

Back to the Future: Electronic Marking of Objective Structured Clinical Examinations and Admission Interviews Using an Online Management Information System in Schools of Health Sciences

Thomas JB Kropmans^{1*}, Liam Griffin², David Cunningham², Domhnall Walsh², Winny Setyonugroho³, Catherine Anne Field⁴, Eva Flynn⁵ and Kieran M Kennedy⁶

¹College of Medicine, Nursing and Health Sciences; School of Medicine; discipline of Medicine, domain Medical Informatics & Medical Education National University of Ireland Galway, Ireland

²Qpercom Ltd, Qualitative Performance and Competence Measures, Ireland

³PhD student School of Medicine/Discipline of Medicine, Medical Informatics and Medical Education, Ireland

⁴School of Medicine/Discipline of Health Promotion, Ireland

⁵College of Medicine, Nursing and Health Sciences; School of Medicine; discipline of General Practice, National University of Ireland Galway, Ireland

⁶ College of Medicine, Nursing and Health Sciences; School of Medicine; discipline of Medicine, National University of Ireland Galway, Ireland

Abstract

Background: The Objective Structured Clinical Examination (OSCE) and Multi Mini Interviews (MMI) are established tools in the repertoire of clinical assessment methods in Schools of Medicine and Health Sciences worldwide. The use of OSCEs facilitates the assessment of psychomotor skills as well as knowledge and attitudes. Identified benefits of OSCE assessment include development of students' confidence in their clinical skills and preparation for clinical practice. However, a number of challenges exist with the traditional paper methodology, including documentation errors and inadequate student feedback, electronic assessment is therefore new future.

Objectives: To explore electronic OSCE delivery and evaluate the benefits of using an electronic OSCE management system.

Design: A pilot study was conducted using electronic software in the management of a five station OSCE assessment with a cohort of first year undergraduate medical students delivered over two consecutive years (n=383) in one higher education institution in Ireland.

Methods: All OSCE documentation was converted to electronic format. Assessors were trained in the use of the OSCE management software package and laptops were procured to facilitate electronic management of the OSCE assessment. Following the OSCE assessment, assessors were invited to evaluate the experience.

Results: Electronic software facilitated the storage and analysis of overall group and individual results thereby offering considerable time savings. Submission of electronic forms was allowed only when fully completed thus removing the potential for missing data.

Conclusions: Analysis of results highlights issues around inter-rater reliability and validity of measurement tools. Regression analysis, as a standard setting method, increases fairness of result calculations as compared to static cut-off scores.

Keywords: Objective structured clinical examination; OSCE; Multi mini interview; MMI; e-Assessment; Borderline regression analysis; Generalizability theory

Introduction

The Objective Structured Clinical Skills Examination (OSCE) is a well-established method of assessing non-cognitive skills in multi mini interviews and clinical competence in so called OSCEs among health practitioners or students that apply for a health sciences degree [1]. The OSCE originated in the UK as an objective means to assess medical students' skills [2]. The examination involves students progressing through a series of stations where they are assessed by an examiner with pre-determined marking criteria [3]. In these series of stations either clinical tasks and/or non-cognitive skills for admission interviews are being assessed.

Several authors have highlighted the importance of using OSCEs as an assessment method in health care education [1,4-6]. The OSCE facilitates the assessment of student's competency with clinical skills in a controlled simulated environment instead of in the practice setting [5]. According to McWilliam and Botwinski, students recognise the value of the OSCE experience to their education [4].

A number of benefits have been attributed to the use of OSCEs including, the development of students confidence [7], the preparation of students for clinical practice and the achievement of deeper more meaningful learning [6]. Importantly, the use of OSCEs facilitates the assessment of psychomotor skills as well as knowledge and attitudes [5] OSCEs provide students with feedback on their clinical performance and facilitate the identification of strengths and weaknesses [4]. The OSCE has been reviewed positively as an assessment method for clinical

***Corresponding author:** Thomas JB Kropmans, College of Medicine, Nursing and Health Sciences; School of Medicine; Domain Medical Informatics and Medical Education, Ireland, Tel: +3539152441; E-mail: thomas.kropmans@nuigalway.ie

Received January 19, 2015; **Accepted** February 15, 2015; **Published** February 18, 2015

Citation: Kropmans TJB, Griffin L, Cunningham D, Walsh D, Setyonugroho W, et al. (2015) Back to the Future: Electronic Marking of Objective Structured Clinical Examinations and Admission Interviews Using an Online Management Information System in Schools of Health Sciences. J Health Med Informat 6: 182. doi:10.4172/2157-7420.1000182

Copyright: © 2015 Kropmans TJB, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

competence and for responding to student diversity in education [8]. However, there are a number of notable disadvantages associated with OSCEs. In particular, some students find them stressful and they are resource intensive in terms of staff, equipment and clinical skills laboratories [5]. However, Alinier [7] suggests that the educational benefits surpass the issues associated with resources [7].

Traditionally OSCEs have been assessed with paper based methods. However, a number of issues have been highlighted with this method including illegible handwriting, missing details (students' names and student numbers) and lost assessment sheets [9]. Furthermore, it is known that manual calculation of results and entering them into a database is time-consuming and is subject to human errors and feedback is rarely provided to students on their performance after paper based assessments [9]. Despite these issues there is a scarcity of literature regarding the use of computer or OSCE software and the assessment of OSCEs. Segall et al [10] compared the usability of paper and pencil method and Personal Digital Assistant (PDA) based quizzes and found the PDA based quiz was more efficient and superior to the traditional based method [10]. Similarly, Treadwell [9] compared the conduction of a paper based OSCEs with an electronic method. The findings indicated that the electronic method was just as effective and more efficient (less time consuming) than the traditional paper based method. In addition, the electronic system was highly rated by the assessors, who found it less invasive and reported they had more time to observe the students and permitted greater observation of the students when using the paper assessment. Schmitz [11] highlights a number of advantages to use an electronic handheld device to assess OSCEs including, speed of data gathering, simplicity of data evaluation and fast automatic feedback [11]. Segall et al [10] support computer based assessment suggesting that grading is more accurate, feedback is immediate, security is enhanced and less time is spent by instructors on grading and data entry [10]. Cunningham and Kropmans developed an OSCE Management Information System (OMIS) that is currently used within 19 prestigious universities worldwide to retrieve, store and analyse all OSCE and MMI data. The aim of this study was to explore the benefits of an online OSCE Management Information System for School of Medicine OSCEs, by means of the analysis of two cohort studies [12,13].

Methods

In this cross sectional study we analysed the outcome of a fully fledged electronic administered Objective Structured Clinical Examination (OSCE), of two cohorts of students assessing the clinical outcome of the first year MD139 module Medical Professionalism using an in house developed online OSCE Management Information System (OMIS) [12]. This MD139 OSCEs comprised of 5 individual stations. Both consecutive student cohorts (i.e. those from the 2012-2013 and 2013-2014 academic years) completed a urine analysis station, a chest X-ray station, a BMI and Vital signs station and finally a Basic Life Support station, each of which was of 5 minutes duration. The total number of first year students that completed the OSCE was 383. The 2012-2013 cohorts comprised 213 students, whereas the 2013-2014 cohorts comprised 170 students. The station checklists for both OSCEs were identical. The novel online OSCE Management Information System, which was developed "in-house" at the National University of Ireland Galway, was used to administer both examinations [12,13]. OMIS retrieves stores and analyses assessment data electronically. Student feedback can be sent to students electronically using the Student Feedback Email System. We used item checklists to assess student competency with each task. The number of items per

assessment form varied from 13, for the Basic Life Support station, to 22, for the Urine analysis and for Vital signs stations, with a maximum score of 60 marks for all clinical stations. The overall professional impression of the examiners (Global Rating Scale) was rated on a 5 item Likert Global Rating Scale (GRS) which included the following options: Fail (0), Borderline (1), Pass (2), Good (3) and Excellent (4). The numerical values of the GRS options were not incorporated in the final student scores, but were instead used for standard setting using an online Borderline Regression Analysis function that is built in to the OMIS. The static pre-determined cut-off score for medicine studies is 50% (<50% means a fail score; >50% means a pass score).

Statistical Analysis

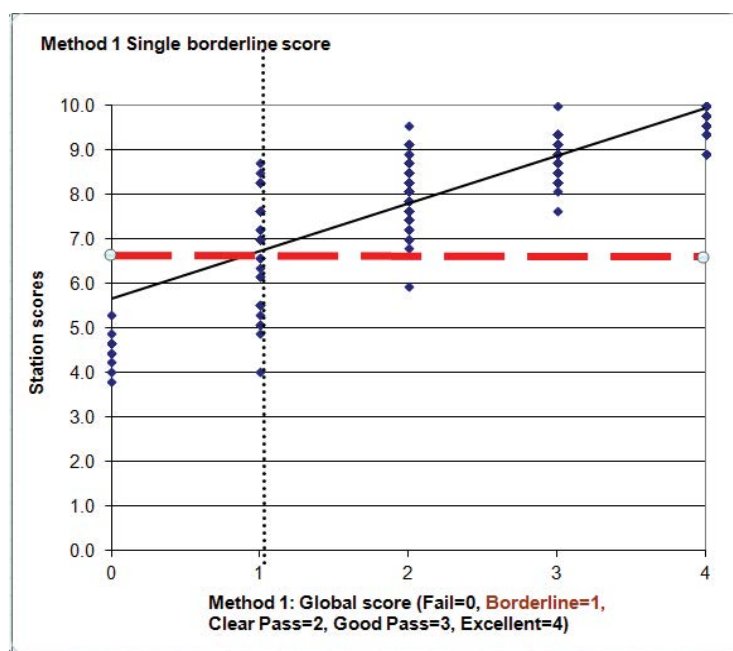
OMIS produces an online analysis of items and overall total (raw) scores and adjusted (raw) scores using standard setting of student performance after regression analysis. The mean result, standard deviation (SD), minimum and maximum and range and mid range are produced instantly, in real time, during the examination. Internal consistency of OSCE station item forms (Cronbach's Alpha) is used to provide insight into the consistency of items in each station predicting the overall score of the student of that specific station. Borderline Regression Analysis (Borderline Group Average versus Borderline Regression Method) calculates a 'flexible cut-off score' complementary to the general static cut score of 50% for each individual station. The overall average regression cut-score is used to adjust the average overall raw score of the students. Borderline Group Average, which is based upon calculation of the average mark of those students that were globally rated by their examiners as 'borderline', is the most simplistic method to use [14]. A complete Borderline Regression Analysis, which is performed over all item marks matched with all of the global ratings (from fail to excellent), can also be used. The flexible cut-off score is calculated using the BRM Cut score (Intercept+1 × Slope)-since borderline=1 using FORECAST method, (Figure 1). All analysis reports and data were exported to Excel to facilitate further detailed analysis.

Data were exported to perform a Generalizability Coefficient analysis using a G- and D-study with EduG software. The G-study generates information about whether the outcome can be generalised to other medicine OSCEs. The D-study provides information on how the generalizability of results can be improved [15].

Results

The summary of results for the 2012-2013 cohort (n=213), as produced by the OMIS, demonstrated an overall internal consistency (*how well predicts the OSCE the overall outcome*) of 0.696, where the Cronbach's Alpha per individual station varied between 0.486 for the Basic Life Support station and 0.769 for the Vital Signs station. In classical psychometric terms internal consistency was moderate [16]. The overall average student performance for the clinical stations was 80.5%, with a minimum score of 16 out of 60 and a maximum of 100% (60 out of 60). The overall average student performance for the Basic Life Support station was 93.7% , with a minimum score of 21 out of 60, and a maximum score of 60 out 60)(Figure 2).

The summary results for the 2013-2014 cohort (n=170), as produced by OMIS, demonstrated an overall internal consistency of 0.666, whereby the Cronbach's Alpha per station varied between 0.426 for the Basic Life Support station and 0.842 for the Vital Signs station. In classical psychometrics terms internal consistency was moderate. The overall average performance (SD) of the students was



Dynamic cut-off score

Figure 1: Single borderline score regression analysis illustrating the effect of a regression analysis in which station scores on the Y-axis are outlined against the professional opinion of the examiners (X-axis) (adjusted from John Patterson, honorary senior lecturer at the Centre for Medical Education of the Barts and London School of Medicine and Dentistry and Assessment Consultant).

Setup	Raw Data	Summary	Detailed	BL Groups	Graphs	Export	All						
		Nr. of Results	Stations Passed	Stations Failed	Mean Result	Mean Result (%)	SD	SD (%)	Min	Max	Range	Mid Range	Station Items Cronbach Alpha (α)
1 - BMI 2011		213	211	2	48.9 / 60	81.6%	5.9	9.9%	26.0	60.0	34.0	43.0	0.757
2 - Urinalysis 2011		213	210	3	44.4 / 60	73.9%	6.1	10.2%	28.0	57.0	29.0	42.5	0.728
3 - Basic life support 2013		213	211	2	56.2 / 60	93.7%	5.2	8.6%	21.0	60.0	39.0	40.5	0.486
4 - Vital Signs 2011		213	213	0	48.2 / 60	80.4%	5.1	8.5%	31.0	60.0	29.0	45.5	0.769
5 - Chest X-Ray 2011		213	204	9	43.7 / 60	72.8%	7.2	11.9%	16.0	60.0	44.0	38.0	0.739
Exam Totals & Mean %		1065	1049	16		80.5%							0.696
Total Students Passed Failed			213 (100%)	0 (0.0%)									

Figure 2: Screenshot OMIS: OSCE Results summary table 2012-13 (OMIS 1.8.5).

84.8% for the clinical stations, with a minimum score of 20 out of 60 and a maximum score of 100% (60 out of 60). The overall average performance of students was 88.5% for the Basic Life Support station, with a minimum score of 28 out of 60 and a maximum score 100% (60 out of 60). In classical psychometric terms internal consistency was moderate (Figure 3).

Borderline regression analysis

Borderline Group Analysis is a simple way of calculating the average cut-off score of those students that were addressed as 'borderline performers' (i.e. examiners not being sure whether the student performance should be marked as fail or pass). Where there are a small number of students in this category, then Borderline Group Average estimates may be very unreliable as shown in the 2012-2013 cohort, where in some stations only 1 or 6 student performances were 'marked as 'borderline' (i.e. 1 borderline scores for the Vital Signs station and 6 for the Basic Life Support station). The 'average' score of these students was 60.0% (1 student) for the Vital Signs station and 70.3% (6 students) in the Basic Life Support station. Using Borderline Group Average in this cohort would not provide similar information for those

stations where *no* students were marked as borderline performers. A fully fledged Borderline Regression Analysis would however provide this information due to the inclusion of all Global Ratings from Fail to Excellent (Figure 1).

A similar situation arises with the Borderline Group Analysis of the 2013-2014 cohort, within which 42 student performances were regarded as borderline (i.e. 9 for urine analysis, 0 for the Chest X-ray station, 4 for the BMI station, 11 for the Vital Sign station and 18 for the Basic Life Support station). Due to the small numbers of students, cut-off scores of 67.8%, N/A, 63.8%, 69.2 and 81.5% may be unreliable but provide an indication about the difficulty of each station (BGS>50 means easy station; BGS<50 means a difficult station).

A fully fledged Borderline Regression Analysis is embedded into the OSCE Management Information System software whereby the forecast method is used to calculate new cut-off scores for each station taking into account the 'difficulty of the station' and the 'hawk and dove effect' of different examiners involved in the OSCE.

Figure 4 shows the item maximum scores for each of the four OSCE

	Nr. of Results	Stations Passed	Stations Failed	Mean Result	Mean Result (%)	SD	SD (%)	Min	Max	Range	Mid Range	Station Items Cronbach Alpha (α)
1 - Urinalysis 2011	170	169	1	49.2 / 60	82.1%	5.5	9.3%	20.0	58.0	38.0	39.0	0.769
2 - Chest X-Ray 2011	170	168	2	50.4 / 60	84.0%	7.4	12.3%	25.0	60.0	35.0	42.5	0.600
3 - BMI 2011	170	168	2	51.4 / 60	85.6%	4.8	8.1%	27.0	60.0	33.0	43.5	0.691
4 - Vital Signs 2011	170	163	7	50.3 / 60	83.8%	8.2	13.6%	23.0	60.0	37.0	41.5	0.842
5 - Basic Life Support 2014	170	169	1	53.1 / 60	88.5%	5.8	9.7%	28.0	60.0	32.0	44.0	0.426
Exam Totals & Mean %	850	837	13		84.8%							0.666
Total Students Passed Failed		170 (100%)	0 (0.0%)									

Figure 3: Screenshot OMIS: OSCE Results summary table 2013 - 14 (OMIS 1.8.5).

Station	Item Maximum Scores																														Total	Raw Station Scores	Station SD	No of Borderline	A	B
	H1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	I13	I14	I15	I16	I17	I18	I19	I20	I21	I22	I23	I24	I25	I26	I27	I28	I29	I30						
1 BMI 2011	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	60	8.2	1.0	14	6.6	6.5
2 Urinalysis 2011	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	60	7.4	1.0	18	5.7	5.9
3 Basic life support 2013	1	2	1	2	2	2	4	4	4	4	3	1	30																		60	9.4	0.9	6	7.0	7.2
4 Vital Signs 2011	2	2	2	2	2	2	2	2	2	2	2	10	1	1	3	1	6	1	1	2	2	2	10								60	8.0	0.8	1	6.0	6.3
5 Chest X-Ray 2011	1	1	3	3	4	2	2	2	2	2	3	3	2	30																	60	7.3	1.2	18	5.2	5.3

Figure 4: Screenshot OMIS: Borderline regression summary table for Cohort 2012-2013 showing Item Scores (I1-I30); Item Total Scores (60); Raw Station Scores and Standard Deviation (SD); the number of students achieving a Borderline Score and finally the Borderline Group Average score (BRA) and Borderline Regression score (BR method 1 and Figure 1).

stations (60 for each of the clinical stations), along with each station’s mean score (out of 10, EU based) and standard deviation. Borderline Group Average could be calculated for all station. With Borderline Regression Method 1, cut-off scores are calculated for all stations based upon analysing item scores and Global Rating Scores of all students varying respectively from 72% in station 3 (was 70% in case of a ‘group’ average) to 53% for station 5.

Discussion

The summary of results section provided instant information about the scores of each individual student from two different cohorts in these two consecutive OSCEs. Although the average results are quite high in both cohorts 16 students in the first and 13 in second cohort failed in one or two of the consecutive stations using a ‘static pass mark’ set prior to the start of the exam at 50%. Due to the availability of a Global Rating Score facility, and an appropriate number of students (n ≥ 100), we performed a Borderline Group Average analysis. The latter is based upon the overall professional impression of the examiner evaluating a student’s performance and incorporates the difficulty of the stations and the variability within examiners. The examiner will mark this overall performance as a pass, borderline, fail, good, or excellent performance (Borderline Regression Method 1). In the borderline group feature (Figures 5 and 6), the average performance of these ‘borderline performing students’ is substantially above the static pass mark of 50% for all in stations of both cohorts of students. Where N/A is indicated this means that no students were marked as Borderline performing students in the second cohort. ‘Borderline performance’ is an indicator of examiner uncertainty with regard to whether or not a student should pass or fail. Whether the regression outcome is high or low, it is an indicator of whether a station is ‘easy’ or ‘difficult’ to pass respectively. Where the ‘Angoff Method’ is a standard setting method used *prior to* an examination, Borderline Regression Analysis is a standard setting method used *after* the examination has taken place and is based upon the professional impression of the examiner

evaluating students’ performances according to a Global Rating Scale [17]. In addition to the Borderline Group Average, OMIS provides a fully fledged borderline regression analysis that takes all scores into account and matches those with the professional impression of the examiners using a regression analysis (Figure 1). We used the simple forecast method, in an Excel template, using all item total marks and the Global Rating Scale in the regression equation [18]. All station dynamic cut-off scores were above the 50% cut-off score indicating that stations were too easy to pass according to the professional impression of the examiners.

The OSCE design used for this clinical skills examination of undergraduate medical students demonstrated poor generalizability of results in this 5 station OSCE. The generalizability would improve by introducing more stations (e.g. an OSCE with 5-10 stations). However, current coefficients do not achieve the standards suggested in other research literature on the subject e.g. OSCEs [17,19-22] The generalizability of results is only appropriate in OSCEs with a minimum of 15-18 stations [23-25].

Scores of ‘borderline performing students’ were way beyond the ‘static cut-off score’ in all stations indicating the OSCE needs a qualification as ‘easy to pass. Making station designers aware of these high marks and training them on existing pre-recorded scenarios and using well described rubrics might reduce the amount of error and should be the focus of additional research. The benefit of student feedback allowing them an opportunity to benchmark themselves against the group and to get relevant timely feedback on their performance is available in the system but not being used in these cohorts. Future research should focus on the impact of instant feedback on the performance of students and used in future comparisons. Although not the subject of this study, the overall impact on time reduction in running the OSCEs and students’ and examiners’ behavior during assessment are features that need to be further researched. In contrast to our previous paper based approach, results and feedback could be released immediately after the exam was finished.

		Item Maximum Scores																				Item Total Scores	Borderline Group Score	Borderline Group Score/10	Borderline Group Score %	Nr. of Borderline Scores		
		I1	I2	I3	I4	I5	I6	I7	I8	I9	I10	I11	I12	I13	I14	I15	I16	I17	I18	I19	I20	I21	I22					
1 - BMI 2011		2	2	2	2	2	2	2	2	2	2	10	2	2	2	2	2	4	2	2	2	10	60	39.6	6.6	66.0%	14	
2 - Urinalysis 2011		2	2	2	2	2	2	2	2	2	10	1	2	1	2	1	1	2	6	2	2	10	60	34.2	5.7	57.0%	18	
3 - Basic life support 2013		1	2	1	2	2	4	4	4	4	3	1	30	60	42.2	7.0	70.3%	6										
4 - Vital Signs 2011		2	2	2	2	2	2	2	2	2	10	1	1	3	1	6	1	1	2	2	2	10	60	36.0	6.0	60.0%	1	
5 - Chest X-Ray 2011		1	1	3	3	4	2	2	2	2	3	3	2	30	60	31.4	5.2	52.4%	18									

*Layout Copyright © 2013 Dr. J.A Patterson BSc PhD Honorary Senior Lecturer at the Centre for Medical Education Barts and the London School of Medicine and Dentistry

Figure 5: Screenshot OMIS: Borderline Group Average analysis in OMIS for cohort 2012-13.

Borderline Regression Method Used:															Statistical Summary:				
0	Borderline Group Average	1													As	89			
1	Pass mark with Regression Method 1														Bs	93			
2	Pass mark with Regression Method 2														Cs	30			
3	Pass mark with Contrasting Method														Ds	2			
Appropriate values MUST be entered in cells shown in red															Es	5			
Total number of active stations = 5					Min. no of stations to be passed = 1					Total Candidates					219				
(ie Candidates must not fail more than 4 stations)					60.3					Fails on BGT pass mark =					7				
Pass mark with Regression Method 1					6.02	5.6	6.38	6.4	5.7	more than minimum no of stations failed =					0				
Raw Max					60	60	60	60	60	TOTAL FAILS (on one criterion or both) =					0				
Borderline Group Pass Mark per station =					5.9	5.5	6.46	6.4	5.1	BG tot %	60.3	sd-raw	7.0	sd-adj	7.0	Fails on each station			
Station					1	2	3	4	5	PASS % =	60.3	7.0	7.0						
Cand. No.	Identifier	First name	Surname	Raw %	ADJ %	GRADE	PASS/FAIL on Marks	Number of Stations failed PASS/FAIL on Stations (7 or more Borderline Group fails)					OVERALL RESULT						
								BMI 2011 /10	Urinalysis 2011 /10	Basic Life Support 2012 /10	Vital Signs 2011 /10	Chest X-Ray /10							
	*****0301			8.2	9.3	6.3	9.3	8.5	83.3	73.0	A	PASS				1 PASS			
	*****0302			9.7	9.5	9.3	8.0	8.3	89.7	79.4	A	PASS				0 PASS			
	*****0303			9.8	9.7	9.5	5.3	8.2	84.9	74.7	A	PASS		F		1 PASS			
	*****0304			9.3	8.5	6.3	8.7	8.7	82.9	72.7	A	PASS		F		1 PASS			
	*****0305			10.0	9.2	8.0	9.0	8.2	88.7	78.4	A	PASS				0 PASS			
	*****0306			9.2	7.7	10.0	8.5	8.5	87.7	77.4	A	PASS				0 PASS			
	*****0307			9.7	9.7	9.0	7.2	8.3	87.7	77.4	A	PASS				0 PASS			
	*****0308			9.0	9.2	8.8	8.8	7.3	86.3	76.1	A	PASS				0 PASS			
	*****0309			9.5	9.2	8.5	7.5	8.5	86.3	76.1	A	PASS				0 PASS			
	*****0310			9.3	9.7	8.2	7.7	8.2	86.0	75.7	A	PASS				0 PASS			
	*****0311			9.3	8.7	8.8	7.7	8.5	86.0	75.7	A	PASS				0 PASS			

Figure 6: Raw scores and adjusted raw scores according to the 'dynamic cut-off score' after Borderline Regression Analysis adjusted after John Patterson, honorary senior lecturer at the Centre for Medical Education of the Barts and London School of Medicine and Dentistry and Assessment Consultant.

References

- Oranye NO, Ahmad C, Ahmad N, Bakar RA (2012) Assessing nursing clinical skills competence through objective structured clinical examination (OSCE) for open distance learning students in Open University Malaysia. Contemp Nurse 41: 233-241.
- Harden RM, Stevenson M, Downie WW, Wilson GM (1975) Assessment of clinical competence using objective structured examination. Br Med J 1: 447-451.
- Pugh D, Touchie C, Wood TJ, Humphrey-Murto S (2014) Progress testing: is there a role for the OSCE? Med Educ 48: 623-631.
- McWilliam P, Botwinski C (2010) Developing a successful nursing Objective Structured Clinical Examination. J Nurs Educ 49: 36-41.
- Baid H (2011) The objective structured clinical examination within intensive care nursing education. Nurs Crit Care 16: 99-105.
- Barry M, Noonan M, Bradshaw C, Murphy-Tighe S (2012) An exploration of student midwives' experiences of the Objective Structured Clinical Examination assessment process. Nurse Educ Today 32: 690-694.
- Alinier G (2003) Nursing students' and lecturers' perspectives of objective structured clinical examination incorporating simulation. Nurse Educ Today 23: 419-426.
- Smith V, Muldoon K, Biesty L (2012) The Objective Structured Clinical Examination (OSCE) as a strategy for assessing clinical competence in midwifery education in Ireland: a critical review. Nurse Educ Pract 12: 242-247.
- Treadwell I (2006) The usability of personal digital assistants (PDAs) for assessment of practical performance. Educ 40: 855-861.
- Segall N, Doolen TL, Porter JD (2005) A usability comparison of PDA-based quizzes and paper-and-pencil quizzes. Computers and Education 45: 417-432.
- Schmitz FM, Zimmermann PG, Gaunt K (2011) Electronic rating of Objective Structured Clinical Examinations: Mobile Digital forms beat paper and pencil checklist in a comparative study. Information Quality in e-Health. 7058: 501-502.
- Cunningham DM, Walsh CD (2008) OSCE Management Information System.

- 1.8.5, OSCE Management Information System.
13. Kropmans TJ, O'Donovan BG, Cunningham D, Murphy AW, Flaherty G, et al. (2012) An online management information system for objective structured clinical examinations. *Computer and Information Science* 5: 38-48.
 14. Mortaz HS, Jalili M, Muijtjens AM, Van Der Vleuten CP (2013) Assessing the reliability of the borderline regression method as a standard setting procedure for objective structured clinical examination. *J Res Med Sci* 18: 887-891.
 15. Black NM, Harden RM (1986) Providing feedback to students on clinical skills by using the Objective Structured Clinical Examination. *Med Educ* 20: 48-52.
 16. Guiton G, Hodgson CS, Delandshere G, Wilkerson L (2004) Communication skills in standardized-patient assessment of final-year medical students: a psychometric study. *Adv Health Sci Educ Theory Pract* 9: 179-187.
 17. Iramaneerat C, Myford CM, Yudkowsky R, Lowenstein T (2009) Evaluating the effectiveness of rating instruments for a communication skills assessment of medical residents. *Adv Health Sci Educ Theory Pract* 14: 575-594.
 18. Kaufman DM, Mann KV, Muijtjens AM, van der Vleuten CP (2000) A comparison of standard-setting procedures for an OSCE in undergraduate medical education. *Acad Med* 75: 267-271.
 19. Newble DI, Swanson DB (1988) Psychometric characteristics of the objective structured clinical examination. *Med Educ* 22: 325-334.
 20. Verhoeven BH, Hamers JG, Scherpbier HC, Hoogenboom AJJA, Van der Vleuten CPM (2000) The effect on reliability of adding a separate written assessment component to an objective structured clinical examination. *Medical Education* 34: 525-529.
 21. Wass V, Jones R, Van der Vleuten C (2001) Standardized or real patients to test clinical competence? The long case revisited. *Med Educ* 35: 321-325.
 22. Wass V, McGibbon D, Van der Vleuten C (2001) Composite undergraduate clinical examinations: how should the components be combined to maximize reliability? *Med Educ* 35: 326-330.
 23. Mitchell SK (1979) Inter-observer agreement, reliability, and generalizability of data collected in observational studies. *Psychological Bulletin* 86: 376-390.
 24. Iramaneerat C, Yudkowsky R (2007) Rater errors in a clinical skills assessment of medical students. *Eval Health Prof* 30: 266-283.
 25. Iramaneerat C, Yudkowsky R, Myford CM, Downing SM (2008) Quality control of an OSCE using generalizability theory and many-faceted Rasch measurement. *Adv Health Sci Educ Theory Pract* 13: 479-493.