ISSN 2348-313X (Print)
International Journal of Life Sciences Research ISSN 2348-3148 (online)
Vol. 2, Issue 4, pp: (193-204), Month: October - December 2014, Available at: www.researchpublish.com

Correlations between the Objective Structured Clinical Examination Score and Written Examinations in Undergraduates

Aisha M. Al-osail[1], Emad M. Al-osail[2], Mohannad A. Al-ghamdi[3]
Abdulaziz M. Al-hawas[4], Abdullah S. Al-bahussain[5], Ahmed A. Aldajani[6]

[1]Medical consultant and 5th year course coordinator, Dammam University, Prince Saud bin Fahd Street, Khobar, Saudi Arabia.
[2,3,4,5,6] Medical student, Dammam University, Prince Saud bin Fahd Street, Khobar, Saudi Arabia.

Abstract: **One of the most important mission and vision for University of Dammam is to send skilled physicians to deal and treat the patients in professional manner, which cannot be without perfect evaluation of the knowledge and clinical skills of the undergraduates. The Objective Structured Clinical Examination (OSCE) was introduced to increase the reliability and validity of the exams.**

A pilot study was conducted over one semester. A total of 299 examinees (4th year = 207 and 5th year = 92) took the OSCE and written exams in three groups. The OSCE consisted of 18–20 clinical stations, which included taking histories, physical examinations, communication skills, and data interpretation. The written exam contained 80 multiple-choice questions.

Cronbach's alphas by group were 0.7, 0.8, and 0.9 (4th year) and 0.62, 0.79, and 0.85 (5th year). Correlations for all stations ranged from 0.6 to 0.8, which indicated good stability and internal consistency. The reliability of the written exam was 0.79 (4th year) and 0.85 (5th year).The validity of the OSCE was assessed by Pearson's correlations, which were 0.63 (4th year), 0.6 (5th year), and 0.66 (overall).

The OSCE is a perfect tool for undergraduate's assessment because of better reliability and validity, which can be measured.

Keywords**: OSCE, assessment, clinical skills, Cronbach's alpha, reliability, validity, Pearson's correlation.**

I. INTRODUCTION

For many decades, the long/short cases exam style was the method of evaluating medical students in many institutions. In that type of exam, students spend 45–60 minutes interviewing and examining real patients, either from a ward or an outpatient clinic, without observation; the students then present the history and exam findings to the examiner. The students do not perform a range of tasks; scoring well on one case does not mean the student will score well in other scenarios. Thus, this type of exam has reliability problems, which are due to having no standardized measures, and validity problems, which could be improved with direct observation. However, this is difficult to accomplish with a large number of students (Newble 1994; Price & Byrne 1994; Swanson, Norman & Lind 1995). A summary of the reasons that the long/short case exam style has gradually been replaced by the Objective Structured Clinical Examination (OSCE) style (Fraser 2001; Meadow 1998) is as follows.

1. Case distribution among students (some cases are difficult and others are easy).
2. Differences in the scoring of examiners.
3. Lack of reliability index for evaluation.

ISSN 2348-313X (Print)
International Journal of Life Sciences Research ISSN 2348-3148 (online)
Vol. 2, Issue 4, pp: (193-204), Month: October - December 2014, Available at: www.researchpublish.com

4. A concentration on primarily one or two systems.
5. Candidates not being observed while taking histories or conducting. physical examinations
6. Lack of blue prints, which is defined by Walvoord and Anderson (2010) as the process of linking test to the learning goals
7. Inconsistency in exam scores in comparison to clinical exam scores of some students.
8. Differences in the levels of cooperation of the patients.
9. Possibility of discussion of investigation and plans being skipped due to time limitations.
10. Probability of psychological and grading effects due to injustice in the scoring of students (Flexner 1910; Norcini 2002; Norcini 2001; van der Vleuten 1996).

These are the reasons that the long case exam style has gradually been replaced by the OSCE (Fraser 2001; Meadow 1998).

The main aim of clinical exam is to evaluate the students' communication history-taking, physical examination and procedure skills, in addition to assessing knowledge, investigations, and planning discussions, which the long/short cases style exam fails to do, especially with an increasing number of medical students.

Harden and Gleeson (1979) implemented the first OSCE as a more reliable and valid exam; changes to the design and structure have increased reliability and validity. The OSCE has many advantages, including the following.

1. Assesses many skills in less time.
2. Standardizes patients/cases for all candidates.
3. Maintains consistency in scoring by examiners .
4. Provides a blue print, which allows more systems to be covered.
5. Offers a reliability index to evaluate the exam.
6. It is a valid test.
7. It is considered a fair exam (Ben-David 2003; Harden & Gleeson 1979; Hodges 2003).

Miller (1990) established a pyramid for assessing the clinical competency of students (Figure 1). The pyramid describes the many domains of clinical education. At the bottom is "knows," which refers to the assessment of knowledge in a written exam, such as the Multiple Choice Question (MCQ) . The next level, "knows how," represents the ability to apply the knowledge gained, which can also be assessed by the MCQ. To this point, only cognitive skills have been measured. "Shows how" refers to the ability to demonstrate the application of knowledge, while "Does" is the application of knowledge to practice, which can be measured through direct observation. These final two sections of the pyramid are the behavioral component of the exam. Using the OSCE style and a written exam, the entire pyramid can be covered. According to Miller (1990), this combination represents a valid exam.

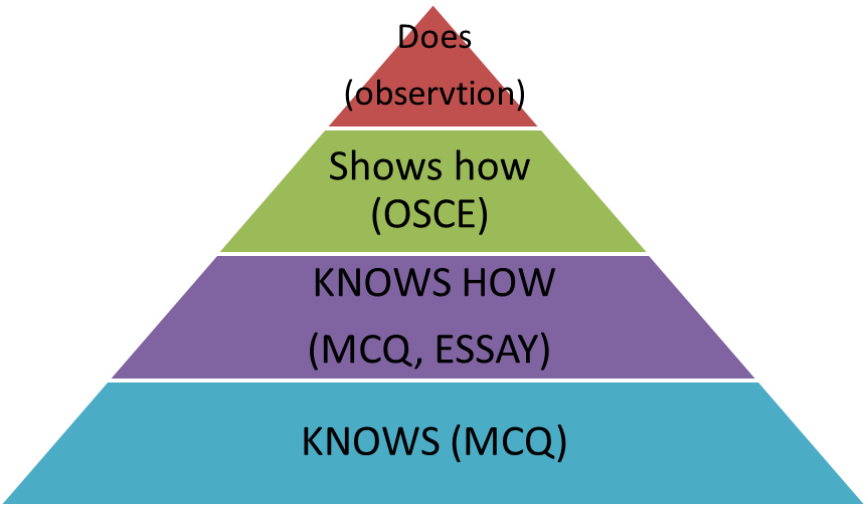

Figure 1: Adapted version of Miller's pyramid

Many studies have reported experiences of introducing the OSCE at universities worldwide; most of the detailed articles cover the role of pre-clinical courses, such as anatomy (Yaqinuddin et al. 2012) and clinical courses, such as the reliability and validity of a pediatric course (Matsell, Wolfish & Hsu 1991), and radiology course (Morag et al. 2001) and psychiatry course (Park et al. 2004; Walters, Osborn & Raven 2005). Grover et al. (2012) reported the experiences in a dermatology department, and the assessment of the clinical competency proved that OSCE was a valid tool in dermatology course assessment, surgery, family medicine, anesthesia, and clerkship based on the reliability and validity of the OSCE. These studies have covered internal medicine in detail for residency programs or specific systems, such as nephrology or cardiology, but no study reported details of internal medicine courses and exams for medical students. The reports covered many issues, including reliability, validity, feedback from students, and gender difference and its use as a teaching tool (Grover et al. 2012; Matsell, Wolfish & Hsu 1991; Morag et al. 2001; Park et al. 2004; Shams et al. 2013; Walters, Osborn & Raven 2005; Yaqinuddin et al. 2012).

The reliability of OSCE exam has been measured by many tools, such as Cronbach's alpha; this tool can evaluate "stability," which can demonstrate the differences in students' performance at each station and provide a global rating, which evaluates the overall performance and if the checklist used is appropriate for the skill level of the students. The R2 coefficient measures the proportional change in the dependent variable (the checklist score) by changes in the independent variable (the global grade); this is a marker of internal consistency (Auewarakul et al. 2005; Eberhard et al. 2011; Iramaneerat et al. 2008; Lawson 2006; Schoonheim-Klein 2008; Tavakol & Dennick 2011). The final method of establishing the validity of the exam is the comparison of the OSCE scores with written exam scores using Pearson's correlation (Pell et al. 2010).

In Saudi Arabia, King Saud University and King Abdul-Aziz University were the first Saudi universities to report their experiences using the OSCE in the field of Surgery, which was reported in detail by Alnaami (2008) at King Saud University (for students in their final year). This report concentrated on reliability and validity. The use of the OSCE in Family Medicine at King Saud University was reported by Haheel and Naeem (2013); the undergraduate perceptions of the exams were positive. For the field of Dermatology, it was concluded that the OSCE was a gold standard exam, but again, there were no details about internal medicine (Al-Naami 2008; Raheel & Naeem 2013).

In 2014, the Internal Medicine department at Dammam University decided to shift from the old style long/short cases exams to OSCE style after becoming aware of the more obvious disadvantages of the former exam style. These included the difficulty in establishing the reliability and validity of the exam, especially after the increase in the number of medical students accepted and the many students entering a residency or fellowship program in US, Canada, UK, or Australia, where the OSCE is part of the licensing requirements [e.g., the Medical Licensing Examination (USMLE), Medical Council of Canada Evaluating Examination (MCCEE), and Professional and Linguistic Assessments Board (PLAB)]. In addition, the OSCE will assist in the goals of sending competent physicians into the community and evaluating the quality and contents of courses.

Study Aims

The aims of this study were:

- Evaluate the reliability and validity of the OSCE.
- Assess if different reliability results affect the validity of the exam.
- Develop a standard for all examinees.
- Identify weakness in courses.
- Ensure the competency of graduates.

II. METHOD

Participants

This pilot study was conducted one semester (February-May) with 299 medical students (4^{th} year = 207 and 5^{th} year = 92), who took the exam in three groups (March, April, and May 2014). At the end of the semester, the students took the written exam, which consisted of 80 multiple-choice questions.

ISSN 2348-313X (Print)
International Journal of Life Sciences Research ISSN 2348-3148 (online)
Vol. 2, Issue 4, pp: (193-204), Month: October - December 2014, Available at: www.researchpublish.com

Procedure

An introductory lectures about the OSCE was given for faculties and each student group on the first day of the course.

The blue print inserted for each exam. The OSCE exam had 18–20 clinical stations and covered all clinical parts (history, physical examinations, communication skills, and data interpretation). Each station took 7 minutes to complete. Students were distributed in groups as shown in Table 1 and 2.

Table 1: Reliability measures for the 4th year OSCE

Day/data	Gender	Students/ days	Stability[a]	Internal consistency[b]	*p*-value	Internal consistency[c]
First group	Male	56/2 days	0.7	0.7	< 0.001	0.72(72%)
Second group	Female	97/3 days	0.8	0.7	< 0.001	0.82(82%)
Third group	Male	54/2 days	0.9	0.8	< 0.001	0.782(78.2%)

Note. [a]Cronbach's alpha. [b]Spearman Rank Correlation. [c]R2 coefficient determinants.

Table 2: Reliability measures for the 5th year OSCE

Day/year	Gender	Students/days	Stability[a]	Internal consistency[b]	*p*-value	Internal consistency[c]
First group	Female	29/ 1 day	0.621	0.6	< 0.001	0.61(61%)
Second group	Female	21/ 1 day	0.799	0.621	< 0.001	0.8 (80%)
Third group	Male	42/ 1 day	0.854	0.75	< 0.001	0.85(85%)

Note. [a]Cronbach's alpha. [b]Spearman Rank Correlation. [c]R2 coefficient determinants

The highest final score is 100%; the OSCE exam accounts for 40% in the 4th year and 50% in the 5th year, a continuous assessment accounts for 10%, and the written exam accounts for 50% in the 4th year and 40% in the 5th year. All 299 students took the clinical and written exam; after each exam, the coordinator of the course met with faculty and students to assess and correct any problems with the OSCE to ensure better reliability in the future.

Ethical Considerations

The Institutional Review Board of the University of Dammam approved the study (Approval number: IRB-2014-01-317). Informed consent was obtained from all participants.

Data analysis

The exam reliability was measured by Cronbach's alpha, the global rating, and the coefficient of determination, R2.and Spearman's Rank Correlation. At the end of the semester, a written exam held, which was analyzed (mean, median, mode) separately for each year. The validity was assessed by Pearson's correlation. Factor Analysis for each station was conducted to understand any deficits in the exam.

III. RESULTS

The reliability for the OSCE was assessed by Cronbach's alpha, which indicated the stability of the stations on the three exams for the 4th year students. The alphas were 0.7, 0.8, and 0.9. For the 5th year students, the alphas were 0.621, 0.799, and 0.854. Spearman's rank correlation and R2 coefficient determinants were used for internal consistency (correlating the checklist and global score) . The correlations were 0.7, 0.7, 0.8 for the 4th year and 0.6, 0.621 and 0.75 for the 5th year ($p < .001$), which indicated a strong correlation. The R2 coefficient determinants, which were used to measured the linear correlation between the check list and the global score, were 72%, 82%, and78.2% for the 4th year students, and 61%, 80%, and 85% for the 5th year students. (Table 1 and 2).

ISSN 2348-313X (Print)
International Journal of Life Sciences Research ISSN 2348-3148 (online)
Vol. 2, Issue 4, pp: (193-204), Month: October - December 2014, Available at: www.researchpublish.com

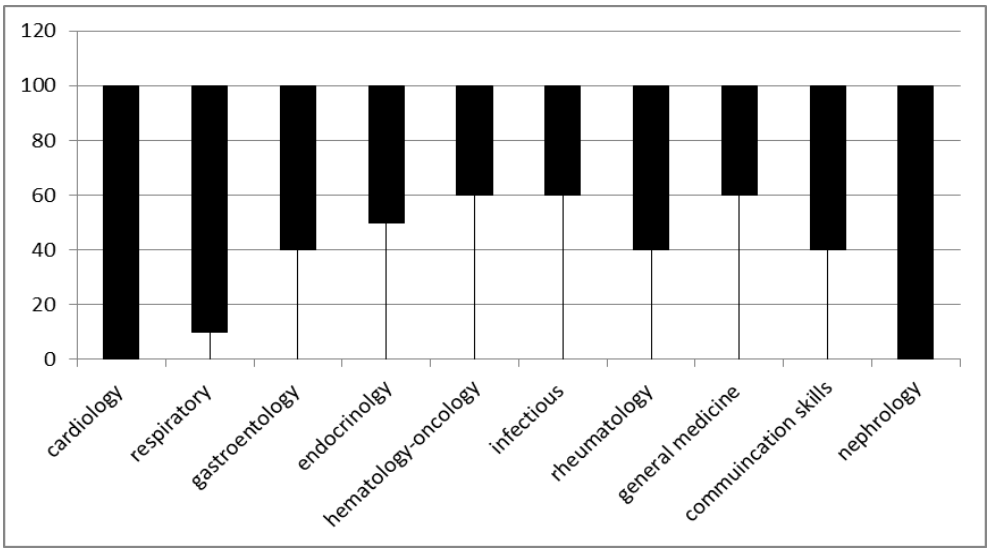

Figure 2: OSCE scores by systems

The Cronbach's alpha for the stations was between 0.5–0.9. Figure 2 shows Cronbach's alpha for the stations based on the systems.

The OSCE score for 4th year students was between (18.7–36.9) with a mean = 27.6, median = 27.9, standard deviation (SD) = 4.07, and skewness = -0.07, which is almost 0, and revealed a normal distribution. Kurtosis (2.37) indicated the curve was flatter than a normal distribution with a wider peak. The probability for extreme values was less than for a normal distribution, and the values had a wider spread around the mean. For the 5th year students, it was between (20–45.9/50) with mean = 37.8, median = 38.7, skewness = -1, and standard deviation = 4.56, which indicated most of the scores were around the right side of the mean and the extreme values were on the left. OSCE score analysis for the students is shown in detail on Table 3 and 4.

Table 3: The analysis for the 4th year OSCE score (total possible = 40)

Statistical parameters	Result
Minimum:	18.7
Maximum:	36.9
Range:	18.2
Count:	207
Mean:	27.6
Median:	27.9
Mode:	26.8, 25.5, 30.8, 28.3
Standard Deviation:	4.07
Variance:	16.5
Mid Range:	27.8
Quartiles:	Quartiles: Q_1 --> 24.9 Q_2 --> 27.9 Q_3 --> 30.8
Interquartile Range (IQR):	5.9
Mean Absolute Deviation:	3.30
Root Mean Square (RMS):	27.9
Std Error of Mean:	0.28
Skewness:	-0.07
Kurtosis:	2.37
Coefficient of Variation:	0.14
Relative Standard Deviation:	14.7%

ISSN 2348-313X (Print)
International Journal of Life Sciences Research ISSN 2348-3148 (online)
Vol. 2, Issue 4, pp: (193-204), Month: October - December 2014, Available at: www.researchpublish.com

Table 4: The analysis for the 5th year OSCE score (total possible = 50

Statistical parameters	Results
Minimum:	20
Maximum:	45.9
Range:	25.9
Count:	92
Mean:	37.8
Median:	38.7
Mode:	39.35
Standard Deviation:	4.56
Variance:	20.8
Mid Range:	32.9
Quartiles:	Quartiles: Q_1 --> 35.5 Q_2 --> 38.7 Q_3 --> 40.7
Interquartile Range (IQR):	5.2
Mean Absolute Deviation:	3.46
Root Mean Square (RMS):	38.15
Std. Error of Mean:	0.47
Skewness:	-1.00
Kurtosis:	4.67
Coefficient of Variation:	0.12
Relative Standard Deviation:	12.0%

The reliability of the written exam was = 0.79 for the 4th year and 0.854 in the 5th year, which was considered very good and was superior for the 5th year students.

The score analysis for the written exam is shown on Tables 5 and 6 in detail; the minimum score was 18.1 and the maximum was 43.1 (out of 50%) for the 4th year students, with mean of 33.6, median of 33.75, SD = 4.35, and relative SD = 12.9, while the minimum score was 15.5 and the maximum was 36 (out of 40%) for the 5th year students, with a mean of 29.5, median of 29.75, SD of 3.76, and relative SD = 12.7%. The SD and relative SD are small in both the 4th and 5th year results, while the skewness and the kurtosis are almost same in the 4th and 5th year written exams.

Table 5: The analysis of the 4th year written exam

Statistical parameters	Results
Minimum:	18.125
Maximum:	43.125
Range:	25
Count:	207
Mean:	33.6
Median:	33.75
Mode:	34.37
Standard Deviation:	4.35
Variance:	18.9
Mid Range:	30.625
Quartiles:	Quartiles: Q_1 --> 30.62 Q_2 --> 33.75 Q_3 --> 36.25
Interquartile Range (IQR):	5.625
Mean Absolute Deviation:	3.48
Root Mean Square (RMS):	33.9

ISSN 2348-313X (Print)
International Journal of Life Sciences Research ISSN 2348-3148 (online)
Vol. 2, Issue 4, pp: (193-204), Month: October - December 2014, Available at: www.researchpublish.com

Std Error of Mean:	0.302
Skewness:	-0.34
Kurtosis:	3.36
Coefficient of Variation:	0.12
Relative Standard Deviation:	12.93%

Table 6: The analysis of the 5th year written exam

Statistical parameters	Results
Minimum:	15.5
Maximum:	36
Range:	20.5
Count:	92
Mean:	29.5
Median:	29.75
Mode:	29
Standard Deviation:	3.76
Variance:	14.1
Mid Range:	25.75
Quartiles:	Quartiles: Q_1 --> 27 Q_2 --> 29.7 Q_3 --> 32.5
Interquartile Range (IQR):	5.5
Mean Absolute Deviation:	2.98
Root Mean Square (RMS):	29.7
Std Error of Mean:	0.39
Skewness:	-0.64
Kurtosis:	3.73
Coefficient of Variation:	0.12
Relative Standard Deviation:	12.7%

For the validity of the exam, we compared the results of the OSCE score and the written exam score using Pearson's correlation. The correlation was 0.63 for the 4th year and 0.6 for the 5th year; the overall correlation was 0.66, which indicated a strong correlation between the OSCE scores and the written exam score (Figure 3, 4, and 5).

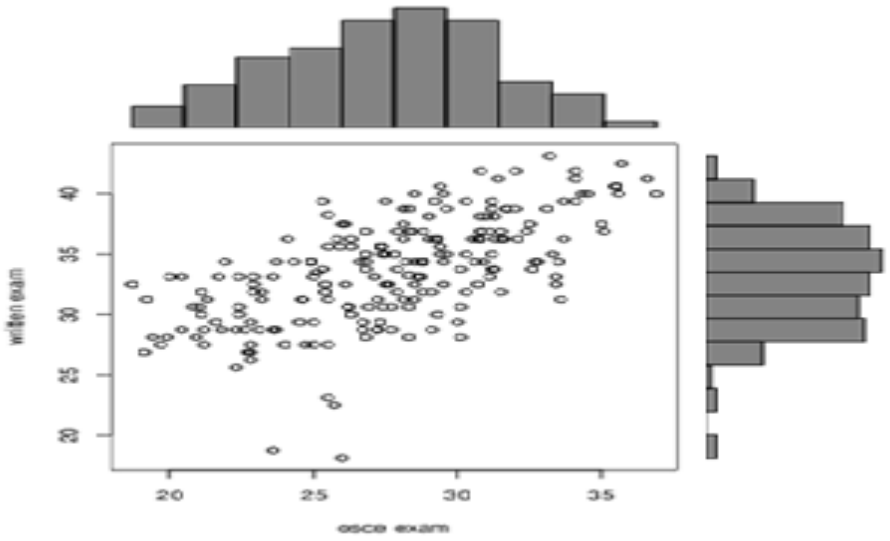

Figure 3: Pearson's correlation for the 4th year

ISSN 2348-313X (Print)
International Journal of Life Sciences Research ISSN 2348-3148 (online)
Vol. 2, Issue 4, pp: (193-204), Month: October - December 2014, Available at: www.researchpublish.com

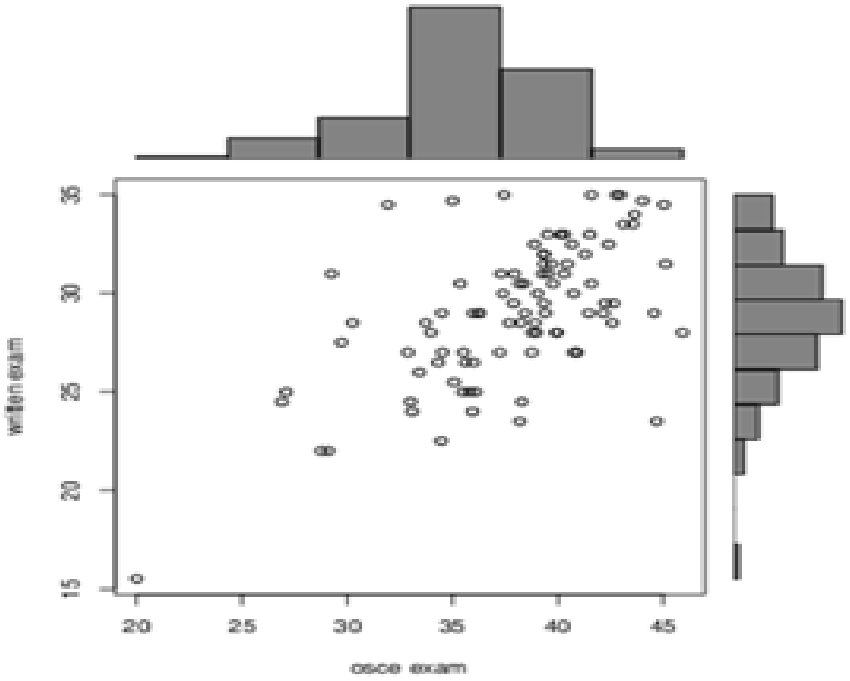

Figure 4: Pearson's correlations for the 5th year

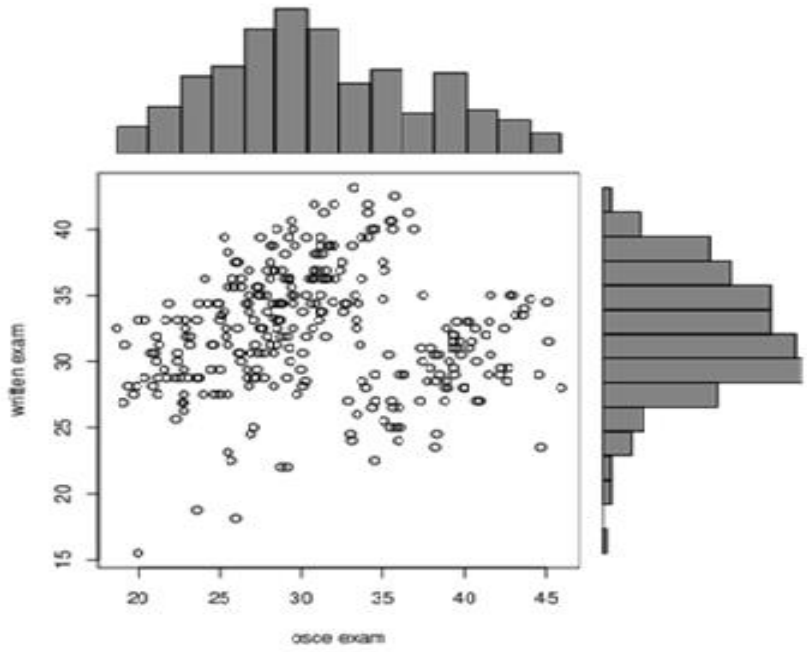

Figure 5: Pearson's correlations for all medical students

Finally, a factor analysis was conducted (with rotated factor), which ensured the components of OSCE stations and construction of the exam: for the 4th year, the value of the rotated factor was ranged from 0.1–0.99, while for the 5th year, it ranged from 0.219–0.9. (Table 7 and 8). This added to the validity of the exam.

ISSN 2348-313X (Print)
International Journal of Life Sciences Research ISSN 2348-3148 (online)
Vol. 2, Issue 4, pp: (193-204), Month: October - December 2014, Available at: www.researchpublish.com

Table 7: Factor analysis for the 4th year results

Stations	Factor1
V1	0.64
V2	0.547
V3	0.713
V4	0.499
V5	0.694
V6	0.621
V7	0.154
V8	0.39
V9	0.613
V10	0.604
V11	0.675
V12	0.795
V13	0.804
V14	0.684
V15	0.752
V16	0.682
V17	0.991
V18	0.991

Table 8: Factor analysis for the 5th year results

Stations	Factor1
V1	0.691
V2	0.59
V3	0.691
V4	0.583
V5	0.7
V6	0.7
V7	0.655
V8	0.521
V9	0.678
V10	0.623
V11	0.723
V12	0.8
V13	0.721
V14	0.655
V15	0.68
V16	0.739
V17	0.3
V18	0.99
V19	0.219
V20	0.982

IV. DISCUSSION

This is a pilot study conducted in the Internal Medicine department at University of Dammam. First, by evaluating the examinee sample who involved in the exam was an excellent sample to start the exam , and the reliability of the OSCE stations was very good and increased after each exam . Second point, many reliability tool measures used (Cronbach's alpha, for the stability of the stations measurement , and Spearman's rank correlation and the R2 coefficient determinants for internal consistency. There was a rise in the internal consistency after each exam secondary to the developments after each exam, with orientation lectures of the faculty, avoiding the mistakes in the station checklists and distribution, and without forgetting better orientation of students who learn from the previous participants. Third issue, we can see that SD were small in all exams (OSCE and written) which indicates that the exam was reliable. The validity of the exam was strong, and we anticipate achieving very strong which encouraging the course coordinators to achieve better result in future.

The outcomes of this pilot study are promising the other departments to conduct the OSCE. By reviewing the other publications that discussed the implementation of the OSCE, this pilot study is the first report in internal medicine course; many reports were focused mainly on the reliability and validity of the exam, teaching item (Brazeau, Bopyd & Crosson 2002). Our pilot study confirmed improvement in carrying the OSCE; this seen by the enhancements in the indexes that were used for reliability measurement for exam. A last point we noticed different days will not change the validity of the exam (Schoonheim-Klein et al. 2008).

Limitations

First, this is a single department and institute study, which involved only 4th and 5th year medical students who agreed to the new examination format. The second limitation is that the topic of management was not there even though it is involved in the course.

V. CONCLUSION

The OSCE is a perfect tool for undergraduate's assessment because of better reliability and validity, which can be measured by many tools. OSCE exam proved to be a reliable, valid and can be conducted in a multi- system courses. More reports are needed in the futures in internal medicine course and other clinical courses.

Practice Points

1. OSCE is an important tool for curriculum development and improvement through identification of weak and strong parts of a course, especially in multi-systems.
2. OSCE is an excellent style of examination, especially for multi-system courses like internal medicine, for evaluating multiple skills.
3. Factor analysis can be used to evaluate examination structure and should be conducted for all examinations.

ACKNOWLEDGEMENTS

We would like to acknowledge the university, internal medicine department, including our chair, Dr.Waleed Albaker, who supports the idea of replacing the long/short cases exam with the OSCE, faculty members, specialists, residents, Dr. Abeer Alsaid, Mr. Zee Shan, and the medical students who were interested to participate in OSCE exam.

Declaration of interest

The authors report no declaration of interests.

Notes on contributors

Aisha M. Alosail, MBBS, MD, is a medical consultant and 5th year internal medicine course coordinator for the medical college at University of Dammam. Emad M. Alosail, Mohannad Ali AlGamdi, Abdulaziz Mohammed Alhawas, Abdullah Sami Albahussain, and Ahmed Abdulfatah Aldaijani are 5th year medical students at University of Dammam.

REFERENCES

[1] Al-Naami MY. 2008. Reliability, validity, and feasibility of the Objective Structured Clinical Examination in assessing clinical skills of final year surgical clerkship. Saudi Med J 29(12):1802–1807.

[2] Auewarakul C, Downing S, Praditsuwan R, Jaturatamrong U. 2005. Item analysis to improve reliability for an internal medicine undergraduate OSCE. Adv Health Sci EducTheory Pract 10:105–113.

[3] Ben-David MF. 2003. Life beyond the OSCE. Med Teacher 25:239–240.

[4] Brazeau C, Boyd L, Crosson J. 2002. Changing an existing OSCE to a teaching tool: the making of a teaching OSCE Acad Med 77(9):932.

[5] Eberhard L, Hassel A, Bäumer A, Becker J, Beck-Muotter J, Bömicke W, Corcodel N, Cosgarea R, Eiffler C, Giannakopoulos NN, Kraus T, Mahabadi J, Rues S, Schmitter M, Wolff D, Wege KC. 2011. Analysis of quality and feasibility of an objective structured clinical examination (OSCE) in preclinical dental education. Eur J Dent Educ 15:1–7.

[6] Flexner A. 1910. Medical Education in the United States and Canada. Science and Health Publications: Bethesda, Maryland

[7] Fraser R. 2001. Does observation add to the validity of the long case? [Author response] Med Educ 35:1131–1133.

[8] Pell G, Fuller R, Homer M, Roberts T; International Association for Medical Education. How to measure the quality of the OSCE: A review of metrics - AMEE guide no. 49. Med Teach. 2010;32(10):802-11. doi: 10.3109/0142159X.2010.507716.

[9] Grover C, Bhattacharya SN, Pandhi D, Singal A, Kumar P. 2012. Computer Assisted Objective Structured Clinical Examination: a useful tool for dermatology undergraduate assessment. Indian J Dermatol Vererol Leprol 78(4):519. doi: 10.4103/0378-6323.98095

[10] Harden RM, Gleeson FA. 1979. ASME Medical Educational Booklet 8. Assessment of medical competence using an objective structured clinical examination (OSCE). J Med Educ 13:41–54.

[11] Hodges B. 2003. Validity and the OSCE. Med Teacher 25:250–254.

[12] Iramaneerat C, Yudkowsky R, CM. M, Downing S. 2008. Quality control of an OSCE using generalizability theory and many-faceted Rasch measurement. Adv Health Sci Educ Theory Pract 13:479–493.

[13] Lawson D. 2006. Applying generalizability theory to highstakes objective structured clinical examinations in a naturalistic environment. J Manipulative Physiol Ther 29:463–467.

[14] Matsell DG, Wolfish NM, Hsu E. 191. Reliability and validity of the objective structured clinical examination in paediatrics. Med Educ 25(4):293–299.

[15] Meadow R. 1998. The structured exam has taken over. BMJ 317:1329.

[16] Miller GE. 1990. The assessment of clinical skills/competence/performance. Acad Med 65:s63–s67.

[17] Morag E, Lieberman G, Volkan K, Shaffer K, Novelline R, Lang EV. 2001. Clinical competence assessment in radiology: introduction of an objective structured clinical examination in the medical school curriculum. Acad Radoil 8(1):74–81.

[18] Newble DI. 1994. The observed long case in clinical assessment. Med Educ 25:369–73.

[19] Norcini JJ. Does observation add to the validity of the long case? [Letter.] Med Educ 2001; 35:1131–3.

[20] Norcini JJ. 2002. The death of the long case? BMJ 324:408–409.

[21] Park RS, Chibnall JT, Blaskiewicz RJ, Furman GE, Powell JK, Mohr CJ. 2004. Construct validity of an objective structured clinical examination (OSCE) in psychiatry: associations with the clinical skills examination and other indicators. Acad Psychiatry 28(2):122–128.

[22] Pell G, Fuller R, Homer M, Roberts T, International Association for Medical Education. 2010. How to measure the quality of the OSCE: A review of metrics - AMEE guides no. 49. Med Teach 32(10):802-11. doi: 10.3109/0142159X.2010.507716.

[23] Price J, Byrne GJA. 1994. The direct clinical examination: an alternative method for the assessment of clinical psychiatric skills in undergraduate medical students. Med Educ 28:120–125.

[24] Raheel H, Naeem N. 2013. Assessing the Objective Structured Clinical Examination: Saudi family medicine undergraduate medical students' perceptions of the tool. J Pak Med Assoc 63(10):1281–1284.

[25] Schoonheim-Klein M, Muijtjens A, Habets L, Manogue M, Van der Vleuten C, Hoogstraten J, Van der Velden U. 2008. On the reliabilityof a dental OSCE, using SEM: effect of different days. Eur J Dent Educ 12:131–137. doi: 10.1111/j.1600-0579.2008.00507.x.

[26] Shams T, El-Masry R, Al Wadani H, Amr M. 2013. Assessment of current undergraduate anesthesia course in a Saudi University. Saudi J Anaesth 7(2):122–127. doi: 10.4103/1658-354X.114049

[27] Swanson DB, Norman GR, Linn RL. 1995. Performance based assessment: lessons learnt from the health professions. Educ Res 24 (5):5–11.

[28] Tavakol M, Dennick R. 2011. Making sense of Cronbach's alpha. International Journal of Medical Education 2:53–55.

[29] van der Vleuten CPM. 1996. Making the best of the "long case". Lancet 347:704–705.

[30] Walters K, Osborn D, Raven P. 2005. The development, validity and reliability of a multimodality objective structured clinical examination in psychiatry. Med Educ 39(3):292–298.

[31] Walvoord BE, Anderson VJ. 2010. Effective Grading: A Tool for Learning and Assessment in College. 2nd ed. San Francisco: Jossey-Bass.

[32] Yaqinuddin A, Zafar M, Ikram MF, Ganguly P. 2013. What is an objective structured practical examination in anatomy? Anat Sci Educ 6(2):125–33. doi: 10.1002/ase.1305. Epub 2012 Aug 28.

Corresponding Author:

Aisha M. Alosail

Medical consultant and 5th year course coordinator

University of Dammam , Prince Saud bin Fahd Street

Khobar, P.O. Box 3669, Saudi Arabia, 31952

aosail@ud.edu.sa

aosail@yahoo.com