019: 107

Search (((((((Objective structured clinical examination) OR OSCE) OR group OSCE) OR GOSCE) OR mini-OSCE) OR mini OSCE)) AND ((formative assessment) OR formative feedback)) AND ((((((medical) OR medicine)) OR ((dental) OR dentistry))) AND (((student) OR students)) OR undergraduate))

Medline via Ovid: 14 November 2019: 567

Sl. No.	Searches	Results
1.	objective structured clinical examination.mp.	1572
2.	OSCE.mp.	1926
3.	group OSCE.mp.	4
4.	GOSCE.mp.	11
5.	mini OSCE.mp.	3
6.	1 or 2 or 3 or 4 or 5	2300

BEME PROTOCOL

7.	Students, Dental/	6312
8.	Students, Medical/	32096
9.	7 or 8	38115
10.	exp Formative Feedback/	738
11.	6 and 10	20
12.	Educational Measurement/ or Curriculum/ or Humans/ or Clinical Competence/ or Learning/ or formative assessment.mp.	18147097
13.	6 and 12	1953
14.	11 or 13	1953
15.	9 and 14	567
16.	from 15 keep 1-567	567

ERIC via EBSCO: 14 November: 13

Search ID	Search terms	Results
S1	objective structured clinical examination OR osce OR gosce OR group osce OR mini osce	144
S2	formative assessment OR formative feedback OR feedback	37181
S3	medical students OR dental students OR undergraduate students	124182
S4	S1 AND S2 AND S3	13

British Education Index: 14 November: 2

Search ID	Search terms	Results
S1	objective structured clinical examination OR osce OR gosce OR group osce OR mini osce	282
S2	formative assessment OR formative feedback OR feedback	328
S3	medical students OR dental students OR undergraduate students	7236
S4	S1 AND S2 AND S3	2

BEME PROTOCOL

Australian Education Index via ProQuest: 14 November: 5

Search ID	Search terms	Results
S1	objective structured clinical examination OR osce OR gosce OR group osce OR mini osce	41
S2	formative assessment OR formative feedback OR feedback	5291
S3	medical students OR dental students OR undergraduate students	7898
S4	S1 AND S2 AND S3	5

BEME Review Title: Formative assessment in OSCE

Contact Reviewer: Prashanti Eachempati

Summary of Searches

April 2019

Searches carried out by

Sumanth Kumbargere Nagraj

Faculty of Dentistry, Melaka Manipal Medical College, Malaysia

Database	Version/issue	Date of search	Records retrieved	After removal of duplicates
MEDLINE via Pubmed	Preliminary search	21 April 2019	107	107
MEDLINE via OVID	Preliminary search	21 April 2019	567	567
ERIC via EBSCO	Preliminary search	21 April 2019	13	12
British Education Index via EBSCO	Preliminary search	21 April 2019	2	1

BEME PROTOCOL

Australian Education Index via EBSCO	Preliminary search	21 April 2019	5	4
Total retrieved from electronic searches: 694				
Total left after deduplication: 691				
Total sent to authors for this search: 691				

- We will also search bibliographies of included studies
- Continually monitor the literature for newly published relevant articles (regular online automatic updates will be created)
- The next search will be conducted within 6 months of the protocol publication

4. Study selection criteria:

Inclusion criteria:

1. Controlled trials (randomized and non-randomized), quasi-experimental designs such as repeated measures or pretest/post-test, cohort studies including cross-sectional studies, case control studies, qualitative studies, mixed methods studies reporting on formative OSCE for undergraduate medical or dental students will be included.
2. Studies reported in all languages will be included and non-English articles will be translated. However, if non-English articles report qualitative data, they will not be included, as the interpretation of data may vary when translated.

Exclusion criteria:

1. Studies on formative OSCE for postgraduates, residents
2. Studies reporting mock OSCE
3. Conference proceedings or abstracts

5. Extracting Data:

5.1 Data to be extracted include the following:

1. Publication characteristics (i.e., publication type, journal, year, authors)
2. Country, state, study venue
3. Method of study identification (e.g., electronic, bibliography, expert recommendation,)

BEME PROTOCOL

4. Population studied (medical or dental)
5. Study aim (and whether this was implied or stated)
6. Is formative OSCE the primary objective (or are others present)
7. Inclusion/exclusion criteria
9. Study design, including presence or absence of control group and whether subjects were randomized
10. Was the study qualitative or quantitative?
11. Was the study design retrospective or prospective?
12. Did authors report *a priori* hypotheses?
13. Sample size
14. Data collection methods
15. Methods used to measure effectiveness of formative OSCE
16. Intervention(s) and comparators
17. Primary outcomes/themes studied and how they were measured
18. Primary conclusions (with regard to both theoretical and practical implications of formative OSCE)
19. Measure of study quality using Medical Education Research Study Quality Instrument (MERSQI) (Reed et al., 2007) and/or BEME Quality Indicators (Buckley et al., 2009) (for quantitative studies) and Standards for Reporting Qualitative Research items (O'Brien et al., 2014) (for qualitative)
20. Conclusions or findings with regard to formative OSCE

Using a similar method described in Aileen Barrett's BEME Protocol (Barrett et al., 2016), four of the study authors will work in pairs (Prashanti Eachempati (PE), Kiran Kumar (KK), Noorliza Mastura Ismail (NMI) and Htoo Htoo (HH)) and independently review 20% of the retrieved articles (randomly selected) each using a customized BEME Coding sheet which we will develop. Studies that are considered eligible for inclusion will be read fully in duplicate and their suitability for inclusion will be determined. The reviewer's data extraction will then be validated for accuracy by a moderator from the group (Sumanth KN (SKN)) for inter-rater reliability to a kappa of 0.80 agreement.) Once this agreement has been reached, all articles will be evenly distributed among the four study authors and data extractions completed in

BEME PROTOCOL

duplicate. In case of disagreements, the two reviewers will read the article in question in full text. If no resolution found, the moderator from the group (SKN) will review the paper in an effort to reach a consensus view. The third-party adjudication will be recorded using an Excel spreadsheet for administration purposes. Here the adjudicator will record relevant bibliographic information, annotations from the primary reviewers about inclusion and exclusion decisions. The adjudicator will have two separate columns, one to record the decision and another to record reasons for decision and recommendations for reviewers to consider. Additional data will be sought from authors where necessary.

6. Appraisal of studies:

Methodological quality will be evaluated using either the BEME Quality Indicators (Buckley et al., 2009) (Appendix 1) and/or Medical Education Research Study Quality Instrument (MERSQI) (Reed et al., 2007) (for quantitative studies) (Appendix 2) and Standards for Reporting Qualitative Research items (O'Brien et al., 2014) (for qualitative) (Appendix 3). Recognizing limitations around reporting quality, we will include a formal risk of bias assessment.

Two coders (PE and SKN) will appraise the quality of each study once it has been deemed suitable for inclusion. Differences of opinion will be resolved by discussion. In case of disagreement, the two coders will meet with a third party (Abdul Rashid (AR) and Ramnarayan K (RK)) who will appraise the study to reach a consensus view. The third part adjudication will be recorded using an Excel spreadsheet for administration purposes. Here the adjudicator will record relevant bibliographic information, annotations from the primary reviewers about their evaluation. The adjudicator will have two separate columns, one to record the decision and another to record reasons for decision and recommendations for reviewers to consider. All articles meeting the inclusion criteria will be considered and quality of the articles will be appraised and reported.

7. Synthesis of extracted evidence and translation into practice

Data extracted will be synthesized and presented both quantitatively and qualitatively. Search results and inclusion/exclusion data will be presented using a PRISMA-style search decision flowchart. This flow diagram will depict the flow of information through the different phases of the systematic review. It maps out the number of records identified, included and excluded, and the reasons for exclusions. Characteristics of included studies will be presented in tabular format. In addition to tabular presentation, synthesis will involve descriptively comparing,

BEME PROTOCOL

contrasting, and discussing studies and outcomes that address our primary research questions, as well as salient themes that become evident during the course of the review. Figures will be created if appropriate to better illustrate important themes.

The discussion section will include a narrative synthesis addressing each of the research questions, other themes identified during the literature search, study limitations and knowledge gaps identified, and directions for future research.

If the trials reporting how different types of formative OSCE promote learning, present qualitative data, thematic analysis will be done using Braun and Clarke’s method (Braun and Clarke, 2014). If the results are presented as quantitative data, and the data is sufficiently homogenous (for example, interventions, comparisons, outcomes, and study designs are similar), we will follow standard methods for meta-analysis as per the Cochrane handbook (Higgins and Thomas, 2019). Similarly, for the outcome on performance in summative OSCE after formative OSCE, we expect mixed data and will follow the methods described above. (both qualitative and quantitative).

If quantitative data is expressed dichotomously, we will use risk ratios (RRs), and for continuous data, we will assess the mean difference (MD). We will present the results with 95% confidence intervals (CI). If ordinal scales are used, we will use proportional odds ratio. If the data obtained is too heterogenous to be combined for quantitative statistical meta-analysis, we will perform a qualitative review of the evidence.

Kirkpatrick’s levels will be used to analyze behavioral changes after formative OSCE (Steinert et al., 2006). Thematic analysis will be done for the perceptions obtained from faculty and students, using Braun and Clarke’s method (Braun and Clarke, 2014).

Level 1	REACTION- Learner reactions	Participants and faculty feedback on formative OSCE
Level2A	LEARNING- Modification of attitude/skills	Improvement in learning and performance as observed by self/ peers/faculty
Level2B	LEARNING- Acquisition of knowledge or skills	Improvement in knowledge and skill as indicated by their performance in summative examination and perception of faculty/peers/self, regarding their learning
Level3	BEHAVIOUR - Change in behaviors	Willingness of learners to apply new knowledge and skills/ behavioral change in daily clinical practice to adhere to clinical guidelines – Self reported and observed

BEME PROTOCOL

Level4A	RESULTS - Change in the system / organizational practice	Wider changes in the organization, attributable to the intervention
Level4B	RESULTS – Benefits to patients	Perceptions of patients / improvement in patient care – self reported and observed

References:

1. Aeder, L., Altshuler, L., Kachur, E., Barrett, S., Hilfer, A., Koepfer, S., Schaeffer, H. and Shelov, S.P. (2007). The " Culture OSCE"-Introducing a formative assessment into a postgraduate program. *Education for health*, 20(1), p.11.
2. Akhigbe, T. (2018). Summative Objective Structured Clinical Examination Assessment: A Mini Review. *International Journal of Medical Reviews*, 5(4), pp.140-142.
3. Alkhateeb, N. (2018). Impact of Formative OSCE on Students' Summative Clinical Performance - Full Text View - ClinicalTrials.gov. [online] Clinicaltrials.gov. Available at: <https://clinicaltrials.gov/ct2/show/NCT03599232> [Accessed 14 Nov. 2019].
4. Alsenany, S. and Al Saif, A. (2012). Developing skills in managing Objective Structured Clinical Examinations (OSCE). *Life Science Journal*, 9(3), pp.597-602.
5. Archer, J. (2010). State of the science in health professional education: effective feedback. *Medical Education*, 44(1), pp.101-108.
6. Bagri, A., Zaw, K., Milanez, M., Palacios, J., Qadri, S., Bliss, L., Roos, B. and Ruiz, J. (2009). Geriatric Medicine Fellows' Experiences and Attitudes Toward an Objective Structured Clinical Examination (OSCE). *Educational Gerontology*, 35(4), pp.283-295.
7. Barrett, A., Galvin, R., Steinert, Y., Scherpbier, A., O'Shaughnessy, A., Horgan, M. and Horsley, T. (2016). A BEME (Best Evidence in Medical Education) review of the use of workplace-based assessment in identifying and remediating underperformance among postgraduate medical trainees: BEME Guide No. 43. *Medical Teacher*, 38(12), pp.1188-1198.
8. Bennett, V. and Furmedge, D. (2013). The hidden value of a mock OSCE. *The Clinical Teacher*, 10(6), pp.407-408.
9. Biran, L. (1991). Self-assessment and learning through GOSCE (group objective structured clinical examination). *Medical Education*, 25(6), pp.475-479.

BEME PROTOCOL

10. Braun, V. and Clarke, V. (2013). *Successful qualitative research*. 1st ed. London: SAGE.
11. Brazeau, C., Boyd, L. and Crosson, J. (2002). Changing an Existing OSCE to a Teaching Tool. *Academic Medicine*, 77(9), p.932.
12. Browne, G., Bjelogric, P., Issberner, J. and Jackson, C. (2012). Undergraduate student assessors in a formative OSCE station. *Medical Teacher*, 35(2), pp.170-171.
13. Buckley, S., Coleman, J., Davison, I., Khan, K., Zamora, J., Malick, S., Morley, D., Pollard, D., Ashcroft, T., Popovic, C. and Sayers, J. (2009). The educational effects of portfolios on undergraduate student learning: A Best Evidence Medical Education (BEME) systematic review. BEME Guide No. 11. *Medical Teacher*, 31(4), pp.282-298.
14. Chisnall, B., Vince, T., Hall, S. and Tribe, R. (2015). Evaluation of outcomes of a formative objective structured clinical examination for second-year UK medical students. *International Journal of Medical Education*, 6, pp.76-83.
15. Elliot, D., Fields, S., Keenen, T., Jaffe, A. and Toffler, W. (1994). Use of a group objective structured clinical examination with first-year medical students. *Academic Medicine*, 69(12), pp.990-2.
16. Eva, K. and Regehr, G. (2005). Self-Assessment in the Health Professions: A Reformulation and Research Agenda. *Academic Medicine*, 80(Supplement), pp.S46-S54.
17. Eva, K. and Regehr, G. (2008). "I'll never play professional football" and other fallacies of self-assessment. *Journal of Continuing Education in the Health Professions*, 28(1), pp.14-19.
18. Eva, K., Armson, H., Holmboe, E., Lockyer, J., Loney, E., Mann, K. and Sargeant, J. (2011). Factors influencing responsiveness to feedback: on the interplay between fear, confidence, and reasoning processes. *Advances in Health Sciences Education*, 17(1), pp.15-26.
19. Harden, R., Stevenson, M., Downie, W. and Wilson, G. (1975). Assessment of clinical competence using objective structured examination. *BMJ*, 1(5955), pp.447-451.
20. Higgins, J. and Thomas, J. (2019). *Cochrane Handbook for Systematic Reviews of Interventions*. [online] Training.cochrane.org. Available at: <https://training.cochrane.org/handbook/current> [Accessed 5 Nov. 2019].

BEME PROTOCOL

21. Junod Perron, N., Louis-Simonet, M., Cerutti, B., Pfarrwaller, E., Sommer, J. and Nendaz, M. (2016). The quality of feedback during formative OSCEs depends on the tutors' profile. *BMC Medical Education*, 16(1).
22. Khan, K., Ramachandran, S., Gaunt, K. and Pushkar, P. (2013). The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part I: An historical and theoretical perspective. *Medical Teacher*, 35(9), pp.e1437-e1446.
23. Kogan, J., Conforti, L., Bernabeo, E., Durning, S., Hauer, K. and Holmboe, E. (2012). Faculty staff perceptions of feedback to residents after direct observation of clinical skills. *Medical Education*, 46(2), pp.201-215.
24. Konopasek, L., Kelly, K., Bylund, C., Wenderoth, S. and Storey-Johnson, C. (2014). The Group Objective Structured Clinical Experience: Building communication skills in the clinical reasoning context. *Patient Education and Counseling*, 96(1), pp.79-85.
25. Larsen, T. and Jeppe-Jensen, D. (2008). The introduction and perception of an OSCE with an element of self- and peer-assessment. *European Journal of Dental Education*, 12(1), pp.2-7.
26. Lele, S.M. (2011). A mini-OSCE for formative assessment of diagnostic and radiographic skills at a dental college in India. *Journal of dental education*, 75(12), pp.1583-1589.
27. Meagher, F., Butler, M., Miller, S., Costello, R., Conroy, R. and McElvaney, N. (2009). Predictive validity of measurements of clinical competence using the Team Objective Structured Bedside Assessment (TOSBA): Assessing the clinical competence of final year medical students. *Medical Teacher*, 31(11), pp.e545-e550.
28. Norcini, J.J. (2003). Work based assessment. *BMJ*, 326(7392), pp.753-755.
29. Norcini, J. and Burch, V. (2007). Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Medical Teacher*, 29(9-10), pp.855-871.
30. O'Brien, B., Harris, I., Beckman, T., Reed, D. and Cook, D. (2014). Standards for Reporting Qualitative Research. *Academic Medicine*, 89(9), pp.1245-1251.
31. O'Sullivan, P., Chao, S., Russell, M., Levine, S. and Fabiny, A. (2008). Development and Implementation of an Objective Structured Clinical Examination to Provide Formative Feedback on Communication and Interpersonal Skills in Geriatric Training. *Journal of the American Geriatrics Society*, 56(9), pp.1730-1735.

BEME PROTOCOL

32. Pelgrim, E., Kramer, A., Mokkink, H. and van der Vleuten, C. (2013). Reflection as a component of formative assessment appears to be instrumental in promoting the use of feedback; an observational study. *Medical Teacher*, 35(9), pp.772-778.
33. Perron, N., Louis-Simonet, M., Cerutti, B., Pfarrwaller, E., Sommer, J. and Nendaz, M. (2016). Feedback in formative OSCEs: comparison between direct observation and video-based formats. *Medical Education Online*, 21(1), p.32160.
34. Pugh, D., Desjardins, I. and Eva, K. (2017). How do formative objective structured clinical examinations drive learning? Analysis of residents' perceptions. *Medical Teacher*, 40(1), pp.45-52.
35. Reed, D., Cook, D., Beckman, T., Levine, R., Kern, D. and Wright, S. (2007). Association Between Funding and Quality of Published Medical Education Research. *JAMA*, 298(9), p.1002.
36. Rush, S., Ooms, A., Marks-Maran, D. and Firth, T. (2014). Students' perceptions of practice assessment in the skills laboratory: An evaluation study of OSCAs with immediate feedback. *Nurse Education in Practice*, 14(6), pp.627-634.
37. Sargeant, J., Mann, K. and Ferrier, S. (2005). Exploring family physicians' reactions to multisource feedback: perceptions of credibility and usefulness. *Medical Education*, 39(5), pp.497-504.
38. Shute, V. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1), pp.153-189.
39. Steinert, Y., Mann, K., Centeno, A., Dolmans, D., Spencer, J., Gelula, M. and Prideaux, D. (2006). A systematic review of faculty development initiatives designed to improve teaching effectiveness in medical education: BEME Guide No. 8. *Medical Teacher*, 28(6), pp.497-526.
40. Sulaiman, N., Shorbagi, S., Abdalla, N., Daghistani, M., Mahmoud, I. and Al-Moslih, A. (2018). Group OSCE (GOSCE) as a formative clinical assessment tool for pre-clerkship medical students at the University of Sharjah. *Journal of Taibah University Medical Sciences*, 13(5), pp.409-414.
41. Teunissen, P., Stapel, D., van der Vleuten, C., Scherpbier, A., Boor, K. and Scheele, F. (2009). Who Wants Feedback? An Investigation of the Variables Influencing Residents' Feedback-Seeking Behavior in Relation to Night Shifts. *Academic Medicine*, 84(7), pp.910-917.

BEME PROTOCOL

42. Thoresen, M. (n.d.). *Formative OSCE - CAME - Center of Assessment in Medical Education - NTNU*. [online] Ntnu.edu. Available at: <https://www.ntnu.edu/came/osce> [Accessed 1 Nov. 2019].
43. Young, I., Montgomery, K., Kearns, P., Hayward, S. and Mellanby, E. (2014). The benefits of a peer-assisted mock OSCE. *The Clinical Teacher*, 11(3), pp.214-218.
44. Zayyan, M. (2011). Objective Structured Clinical Examination: The Assessment of Choice. *Oman Medical journal*, pp.219-222.

BEME PROTOCOL

	Pre- review work:2019 March- Sept	Sept	Oct	Nov	Dec	Jan 2020	Feb	March	April	May	June	July	Aug	Sept
Title registration, protocol development and review														
Data search and data synthesis														
Coding of studies and data analysis														
Data synthesis														
Preparation of paper for publication														

Appendix 1:

BEME Quality Indicators	
Research Question	Is the research question or hypothesis clearly stated?
Study Subjects	Is the subject group appropriate for the study being carried out?
Data Collection	Are the methods used appropriate for the research question and context?
Completeness of data	Attrition rates/acceptable questionnaire response rates
Risk of bias assessment	Is a statement of author positionality and a risk of bias assessment included?
Analysis of results	Are the statistical and other methods of results analysis used appropriate?
Conclusions	Is it clear that the data justify the conclusions drawn?
Reproducibility	Could other researchers repeat the study?
Prospective	Is the study prospective?
Ethical Issues	Are all ethical issues articulated and managed appropriately?

Triangulation	Were results supported by data from more than one source?

Appendix 2:

Domain	MERSQI Item	Score	Max Score
Study design	Single Group Cross-sectional or single group posttest only	1	3
	Single group pretest & posttest	1.5	
	Nonrandomized, 2 groups	2	
	Randomized controlled trial	3	
Sampling	<i>Institutions studied:</i>		3
	1	0.5	
	2	1	
	3	1.5	
	<i>Response rate, %:</i>		
	Not applicable		
	<50 or not reported	0.5	
	50- 74	1	

	>75	1.5	
Type of data	Assessment by participants	1	3
	Objective measurement	3	
Validity of evaluation instrument	<i>Internal structure:</i>		3
	Not applicable		
	Not reported	0	
	Reported	1	
	<i>Content:</i>		
	Not applicable		
	Not reported	0	
	Reported	1	
	<i>Relationships to other variables:</i>		
	Not applicable		
	Not reported	0	
	Reported	1	
Data analysis	<i>Appropriateness of analysis:</i>		3
	Inappropriate for study design or type of data	0	
	Appropriate for study design, type of data	1	
	<i>Complexity of analysis:</i>		
	Descriptive analysis only	1	
	Beyond descriptive analysis	2	
Outcomes	Satisfaction, attitudes, perception, opinions, general facts	1	3
	Knowledge, skills	1.5	
	Behaviors	2	
	Patient/health care outcome	3	
Total possible score*			18

Appendix 3:

	Standards for Reporting Qualitative Research (SRQR)*	
	http://www.equator-network.org/reporting-guidelines/srqr/	
		Page/line no(s).
Title and abstract		
	Title - Concise description of the nature and topic of the study Identifying the study as qualitative or indicating the approach (e.g., ethnography, grounded theory) or data collection methods (e.g., interview, focus group) is recommended	
	Abstract - Summary of key elements of the study using the abstract format of the intended publication; typically includes background, purpose, methods, results, and conclusions	
Introduction		

	Problem formulation - Description and significance of the problem/phenomenon studied; review of relevant theory and empirical work; problem statement	
	Purpose or research question - Purpose of the study and specific objectives or questions	
Methods		
	Qualitative approach and research paradigm - Qualitative approach (e.g., ethnography, grounded theory, case study, phenomenology, narrative research) and guiding theory if appropriate; identifying the research paradigm (e.g., postpositivist, constructivist/ interpretivist) is also recommended; rationale**	
	Researcher characteristics and reflexivity - Researchers' characteristics that may influence the research, including personal attributes, qualifications/experience, relationship with participants, assumptions, and/or presuppositions; potential or actual interaction between researchers' characteristics and the research questions, approach, methods, results, and/or transferability	
	Context - Setting/site and salient contextual factors; rationale**	
	Sampling strategy - How and why research participants, documents, or events were selected; criteria for deciding when no further sampling was necessary (e.g., sampling saturation); rationale**	
	Ethical issues pertaining to human subjects - Documentation of approval by an appropriate ethics review board and participant consent, or explanation for lack thereof; other confidentiality and data security issues	
	Data collection methods - Types of data collected; details of data collection procedures including (as appropriate) start and stop dates of data collection and analysis, iterative process, triangulation of sources/methods, and modification of procedures in response to evolving study findings; rationale**	
	Data collection instruments and technologies - Description of instruments (e.g., interview guides, questionnaires) and devices (e.g., audio recorders) used for data collection; if/how the instrument(s) changed over the course of the study	

	Units of study - Number and relevant characteristics of participants, documents, or events included in the study; level of participation (could be reported in results)	
	Data processing - Methods for processing data prior to and during analysis, including transcription, data entry, data management and security, verification of data integrity, data coding, and anonymization/de-identification of excerpts	
	Data analysis - Process by which inferences, themes, etc., were identified and developed, including the researchers involved in data analysis; usually references a specific paradigm or approach; rationale**	
	Techniques to enhance trustworthiness - Techniques to enhance trustworthiness and credibility of data analysis (e.g., member checking, audit trail, triangulation); rationale**	
Results/findings		
	Synthesis and interpretation - Main findings (e.g., interpretations, inferences, and themes); might include development of a theory or model, or integration with prior research or theory	
	Links to empirical data - Evidence (e.g., quotes, field notes, text excerpts, photographs) to substantiate analytic findings	
Discussion		
	Integration with prior work, implications, transferability, and contribution(s) to the field - Short summary of main findings; explanation of how findings and conclusions connect to, support, elaborate on, or challenge conclusions of earlier scholarship; discussion of scope of application/generalizability; identification of unique contribution(s) to scholarship in a discipline or field	
	Limitations - Trustworthiness and limitations of findings	
Other		
	Conflicts of interest - Potential sources of influence or perceived influence on study conduct and conclusions; how these were managed	

	<p>Funding - Sources of funding and other support; role of funders in data collection, interpretation, and reporting</p>	
	<p>*The authors created the SRQR by searching the literature to identify guidelines, reporting standards, and critical appraisal criteria for qualitative research; reviewing the reference lists of retrieved sources; and contacting experts to gain feedback. The SRQR aims to improve the transparency of all aspects of qualitative research by providing clear standards for reporting qualitative research.</p>	
	<p>**The rationale should briefly discuss the justification for choosing that theory, approach, method, or technique rather than other options available, the assumptions and limitations implicit in those choices, and how those choices influence study conclusions and transferability. As appropriate, the rationale for several items might be discussed together.</p>	