Evaluation of Multifaceted Assessment of the Fifth-Year Medical Students in Family Medicine Clerkship, Saudi Arabia Experience

Background: There are few data on assessment methods in the field of family medicine in Arab countries. Imam Abdulrahman Bin Faisal University (IAU) in 2017 established the family medicine clerkship rotation for the fifth-year medical students. The objective of the study is to evaluate the multifaceted assessment of family medicine clerkship and their correlation to each other among fifth-year medical students in IAU.

Methods: We examined the performance of 455 fifth-year medical students from two batches (2018 and 2019) in the continuous course assessment, Objective Structured Clinical Examinations (OSCE), extended matching questions (EMQ), and multiple-choice questions exams (MCQ).

Results: The students’ total mark range was 55.16 to 95.64. Batch 2018 and females had significantly higher marks (p=0.000). On the other hand, no significant associations were seen between the total marks and the semester in which the clerkship was taken or the order of group in each semester (p= 0.464 and 0.980 respectively). Significant positive correlations were seen between all forms of assessment during the family medicine clerkship (p=0.000).

Conclusion: This study demonstrates that different assessment methods used to evaluate medical students during the family medicine clerkship were correlated. The performances of the female and the first batch were superior. This finding is the basis for introducing a platform for developing modern learning and the improvement of the assessment standards in family medicine clerkship rotation.

Keywords: family medicine, clerkship, medical students, assessment

Introduction

Family medicine is a medical specialty that deals with the individual and family, both sexes, all ages, and all body systems in a comprehensive and continuity model of care.\(^1\) This specialty integrates clinical, behavioral, biological, and social sciences.\(^1\) There is a global need for family physicians and teaching family medicine concepts at the undergraduate level in medical schools.\(^2,3\) In Saudi Arabia, with the emergence of the 2030 vision, efforts are made to encourage medical students towards the specialty of family medicine.\(^4\) A study conducted in 2015, indicated that family medicine course is given as a 4 to 8 weeks rotation in 13 out of 15 medical colleges around Saudi Arabia.\(^5\) Medical curriculum at Imam Abdulrahman Bin Faisal University, in Saudi Arabia, has been subjected to continuous and rigorous evaluation since its inception in 1975. The college of medicine launched a new revised study plan in 2014. The curriculum has been developed in...
alignment with the vision and mission and goals of the University and the College of Medicine and responded to the Saudi National requirements and international standards. The revised MBBS curriculum is a six-year undergraduate degree consists of one preparatory year required for all health specialties students, followed by five academic years in addition to one mandatory Internship training year. The curriculum is adopting a problem-based approach, an early introduction to the patient and the community, and the basic and clinical biomedical sciences with the social and behavioral aspects of medicine. It is designed as a thematic integrated program incorporating four themes across the five years of the study plan. These four themes are: Personal and Professional Development; Population, Society, Health and Illness; The Scientific Basis of Clinical Practice; and Clinical Skills. During the early years of the program, the basic medical, community based, and behavioral sciences are introduced within inter-disciplinary units. The fifth year is largely taken up with core clinical rotations in women’s and children’s health, family medicine and mental health.6,7 The family medicine course involves fifty topics in family medicine. It covers basic concepts in family medicine, including principles of family medicine, communication skills, anticipatory care, and common health problems in family medicine. Different methods of assessment are used in undergraduate levels such as the Objective Structured Clinical Examination (OSCE), the extended matching questions (EMQ) test, and the multiple-choice questions (MCQ) exam.8 OSCE was first introduced in 1975.9 It is increasingly used for the evaluation of students because of its reliability and validity.10 It is the best method to assess many skills, such as: taking the history, performing the physical examination, interpretation of investigations, communication skills, and management.11 A study conducted in 2010 at the college of medicine in King Saud University, Riyadh, Saudi Arabia, revealed that 52% of undergraduate medical students perceived that OSCE is a fair method of assessment, and 65% believed that it is a standardized tool for evaluation.12 EMQ test is a form of written exam. Researchers found that it a reliable test in assessing students’ core knowledge and clinical reasoning with less chance of students guessing the correct answers in compared with MCQ.13 On the other hand, MCQ examination has strong discriminative capacity than EMQ in identifying students with poor levels.13 Previous studies revealed that female students perform in the exams better than male students at the undergraduate level.14 Others found that there are no significant differences in exam scores between males and females.15

The objective of this study was to assess and compare the outcome of different methods of family medicine clerkship assessment among the fifth-year medical students and to assess the influence of gender and timing of the course on students’ performance at Imam Abdulrahman Bin Faisal University, Saudi Arabia.

Methodology
This research was a cross-sectional study conducted by the Department of Family and community medicine at the Imam Abdulrahman Bin Faisal University between 2018–2019 after obtaining approval from the college authorities. All fifth-year students from two batches (2018 and 2019) were included (N=455). All participants gave their informed consent. Confidentiality of the data was assured, and it will be used only for research purposes. For each batch, there are two semesters. For each semester, we have two groups, males’ group and females’ group, each group spends eight weeks rotation. During the rotation, students have lectures in the morning and clinics in the afternoon. According to the college regulations, the Grade Point Average (GPA) is as the following: grade A if the total mark ≥90, grade B if the mark ≥80 to 89, grade C if the mark ≥70 to 79, grade D if the mark ≥60 to 69, and grade F if the mark <60. Students’ records were the source of data collected.

Assessment Formats
The assessment of the course consists of; continuous course assessment, extended matching questions (EMQ) and multiple-choice questions exams (MCQ) at the end of the rotation, and Objective Structured Clinical Examinations (OSCE) at the end of the semester. The final written exam included MCQ and EMQ exams. For batch 2018, the final written exam was at the end of the year and was combined with all rotations taken in the fifth year (psychiatry, obstetrics and gynecology, and pediatrics). For batch 2019, the final written exam was at the end of each semester and was combined with the psychiatry exam only. A total mark of all assessment formats in family medicine clerkship was 100%.

Continuous Course Assessment
Total marks of 34% divided into 5.5% logbook, 5.5% presentation of the data analysis from the logbook, 5.5% reflective journal, and 17.5% attendance. If attendance is
less than 85%, the student will not be allowed to attend any exam and must repeat the course in the summer according to university regulations.

Multiple-Choice Questions
Students have two MCQs exams, one at the end of the rotation and the other as part of the final written exam with a total mark of 11% for each one. The MCQ paper at each examination consisted of 50 questions. The questions’ bank consists of 500 questions that were designed to cover all lectures. One question was included to represent each lecture. Each MCQ question consisted of a stem of a short sentence, along with four response options, with one correct answer. The type of MCQs is the single-best answer. The correcting machine corrected the MCQ answer sheets.

Extended Matching Questions
Students have two EMQ exams, one at the end of the rotation and the other as part of a final written exam with a total mark of 11% for each one. EMQs are problem-focused questions, often referring to realistic cases. They have four components: (i) a theme, (ii) a lead-in statement for the questions giving the students instructions on what to do, (iii) the questions giving students pertinent information based on which the student is to select the correct answer and (iv) a list of options or answer possibilities. In the EMQs, students were asked to choose the best answer from a list of five options that were designed by the examiner, each of which could be used once, more than once, or not at all. The exams have ten themes. The correcting machine corrected the EMQ answer sheets.

Objective Structured Clinical Examinations
It is at the end of each semester with a total mark of 22%. The OSCE was based on the curricular constructs that included five primary skills: history taking, physical examination, patient education, approach to the acute case, and anticipatory care. The OSCE was composed of five stations. Each station has one examiner and one simulator. Assessment is based on a structured checklist developed by one faculty and revised by the exam committee. Students spend 8 mins to perform tasks at each station.

Regarding exam validity, family medicine clerkship assessment system is based on exam blueprint distributing the assessment tools (MCQs, EMQs, and OSCE) to cover the whole clerkship contents in order to ensure the content validity. For each assessment tools, there is also a specific content blueprint to cover suitable course contents.

Statistical Analysis
Data analysis was carried out using the Statistical Software for the Social Sciences (SPSS) package (Version 15, Chicago, Illinois, USA). The reliability of the OSCE examination was assessed by Cronbach’s alpha coefficient. The reliability of the MCQs and EMQs examination was assessed by Kuder–Richardson 20 (KR-20). The Borderline Regression method was used to calculate pass marks in OSCE. Median, minimum, maximum, mean, and standard deviations were calculated for examination marks, and frequency and percentage used for the categorical data. Statistical comparison was carried out using the t-test, chi-squared test, and Pearson’s correlation coefficient. P-value < 0.05 was considered statistically significant.

Results
A total of 455 students participated in the study. Females were slightly more than males in both batches. About half of students take the family medicine clerkship in the first semester, and half of them take it as the first group in each semester. (Table 1)

Table 2 displays the students’ scores in different assessment methods used during the clerkship. The total marks score range was 55.16 to 95.64, and the mean was 84.93.

Table 1 Sociodemographic Characteristics of Medical Students, Imam Abdulrahman Bin Faisal University, Saudi Arabia, 2018–2019 (N=455)

| Characteristics                     | Batch                  | 2018 (n=209) | 2019 (n=246) |
|-------------------------------------|------------------------|-------------|-------------|
|                                     | No. (%)                | No. (%)     |             |
| Gender                              |                        |             |             |
| Male                                | 87 (41.6%)             | 121 (49.2%) |             |
| Female                              | 122 (58.4%)            | 125 (50.8%) |             |
| Semester of which the course was    |                        |             |             |
| taken                               |                        |             |             |
| First                               | 104 (49.8%)            | 123 (50.0%) |             |
| Second                              | 105 (50.2%)            | 123 (50.0%) |             |
| Group order in each semester        |                        |             |             |
| First                               | 104 (49.8%)            | 124 (50.4%) |             |
| Second                              | 105 (50.2%)            | 122 (49.6%) |             |
The reliability of the OSCE examination was assessed by Cronbach’s alpha coefficient with a mean of 0.71 and standard deviation of 0.03 for batch 2018, and with a mean of 0.76 and standard deviation of 0.1 for batch 2019. The reliability of the MCQs examination was assessed by Kuder–Richardson 20 (KR-20) showed a mean of 0.87 and standard deviation of 0.05 for batch 2018, and a mean of 0.86 and standard deviation of 0.06 for batch 2019. EMQs reliability was calculated using Kuder–Richardson 20 (KR-20) showed a mean of 0.84 and standard deviation of 0.11 for batch 2018, and a mean of 0.87 and standard deviation of 0.02 for batch 2019. Standard settings were not performed for written exams because family medicine is part of year five study plan and it is not a high-stake examination. However, it was tested in OSCE exams where the Borderline Regression method was used to calculate pass marks. The mean pass marks for five OSCE stations for the four exams in years 2018 and 2019 were 57%, 53%, 58%, and 52%, respectively with an overall average of 54% and standard deviation of 2.61.

Figure 1 shows that most students got grade B in the clerkship (68.8%), 16% of students got grade A, 13.2% got grade C, and only one student got F. Most students with grade A were from batch 2018 (78.1%), while the majority of students with grade D from batch 2019 (75%), (p=0.000). Most students with grades A and B were females (86.3% % and 51.8% respectively), on the other hand, most students with grades C and D were males (65% and 87.5% respectively) (p=0.000). There was no significant difference in scores earned by students from different semesters and different group orders. (Table 3). Table 4 demonstrated similar associations. Batch 2018 and females have significantly higher marks (p=0.000). On the other hand, no significant associations were seen between total marks and the semester in which the clerkship was taken.
Significant positive correlations were seen between all forms of assessment during the family medicine clerkship (p=0.000) (Table 5). Strong positive correlations were found between total marks and MCQs, EMQs, and OSCE examinations (r=0.857, 0.779 and 0.715, respectively). Moderate correlation was found between total marks and continuous assessment (r=0.508). Additionally, moderate correlations were seen between OSCE and all other assessment forms. Similarly, moderate correlations were observed between MCQs and EMQs.

**Discussion**

Exams are a crucial part of medical education to assess medical students. The combination of different methods of assessments helps students because they could have different strengths and weaknesses in different forms. A combination of various methods of assessment enhances the validity and reliability of evaluating medical students.8 Various assessment formats were used in the current study to measure different skills and levels of competencies in Miller’s pyramids.8 MCQs and EMQs exams were used to measure (“Knows” and “Knows how”) levels of Miller’s pyramid. For “Shows how” level, OSCE was used. Continuous course assessment measures the (“does”) level.8 Significant positive correlations were seen between all assessment methods in this study. This result could reflect individual student performance because student scores in different methods of the exam were correlated with the scores in the continuous course assessment. The outcome agrees with the Fallatah et al study that was
conducted among medical students who underwent internal medicine rotation at King Abdulaziz University.17

The results of this study showed that the female students performed better than male in overall scores. This result is consistent with the findings of other studies.14,18–20 Various factors could explain this finding, socially, females in a medical field are working hard on trying to prove themselves in a male-dominant society.21 Physiologically, females have better vigilance speed, associative memory, and environmental scanning.22 Biologically, the brain in male is larger and has deeper fissures and sulci than female. On the other hand, a woman’s brain has more white matter in areas responsible for intelligence.23–26 Similar explanation reported in Shams et al study that was conducted among medical students during anesthesia rotation in King Faisal University.20

This study revealed that the first batch (2018) scored higher than the second batch (2019). Moreover, most of the students who got grade A were from batch 2018, and most of the students with D grade from batch 2019; which further explains the fact that 2018 batch took the family medicine final exam with all other major specialties’ exam in the fifth year at the end of the year (with psychiatry, pediatrics, and obstetrics and gynecology) in the same day. Cumulative studying of these subjects could improve their performance as family medicine is a comprehensive specialty and covers all organ systems and different diseases.1 Moreover, Batch 2018 had better performance than 2019, it may be attributed to the observed better scores of batch 2018 across years and the batch was the first group in the implementation of the revised curriculum. A reassuring finding in this study was that no significant associations were seen between students’ marks and the timing of the clerkship rotation, whether taken in a first or second semester or a first or second group in during the semester. This lack of association ensures fairness between students regarding the planned time to take the clerkship during the academic year.

The strengths of our study included the use of recorded data rather than ask students about their grades; this limits the risk of recall bias. Moreover, IAU has adopted a new curriculum for its college of medicine at the undergraduate level. The change has moved from a traditional module-based medical education to a problem-based learning activity where students discuss cases from year one, additionally, including family medicine as a clinical rotation. Our target population was the first and second batches of students undergoing this new curriculum.7 Moreover, to our best of knowledge, this is the first study investigating different methods of assessment for family medicine clerkship in Saudi Arabia. However, there were some limitations. It was a cross-sectional study; therefore, causality could not be assessed. Moreover, data was collected from one university because we want to standardize the effect of the learning environment and because of variations in the curricula used by universities among the Kingdom. The study recommends future research to compare the outcomes of different assessment methods concerning different curricula.

Table 5 Correlation Matrix Between Students’ Marks in Different Methods of Assessment During the Family Medicine Clerkship, Imam Abdulrahman Bin Faisal University, Saudi Arabia, 2018–2019 (N=455)

|                        | Continuous Assessment | Total MCQs | Total EMQs | OSCE |
|------------------------|-----------------------|------------|------------|------|
| Continuous assessment  | P value               | –          | –          | –    |
| r coefficient          | –                     | –          | –          | –    |
| Total MCQs             | P value               | 0.285      | –          | –    |
| r coefficient          | 0.000                 | –          | –          | –    |
| Total EMQs             | P value               | 0.184      | 0.589      | –    |
| r coefficient          | 0.000                 | 0.000      | 0.000      | –    |
| OSCE                   | P value               | 0.278      | 0.496      | 0.322|
| r coefficient          | 0.000                 | 0.000      | 0.000      | –    |
| Total marks            | P value               | 0.508      | 0.857      | 0.779|
| r coefficient          | 0.000                 | 0.000      | 0.000      | 0.715|
|                        |                       |            |            |      |

Abbreviations: MCQs, Multiple choice questions; EMQs, Extended matching questions; OSCE, Objective structured clinical exam.

Conclusion
This study demonstrates that using multifaceted assessment of fifth-year medical students in family medicine clerkship applying different methods to evaluate competencies and knowledge during the course. Results were both additive and correlated. Results were not affected by chronology of the course relative to different other clinical rotations. The performance of females and the first batch were superior. The findings in this study are the basis for introducing a platform for developing modern learning and improving assessment standards in family medicine clerkship.
Abbreviations
OSCE, Objective Structured Clinical Examination; EMQ, extended matching questions; MCQ, multiple-choice questions; GPA, Grade Point Average; IAU, Imam Abdulrahman Bin Faisal University.

Ethics Approval
The college authorities approved the study. All participants gave their Informed consent. Confidentiality of the data was assured, and it will be used only for research purposes.

Acknowledgments
The authors would like to thank medical students in the College of Medicine, Imam Abdulrahman Bin Faisal University, for their participation. Also, we would like to thank Editage for English language editing.

Disclosure
The authors report no conflicts of interest in this study.

References
1. Aafp.org. Family medicine, definition of. [online]. 2019. Available from: https://www.aafp.org/about/policies/all/family-medicine-definition.html. Accessed December 23, 2019.
2. Mudiyanse RM. Need to teach family medicine concepts even before establishing such practice in a country. Asia Pac Fam Med. 2014;13(1):1. doi:10.1186/1447-056X-13-1
3. Bakshi S, Kaljee L, Parke D. A cross-sectional survey on medical education needs of general practitioners and family medicine: Delhi, Himachal Pradesh and Tamil Nadu, India. J Family Med Prim Care. 2019;8(11):3351–3357. doi:10.4103/jfmpc.jfmpc_644_19
4. The government of Saudi Arabia. National transformation program. 2016. Available from: http://www.vision2030.gov.sa/sites/default/files/NTP_En.pdf. Accessed October 20, 2019.
5. Al-Khaldi YM, Al-Ghamdi EA, Al-Mogbil TI, Al-Khashan HI. Family medicine practice in Saudi Arabia: the current situation and proposed strategic directions plan 2020. J Family Community Med. 2017;24(3):156–163. doi:10.4103/jfcm.JFCM_41_17
6. Alshahrani M, Dhafery B, Al Mulhim M, Alkhadra F, Al Bagshi D, Bukhamsin N. Factors influencing Saudi medical students and interns’ choice of future specialty: a self-administered questionnaire. Adv Med Educ Pract. 2014;5:397. doi:10.2147/AMEP.S69152
7. University council. Revised Study Plan and Course Specification Manual. 2nd ed. Dammam: IAU; 2014.
8. Tabish S. Assessment methods in medical education. Int J Health Sci (Qassim). 2008;2(2):3–7.
9. Harden R, Stevenson M, Downie WW, Wilson GM. Clinical competence in using objective structured examination. Br Med J. 1975;1:447–451. doi:10.1136/bmj.1.5955.447
10. Al-Osail AM, Al-Sheikh MH, Al-Said AH, et al. Objective structured clinical examination as an assessment method for undergraduate medical students. Int J Med Sci Public Health. 2015;4:192–198. doi:10.5455/ijmsh.2015.3110201437
11. Khan KZ, Gaunt K, Ramachandran S, Pushkar P. The Objective Structured Clinical Examination (OSCE): AMEE Guide No.81. Part II: organisation & administration. Med Teach. 2013;35(9):e1447–e1463.
12. Raheel H, Naeem N. Assessing the objective structured clinical examination: Saudi family medicine undergraduate medical students’ perceptions of the tool. J Pak Med Assoc. 2013;63(10):1281–1284.
13. Eijjsvogels T, van den Brand T, Hopman M. Multiple choice questions are superior to extended matching questions to identify medicine and biomedical sciences students who perform poorly. Perspect Med Educ. 2013;2(5–6):252–263. doi:10.1007/s40037-013-0068-x
14. Deepak KK, Al-Umran KU, Al-Sheikh MH, Al-Rubaish A. The influence of gender on undergraduate performance in multiple-choice testing in clinical disciplines at University of Dammam, Saudi Arabia. Al Ameen J Med Sci. 2011;4(2):123–130.
15. Amr M, Amin T. Assessment methods of an undergraduate psychiatry course at a Saudi university. Sultan Qaboos Univ Med J. 2012;12(2):214–220. doi:10.12816/0003115
16. Beullens J, Damme BV, Jaspaert H, Janssen PJ. Are extended-matching multiple-choice items appropriate for a final test in medical education? Med Teach. 2002;24(4):390–395. doi:10.1080/014215902100000843
17. Fallatah H, Tekian A, Park Y, Al Shawa L. The validity and reliability of the sixth-year internal medical examination administered at the King Abdulaziz University Medical College. BMC Med Educ. 2015;15(1). doi:10.1186/s12909-015-0295-4
18. Omigbodun O, Omigbodun A. Influence of gender on undergraduate performance in psychiatry at Ibadan, Nigeria. Med Educ. 2003;37(12):1091–1093. doi:10.1046/j.1365-2923.2003.01712.x
19. McDonough CM, Horgan A, Codd MB, Casey PR. Gender differences in the results of the final medical examination at University College Dublin. Med Educ. 2000;34(1):30–34. doi:10.1046/j.1365-2923.2000.00456.x
20. Shams T, El-Masry R, Al Wadani H, Amr M. Assessment of current undergraduate anaesthesia course in a Saudi University. Saudi J Anaesth. 2013;7(2):122. doi:10.4103/1658-354X.114049
21. Al-Tamimi DM. Saudi women in academic medicine. Are They Succeeding? Saudi Med J. 2004;25:1564–1567.
22. Boulet JR, McKinley DW, Norcini JJ, Whelan GP. Assessing the comparability of standardized patient and physician evaluations of clinical skills. Adv Health Sci Educ Theory Pract. 2002;7(2):85–97. doi:10.1023/A:1015750009235
23. Luders E, Narr KL, Thompson PM, et al. Gender differences in cortical complexity. Nat Neurosci. 2004;7(8):799–800. doi:10.1038/nn1277
24. Haier RJ, Jung RE, Yeo RA, Head K, Alkire MT. The neuroanatomy of general intelligence: sex matters. Neuroimage. 2005;25(1):320–327. doi:10.1016/j.neuroimage.2004.11.019
25. Benbow CP, Lubinski D, Shea DL, Eftekhari-Sanjani H. Sex differences in mathematical reasoning ability at age 13: their status 20 years later. Psychol Sci. 2000;11(6):474–480. doi:10.1111/1467-9280.00291
26. Leagey E. Gender differences in mathematical trajectories. Soc Forces. 2001;80:713–732. doi:10.1353/sof.2001.0102
