Design and Validation of an integrated Objective Structured Clinical Examination (i-OSCE) for non-surgical aesthetics program

Narendra Kumar (✉ narendra.kumar.19@ucl.ac.uk)
  University College London

Eqram Rahman
  University College London

Research Article

Abstract

Introduction: The Objective Structured Clinical Examination (OSCE) is a popular and practical method for evaluating trainee physicians' competencies. To help assess non-surgical aesthetics students' critical thinking and relevant abilities, we developed a novel assessment tool based on the OSCE, the i-OSCE (Integrated objective structured Clinical examination).

Methods: Initially, an expert panel consisting of 5 Aesthetic Practitioners with over fifteen years of experience and a senior clinical academic were selected to develop a blueprint for i-OSCE. Through this blueprint, essential qualities and skills were identified for the assessment. To ensure the process standardisation, training workshops for examiners and simulated patients were organised. The final i-OSCE consisted of 12 stations (four clinical, four critical thinking, and four rest stations lasting 180 minutes.

Results: The Interclass correlation coefficient between the station checklist items was 0.946 (average measure upper bound 0.916, lower bound 0.968; $p < 0.00$), considered to be significant. The Inter-Item Correlation Matrix among the clinical station checklist and critical thinking checklist items also showed statistical significance. The Pearson correlation coefficient (PCC) used to ascertain the correlation between checklist rating and global rating, yielding a high correlation (0.80 to 0.934).

Conclusion: The i-OSCE has been proven to be a useful and reliable assessment tool to evaluate clinical competence and critical thinking in non-surgical aesthetics education.

Introduction

Assessment and evaluation are critical steps in medical education and rely on selecting a proper and robust instrument. Appropriate assessment tool helps to determine the effectiveness of educational programmes and ensures that the future clinicians are competent and suitable for independent clinical practice. However, the currently used assessment tools are insufficient to test the learners' knowledge, skills, behaviour, and critical thinking abilities holistically. In such cases, using the 'test battery' approach becomes more practical to use a mix of assessment tools for measuring an array of learning domains (1, 2).

Traditionally, clinical assessment strategies comprise a combination of 'short' and 'long case' evaluations. However, criticism about its low reliability (3) and modern-day constraints such as increased litigation and student appeals (4) have led institutions to focus on exams that produce trustworthy, more easily defendable outcomes. Accordingly, conventional assessment strategy evolved to overcome the challenges of traditional methodologies, such as reliance on the patient's performance, the examiner's bias, the non-standardised grading scheme, and the candidate's actual performance; the assessment strategy went through an evolution process. Consequently, the assessment process became standard, and the number of variable affective students' performance was reduced and paved the path favouring the introduction of objective structured clinical examination (OSCE), the "gold standard" for clinical

assessments globally. It aims to examine the skills and ability to understand that the assessment results will reflect the trainee physicians' day-to-day clinical performance in real-life scenarios.

In various research studies, OSCEs have been helpful in terms of reliability and validity(5). However, the long examination time is a cause of concern for the trainee physicians and costs to the program directors (6). OSCE mainly focuses on assessing affective, cognitive, and psychomotor learning domains. Nonetheless, performance is affected by various other factors such as knowledge to apply in real-life scenarios; non-clinical skills (decision-making, teamwork, resource management, planning, and critical thinking); attitudes; environment; emotional state; physical state; and personality traits. The drawback of OSCE is that it cannot be easily used to measure non-clinical skills (2)

In clinical education, critical thinking skills are measured by high-fidelity patient simulations, "California critical thinking skills test, California critical thinking disposition inventory, Del Bueno's performance-based development system, health science reasoning test", and Watson-Glaser critical thinking appraisal. However, these are limited in their lack of measuring the particular aspect of health profession-related attributes, inability to evaluate medical professionals' practical reality, and effectively assess psychometric properties (7, 8).

Critical appraisal is a subcategory of critical thinking, which specifies the ability to make clinical decisions by research evidence. Various studies have concluded that critical thinking can be refined, and without this essential ability, there can be drastically negative ramifications on trainee physicians' decisions. There has been evidence of a direct correlation between critical thinking and academic success; unfortunately, many trainee physicians struggle on tests explicitly measuring it (9). While evaluating critical thinking strategies, there is a prominent spotlight on evidence-based practice and its role in education. Numerous systematic reviews have inferred that clinically integrated assessment methods are needed to improve further evidence-based practice skills (10). Critical appraisal has been incorporated into some of the high-stake professional and fellowship examinations where physicians are tested to assess their ability to judge a clinical paper in a short time based on its research design, result and whether to consider this to change one's clinical practice.

However, there is no evidence in the literature reporting the development and implementation of an evaluation tool for assessing clinical skills, analytical thinking, and non-clinical skills in the NSA educational program. Therefore, the current study aims to develop and validate an integrated objective structured clinical examination (i-OSCE) by integrating clinical and critical thinking stations for the NSA postgraduate program.

Methods

OSCE Station Blueprinting

Blueprinting is the standard process of mapping the intended learning outcomes, which comprise knowledge and understanding, intellectual, practical, affective, and psychomotor skills relating to the

postgraduate curriculum on NSA, with the knowledge and skill competencies to be tested in individual stations. An 'expert panel' was formed consisting of five aesthetic practitioners and a senior clinical academic with over fifteen years of experience. A consensual and ceaseless approach was adopted to identify the tasks to be assessed, which are essential and relevant to the NSA practice, thereby validating the content of i-OSCE (Table 1).

Table 1: OSCE Station Blueprinting

Matrix	Station 1 Upper Face	Station 2 Mid / Lower face	Station 3 Total face	Station 4	Station 5	Station 6	Station 7	Station 8
Consultation skills pertaining to the non-surgical facial aesthetics	▓							
Awareness of the relationship between the signs of ageing and the underlying anatomy	▓							
Assessment of Skin quality	▓							
Conduct a full-face assessment to identify treatment needs for optimal results	▓	▓						
Clinical Photography		▓						
Develop a treatment plan that which will deliver the optimal effect and most natural-looking results with the selection of appropriate product		▓	▓					
Safe Injection technique		▓						
Post-treatment Advice		▓						
Complication Management				▓				
Situational Judgement				▓				
Critical understanding of the research relating to facial assessment					▓			
Critical understanding of the research relating to botulinum toxin						▓		
Critical understanding of the research relating to dermal fillers							▓	
Critical understanding of the research relating to complication management								▓

Content development and Validation

A 2-day OSCE writing workshop was conducted for 15 aesthetic practitioners and clinical academicians in the presence of 3 expert facilitators, divided into three small groups. After a brief and structured presentation on OSCE, all the groups had a facilitator lead practice session to construct each station's case scenarios. It was followed by critical feedback from the facilitator to the participants. The three working groups met regularly to construct case scenarios, candidate instruction, standardised patient information sheet, and, most importantly, the marking sheet where the entire scenario was deconstructed to make a performance checklist effectively to match the blueprint theme.

Finally, the expert panel was reconvened again to review the constructed cases with the checklists and parity of competencies across the cases. Fifty stations (25 Clinical and 25 Critical thinking) were selected to store in the repository managed by an OSCE administrator.

Clinical Stations

These stations consisted of a simulated scenario of consultation in facial aesthetics. Here, candidates must take an appropriate history, clinical photography, and facial assessment to reach a specific,

accurate diagnosis. Either standardised patients or patient actors were utilised within these stations. Any clinical examination skills relevant to facial aesthetics were subsidised to fit the station's time limit. Candidates were given a brief history and asked to perform (either all or some aspects) a clinical examination and discuss it with the patient. At the end of the station, candidates were subsequently asked to summarise their findings or provide a brief management plan, including its justification, to the examiner. The other stations developed would assess candidates' professionalism and communication skills.

Critical Thinking (CT) Stations

Twenty-five CT stations were created, where candidates were asked to critically appraise its validity and reliability (formulating PICO, review methodology and critical analysis of the discussion), whether the article published in a peer-reviewed journal and decide to adopt this into clinical practice (applicability).

Marking scheme

Every separate checklist score was weighed based on the allotted task's significance as the station author deemed, which later reached further agreement from an expert panel in a station review meeting. Finally, each station received an independent standardisation to create the pass marks with the help of the borderline regression method, which utilised a combination of the checklist score and the examiner's single 3 points global rating (clear pass, borderline, or clear fail) (Table 2).

Table 2: Example of a checklist/marking scheme

Candidate's Name		Circuit		Examiner's Code	
Station: History taking, Facial assessment and formulating a management plan for the patient					
(2=does well, 1= adequately, 0=poorly or not done)					

Part 1- Initial Approach to the patient			
Introduction	2	1	0
Informed Consent Taken		1	0
Part 2- Information Gathering: Clinical content			
Overall assessment considering medical History, lifestyle	2	1	0
Assessment of the Skin Quality	2	1	0
Full face assessment	2	1	0
Clinical Photography	2	1	0
Part 3 – Information Gathering: communication skills			
Determine for which area(s) is the treatment being sought	2	1	0
Check the candidate's understanding of the treatments available	2	1	0
Educate the patient on the change in their skin/bone/muscle anatomy and the need for a long-term treatment plan	2	1	0
Inform the patient on procedures available		1	0
Create a treatment plan to deliver the best results	2	1	0
Discuss potential risks and benefits of treatment so that they can make an informed decision	2	1	0
Part 4 – Safe procedural practice and Post-treatment Advice			
Selection of appropriate treatment option	2		0
Appropriate injection technique	2	1	0
Post-treatment advice given		1	0
Part 5- Rapport and professionalism			
Polite and courteous and maintain a good rapport throughout	2	1	0
Carries out examination fluently and professionally	2	1	0
SP Mark	2	1	0
Examiner's Comments			
Overall Judgement (Global Scoring)			

Clear Pass		Borderline		Clear Fail	

Assessor Training

All the examiners participated in an hour-long orientation program to familiarise themselves with the OSCE setting, competency testing, and scoring guidelines. Further, they were provided with a guide describing the definitions of the competencies, checklist, global ratings.

Standardised Patient (SP) Training

Healthy volunteers were recruited with the help of a modelling agency to act as 'simulated' patients for all the stations. They went through coaching conducted by professional medical actors and clinical and communication skills experts. SPs were given a task, particularly to the station, and they practised until they played their roles consistently. As they were also responsible for completing their part of the checklist, a calibration video was shown to practice marking and debriefing.

Pilot Study

A pilot study was conducted with ten aesthetic practitioners to examine the feasibility of i-OSCE. Planned eight stations were run with one examiner and one observer. Finally, the results were reviewed, including the feedback from individual stations to amend i-OSCE documents for final implementation during the summative examination.

Final i-OSCE

The final examination was conducted comprised four clinical and four critical thinking stations of 15 minutes each and four rest stations with forty trainee physicians. The total run for the exam was 180 minutes.

Statistical Analysis

For an assessment tool to be accepted as reliable and valid, the most widely used statistical measurement is Cronbach's alpha (11). However, some studies argue that it should not be used for internal reliability as sole measurement, as it is directly proportionate to the examination length; therefore, it indicates the station's stability, not the internal consistency (12,13). However, concurrent use of Pearson correlation coefficient (PCC) or Spearman's rank correlation helps overcome the issue (14). Therefore, PCC was used to investigate the strength of correlation between utilising the checklist and the global rating (clear pass, borderline, or clear fail), which helped provide a measure of the validity of the marking criteria used. For calculating the interrater reliability (IRR), Cronbach's alpha was used through two-way mixed effects; intra-class correlations (ICC) for consistency and internal reliability. For interpretation of ICCs, Cicchetti's classification (IRR less than 0.40 is poor; 0.40−0.59 is fair; 0.60−0.74 is good; 0.75−1.00 is excellent) was used (15). Moreover, content validity was measured with the help of experts. Statistical analysis was carried out by using IBM SPSS Statistics for Mac, Version 27.0 (IBM Corp. Armonk, NY, USA).

Results

The Interclass correlation coefficient between the station checklist items was 0.946 (average measure upper bound 0.916, lower bound 0.968; p<0.00), considered to be significant (Table 3). The Inter-Item Correlation Matrix among the clinical station checklist items and critical thinking checklists also showed statistical significance (Table 4). The Pearson correlation coefficient (PCC) used to ascertain the correlation between checklist rating and global rating (Table 5), yielding a high correlation (0.80 to 0.934).

Table 3: Intraclass Correlation Coefficient

	Intraclass Correlation[b]	95% Confidence Interval		F Test with True ...	
		Lower Bound	Upper Bound	Value	df1
Single Measures	.688[a]	.578	.793	18.611	38
Average Measures	.946[c]	.916	.968	18.611	38

Table 4: Inter-Item Correlation Matrix of station checklist

Inter-Item Correlation Matrix								
	Station_1	Station_2	Station_3	Station_4	Station_5	Station_6	Station_7	Station_8
Station_1	1	0.722	0.744	0.819	0.629	0.674	0.836	0.656
Station_2	0.722	1	0.81	0.726	0.675	0.601	0.75	0.725
Station_3	0.744	0.81	1	0.837	0.579	0.556	0.674	0.607
Station_4	0.819	0.726	0.837	1	0.43	0.503	0.633	0.446
Station_5	0.629	0.675	0.579	0.43	1	0.84	0.799	0.859
Station_6	0.674	0.601	0.556	0.503	0.84	1	0.874	0.803
Station_7	0.836	0.75	0.674	0.633	0.799	0.874	1	0.86
Station_8	0.656	0.725	0.607	0.446	0.859	0.803	0.86	1

Table 5: Checklist Vs Global Assessment Correlation

Station	Pearson Correlation
1	0.865
2	0.82
3	0.934
4	0.8
5	0.856
6	0.848
7	0.891
8	0.875

Discussion

This is the first integrated OSCE validation study combining clinical and critical thinking skills for a postgraduate NSA education to the best of the author's knowledge. OSCE is a flexible assessment method use to evaluate competence by direct observation based on objective assessment criteria. It is composed of several "stations" where examinees are required to conduct a range of clinical tasks against required clinical competence displaying the skills and attitudes over a given duration. The OSCE has been used to assess the skills most important to the healthcare professionals' success, such as data acquisition, interpretation, troubleshooting, engagement, and management of erratic patient behaviour, otherwise difficult to obtain during the classic clinical review (16). Miller's framework for clinical competency development recommended four stages; "knows the facts"; "knows how to elaborate and integrate the understanding"; "shows how" they apply knowledge, skills and attitude for the patient outcome; and finally "does" employ all the skills in their independent practice to serve the community, proven to work reasonably well in medical education settings (17, 18). Evidence suggests that the OSCE helps assess the third stage "shows how" by concentrating on the clinical skills in a safe learning environment.

Critical thinking is considered to be a crucial cognitive method for the creation and utilisation of knowledge. It plays a functional role during problem-solving and decision-making in a social, clinical, or ethical context. Moreover, it is equally valuable for analysing complex data, assessing situations, and implementing the most suitable actions. In a recent article, "critical thinking is described as a cognitive process, purposeful, self-regulatory judgment that has two components of cognitive skills (interpretation, analysis, inference, evaluation, explanation, and self-regulation) and a motivational component (the disposition toward critical thinking)" (19).

More focus has been put in recent years on improving higher-level thought (critical thinking and clinical reasoning) skills to help physicians retain clinical integrity and medical professionalism. More than two-thirds of the reported mistakes in diagnosis are linked to physician's lack of critical thinking ability in the present context. Given the belief that healthcare professionals must be logical thinkers, there is no consensus on the most successful model to teach and evaluate critical thinking and clinical reasoning skills (8). Recent research, which evaluated a wide range of quantitative and qualitative competencies, including behavioural and communication skills, showed that the OSCE was valid and reliable and essential for positive educational effects. Several authors have advocated that emphasising an aim of OSCEs is to develop affability in critical thinking as a precursor to practising (20).

There is no valid assessment instrument combining clinical and critical appraisal skills to evaluate safe practice in non-surgical aesthetics. Therefore, using different stations to evaluate clinical skills and critical thinking ability is beneficial in this evaluation. The various clinical skills measured were consultation skills pertaining to the NSA, knowledge of the signs of ageing and the underlying anatomy, assessment of skin quality, full-face assessment to identify treatment needs for optimal results, clinical photography, development of an efficient and optimal treatment plan, safer injection techniques, post-

treatment advice, complication management, and situational judgments. During the evaluation, the critical appraisal skills measured included understanding the research relating to facial assessment, botulinum toxin science, rheology of the soft-tissue fillers and complication management.

It has shown that the generalisation coefficients appear to differ significantly from 0.40 to 0.85, while the majority of these coefficients vary from 0.5–0.6 (6). In the present study, the average intraclass correlation coefficient measures range between 0.916 to 0.968, which is more than the reliability coefficient threshold of 0.8 or over. The variability in the generalizability coefficients may be attributed to the examinees' variable performance on different OSCE stations (content specificity). I-OSCE is shown to be robust and able to test applicants for their competence to carry out multiple component tasks.

Conclusion

Integrated OSCE has demonstrated to be a reliable and accurate assessment tool for examining the trainee aesthetic physicians' professional competence. This tool has objectively evaluated trainee physicians critical thinking and clinical skills, including clinical reasoning. The program directors should consider the deployment of i-OSCE along with OSPE as an assessment tool in the postgraduate curriculum for non-surgical aesthetics.

Abbreviations

NSA
non-surgical aesthetics
OSCE
Objective structured clinical examination
I-OSCE
Integrated Objective structured clinical examination
OSPE
Objective structured practical examination
SP
Standardised patient

Declarations

Ethics Approval and Consent to Participate

The study was approved by the Faculty Research Ethics Panel of the Faculty of Medical Science, Anglia Ruskin University, Chelmsford, UK.

Consent for publication: Not applicable

References

1.	Tabish SA. Assessment methods in medical education. Int J Health Sci (Qassim) [Internet]. 2008 Jul;2(2):3–7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21475483

2.	Khan KZ, Gaunt K, Ramachandran S, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE Guide No. 81. Part II: Organisation & Administration. Med Teach [Internet]. 2013 Sep 22;35(9):e1447–63. Available from:
http://www.tandfonline.com/doi/full/10.3109/0142159X.2013.818635

3.	Tey C, Chiavaroli N, Ryan A. Perceived educational impact of the medical student long case: a qualitative study. BMC Med Educ [Internet]. 2020 Dec 7;20(1):257. Available from:
https://bmcmededuc.biomedcentral.com/articles/10.1186/s12909-020-02182-6

4.	Pell G, Roberts TE. Setting standards for student assessment. Int J Res Method Educ [Internet]. 2006 Apr;29(1):91–103. Available from:
http://www.tandfonline.com/doi/abs/10.1080/01406720500537486

5.	Kumar N, Bhardwaj S, Rahman E. Multiple mini-interview as a predictor of performance in the objective structured clinical examination among Physician Associates in the United Kingdom : a cohort study. Adv Med Educ Pract. 2018;9:239–45.

6.	Marwaha S. Objective Structured Clinical Examinations (OSCEs), psychiatry and the Clinical assessment of Skills and Competencies (CASC)Same Evidence, Different Judgement. BMC Psychiatry [Internet]. 2011 Dec 16;11(1):85. Available from:
http://bmcpsychiatry.biomedcentral.com/articles/10.1186/1471-244X-11-85

7.	Nguyen K, Ben Khallouq B, Schuster A, Beevers C, Dil N, Kay D, et al. Developing a tool for observing group critical thinking skills in first-year medical students: a pilot study using physiology-based, high-

fidelity patient simulations. Adv Physiol Educ [Internet]. 2017 Dec 1;41(4):604–11. Available from: https://www.physiology.org/doi/10.1152/advan.00126.2017

8.   Majumder MAA, Sa B, Alateeq FA, Rahman S. Teaching and Assessing Critical Thinking and Clinical Reasoning Skills in Medical Education. In 2019. p. 213–33. Available from: http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-5225-7829-1.ch012

9.   West DC, Pomeroy JR, Park JK, Gerstenberger EA, Sandoval J. Critical Thinking in Graduate Medical Education. JAMA [Internet]. 2000 Sep 6;284(9):1105. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.284.9.1105

10.   Maudsley G, Strivens J. Promoting professional knowledge, experiential learning and critical thinking for medical students. Med Educ. 2000 Jul;34(7):535–44.

11.   Swanson DB, Stillman PL. Use of Standardized Patients for Teaching And Assessing Clinical Skills. Eval Health Prof [Internet]. 1990 Mar 29;13(1):79–103. Available from: http://journals.sagepub.com/doi/10.1177/016327879001300105

12.   Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika [Internet]. 1951 Sep;16(3):297–334. Available from: http://link.springer.com/10.1007/BF02310555

13.   Streiner DL. Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency. J Pers Assess [Internet]. 2003 Feb;80(1):99–103. Available from: http://www.tandfonline.com/doi/abs/10.1207/S15327752JPA8001_18

14.   Graham JM. Congeneric and (Essentially) Tau-Equivalent Estimates of Score Reliability. Educ Psychol Meas [Internet]. 2006 Dec 2;66(6):930–44. Available from: http://journals.sagepub.com/doi/10.1177/0013164406288165

15.   Al-Osail AM, Al-Sheikh MH, Al-Osail EM, Al-Ghamdi MA, Al-Hawas AM, Al-Bahussain AS, et al. Is Cronbach's alpha sufficient for assessing the reliability of the OSCE for an internal medicine course? BMC Res Notes [Internet]. 2015 Dec 19;8(1):582. Available from: http://www.biomedcentral.com/1756-0500/8/582

16.   Zayyan M. Objective Structured Clinical Examination: The Assessment of Choice. Oman Med J [Internet]. 2011 Jul 25;26(4):219–22. Available from: http://www.omjournal.org/fultext_PDF.aspx?DetailsID=114&pdf=images/114_M_Deatials_Pdf_.pdf&type=pdf

17.   Miller GE. The assessment of clinical skills/competence/performance. Acad Med [Internet]. 1990 Sep;65(9):S63-7. Available from: http://journals.lww.com/00001888-199009000-00045

18.   Sousa AC, Wagner DP, Henry RC, Mavis BE. Better data for teachers, better data for learners, better patient care: college-wide assessment at Michigan State University's College of Human Medicine. Med

Educ Online [Internet]. 2011 Jan 14;16(1):5926. Available from: https://www.tandfonline.com/doi/full/10.3402/meo.v16i0.5926

19. Nematbakhsh M, Sharif S-M, Athari Z-S, Nasr A. Assessing critical thinking in medical sciences students in two sequential semesters: Does it improve? J Educ Health Promot. 2013;2(1):5.

20. Krusen NE, Rollins D. Design of an OSCE to Assess Clinical Competence of Occupational Therapy Students. J Occup Ther Educ [Internet]. 2019 Jan 1;3(1). Available from: https://encompass.eku.edu/jote/vol3/iss1/11