Original Research

Relation of the final year performance in different assessment modalities in final MBBS Pediatrics, College of Medicine, University of Bahri

Muawia E. A. Idris[1], Nour Elhouda Ata Alla Alla Gabo Rahma[1], Omer A. E. Elfaki[2], Waly Eldin Elnour Elfakey[3], Karimeldin Mohamed Ali Salih[3]

[1]Department of Pediatrics, College of Medicine, University of Bahri, Khartoum, Sudan, [2]Department of Medical Education, College of Medicine, King Khalid University, Abha, Saudi Arabia, [3]Department of Pediatrics and Medical Education, College of Medicine, University of Bahri, Khartoum, Sudan

Address for correspondence:
Karimeldin Mohamed Ali Salih, Department of Pediatrics and Medical Education, College of Medicine, University of Bahri, Khartoum, Sudan.
E-mail: chdcomkku@gmail.com

ABSTRACT

Background: Many tools and instruments can be used for the assessment of medical students including multiple choice questions (MCQs), modified essay questions (MEQs), objective structured practical examination (OSPE), problems, short answer questions (SAQs), extended matching questions (EMQs), and objective structured clinical examination (OSCE). The golden rule in assessment is to use multiple methods of assessment to provide students adequate opportunities to perform well. Using variable assessment tools will allow for making use of their advantages and reduce their disadvantages as much as possible. For correlation of combinations of assessment tools, alignment between objectives and instruction should be made. Many studies demonstrated different correlation with different forms of assessment, namely, MCQs, EMQs, MEQs, OSPE, OSCE, and SAQs. The objectives of this study were to determine the correlation of the students' scores in the four assessment modalities used for the final students in pediatrics and to determine the correlation of the combined score of written assessment and the combined score of practical assessment with the composite score. **Methodology:** This was a descriptive, cross-sectional study. A correlational design was used. The scores of 219 students in four methods of assessment adopted in pediatrics were analyzed. The assessment methods were MCQs, problems, OSPE, and OSCE. SPSS version 21 was used to compute the Pearson's correlation for each of the four methods with each other and with the composite scores. **Results:** The Pearson's r for the combined score of MCQs and problems with the composite score of all four assessment was 0.966 while 0.971 was the Pearson's r for the combined score for OSCE and OSPE with the composite score. Pearson's r of each of the assessment methods with composite scores were determined. These were 0.924, 0.901, 0.953, and 0.824 for MCQs, problems, OSCE, and OSPE, respectively. The correlation of each of the methods of assessment with each other revealed a Pearson's r correlation ranging from 0.977 to 0.819. All the correlations were highly significant. **Conclusion:** Analysis of the students' scores in the four assessment methods adopted in dermatology showed highly significant correlation. Scores in the written and practical assessment were highly correlated.

KEY WORDS: Assessment, objective structured practical examination, objective structured clinical examination, multiple choice questions, problem

INTRODUCTION

Assessment is usually carried out as formative (for diagnostic) and summative (as exist pass way). Many tools and instruments can be used including multiple choice questions (MCQs), modified essay questions (MEQs), objective structured practical examination (OSPE), problems, short answer questions (SAQs), extended matching questions (EMQs), and objective structured clinical examination (OSCE). The golden rule in assessment is to use multiple methods of assessment to provide students

adequate opportunities to perform well [1,2]. Assessment as agreed by different educators is not an easy job since it derives learning, determine human safety, needs to be active, planned, structured, including different instruments, and ensuring continuity [2-5]. Using variable assessment tools will allow for making use of their advantages and reduce their disadvantages as much as possible [6-8]. For correlation of combinations of assessment tools, alignment between objectives and instruction should be made [1]. Wass *et al.* [9] demonstrated different correlation with different forms of assessment, namely,

MCQs, EMQs, MEQs, and SAQs. They concluded that MCQs and SAQs might be suitable for testing knowledge and problem-solving domains, respectively. Adeniyi *et al.* [1,10] concluded that MEQs, MCQs, and practical assessment could predict well the overall score in descending order of importance. Bahri University (BU) has been established in 2011 as a product of the three universities: University of Juba-1977, Upper Nile University-1992, and University of Bahri Alghazal-1992. The three universities moved to the Republic of South Sudan after its separation from Sudan. The curriculum type adopted in the college was the SPICES model. The rationale for this study was to ensure that the assessment adopted for this course is sounded and could result in good quality of graduates. The objectives of this study were first to determine the correlation of the students' scores in the four assessment modalities used for the final students in pediatrics. Second, to determine the correlation of the combined score of written assessment and the combined score of practical assessment with the composite score.

METHODOLOGY

This was a descriptive, cross-sectional, correlation study conducted at the faculty of medicine, BU in the period June-October 2016. Ethical clearance was obtained from the Research and Ethics Committee the College. The researchers used the scores of the students in the different assessment methods used for pediatrics in the final MBBS cohort of students who graduated in the academic year 2015-2016. The assessment methods were as follows:

1. OSCE: Eight active stations of 10 min duration each. Real patients were used in six stations and simulated patients in two. The skills assessed included history taking and communication, general and focused examination, and interpretation of clinical findings to formulate a diagnosis and suggest a management plan.
2. OSPE: Twenty projected slides containing clinical material in the form of patient pictures, medical equipment, x-rays, and laboratory results. Two SAQs were asked on each slide.
3. MCQ: Fifty-one best answer types with five options. Most of the questions were application type testing higher order thinking.
4. Four problems testing problem solving and higher order thinking.

The OSCE and OSPE carried 50% of the total score with 40% and 10%, respectively. The MCQs and problems make the other 50% of the total score with 30% and 20%, respectively. The pass/failure decision and grades were based on the aggregate score of the four components. The data were accessed from the academic office where all the results of exams used to be monitored and stored. The scores were treated anonymously without identification of students' names. The data for the 219 students of the cohort were included in the study. All the data were entered into SPSS version 21. Pearson's correlation coefficient was calculated to determine the correlation between the variables and each variable with the total scores. $P < 0.01$ was taken as statistically significant.

RESULTS

Scores of 219 students were analyzed. The composite score was used to classify the students into the five-grade system adopted by the college. The failure rate was 23.3% [Table 1]. The Pearson's r for the combined score of MCQs and problems with the composite score of all four assessments was 0.966 while 0.971 was the Pearson's r for the combined score for OSCE and OSPE with the composite score. In Table 2, Pearson's r of each of the assessment methods with composite score were shown. These were 0.924, 0.901, 0.953, and 0.824 for MCQs, problems, OSCE, and OSPE, respectively. The correlation of each of the methods of assessment with each other revealed a Pearson's r ranging from 0.779 to 0.819 [Table 3].

DISCUSSION

The raw scores in the different assessments were taken here to form the composite score which formed the basis for grading. Although it had been recommended to standardize the scores or use weighting method to form a composite score, this can be compensated for if the raw scores were based on a blueprint [11]. The raw scores used in this study followed the blueprint where assessment methods used were aligned with the learning outcomes, and the weight of each assessment in the

Table 1: Students' grades distribution (*n*=219)

Grade	Number of students (%)
Distinction	2 (0.9)
Very good	19 (8.7)
Good	48 (21.9)
Pass	99 (45.2)
Fail	51 (23.3)
Total	219 (100)

Table 2: The Pearson's correlation coefficient values of the correlation of the different variables with the composite score

Correlated variable	Pearson's r for correlation with composite scores
MCQs	0.924
Problems	0.901
Combined score of MCQs and problems	0.966
OSPE	0.953
OSCE	0.894
Combined score of OSCE and OSPE	0.971

MCQs: Multiple choice questions, OSPE: Structured practical examination, OSCE: Structured clinical examination *P*<0.001

Table 3: The correlation of the different methods of assessment with each other

Methods of assessment	MCQs	Problems	OSCE	OSPE
MCQs/30	1	0.779**	0.800**	0.789**
Problems/20	0.779**	1	0.816**	0.812**
OSCE/40	0.800**	0.816**	1	0.819**
OSPE/10	0.789**	0.812**	0.819**	1

MCQs: Multiple choice questions, OSPE: Structured practical examination, OSCE: Structured clinical examination. **Correlation is significant at the 0.01 level (two-tailed)

final composite score was considered and aligned. The highly significant correlation of the combined written assessment score and the combined practical assessment scores with the composite score was quite evident (0.966 and 0.971, respectively). Although the weight of the practical assessment is less than the written assessment (40% vs. 60%) of the composite score, its correlation with the composite score is higher than the written. It was also quite impressive and interesting that the four modalities of assessment showed highly significant correlation with each other. This could be explained by the similarities of constructs being tested in the four methods and the good quality of written assessment items being capable of testing application and higher cognitive levels rather than recall and remembering only. The correlation of any one or more modalities of assessment with any different modalities or modalities had been researched in many contexts and different correlations reported [9]. Moderate correlation proved to occur between MCQs and MEQ scores in some studies [12,13]. A strong correlation between MCQs and SEQs has been reported ($r = 0.6$, $P < 0.01$) in other studies [14]. In case of the correlation of MCQS with OSCE, it had been believed that they were well correlated with [14]. In a large-scale multicenter study where 49 medical schools participated, Pearson's correlation coefficient between OSCEs and MCQs was 0.335 ($P = 0.00$) [15]. Sehlule *et al*. [1] studied the correlation of different modes of assessment of medical students in hematology and integrated exam of paraclinical sciences. Their modalities included MCQs, OSCE, OSPE, SAQs, and EMQs. Their findings confirmed the correlation of all the modalities. On the other hand, students' scores resulting from different assessment modalities may not be correlated as shown in other studies [16]. In some other studies, the overall reliability of correlating scores of different assessment modalities had been questioned. It had been pointed out these modalities might differ in many specifications including item format, content, number of items, duration, and the exact weight in the composite score.

The limitations of this study were covering one scores and taking one cohort, making generalize ability inappropriate.

CONCLUSION

Analysis of the students' scores in the four assessment methods adopted in dermatology showed highly significant correlation. Scores in the written and practical assessments were highly correlated.

ACKNOWLEDGMENT

We would like to convey our thanks and appreciation to our colleagues in the department for their nice cooperation, our student for their active participations, our clerk, and our statistician for their help in analysis and typing.

REFERENCES

1. Vuma S, Ramsewak BS. A retrospective co-relational analysis of third year MBBS students' performance in different modalities of assessment in haematology and final integrated examinations. Caribb Teach Sch 2015;5:37-464.
2. Salih KM, Alshehri MA, Elfaki OA. Comparison between students performance in multiple choice and modified essay questions in the MBBS pediatrics examination at the college of medicine, King Khalid University, KSA. J Educ Pract 2016;7:116-20.
3. Ebstien RM. Assessment in medical educational objectives. Engl J 2007;366:387-96.
4. Drew S. Student perceptions of what helps them learn and develop in higher education. Teach High Educ 2001;6:309-31.
5. Scouller K. The influence of assessment method on students' learning approaches: Multiple-choice question examination versus assignment essay. High Educ 1998;35:453-72.
6. Oyebola DO, Adewoye OE, Iyaniwura JO, Alada RA, Fasanmade AA, Raji YA. Comparative study of students' performance in preclinical physiology assessed by multiple choice and short essay questions. Afr J Med Sci 2000;29:201-5.
7. Nnodim JO. Multiple choice testing in anatomy. J Med Educ 1992;26:301-9.
8. Day SC, Norcini JJ, Diserens D, Cebul RD, Schwartz JS, Beck LH, *et al*. The validity of an essay test of clinical judgment. Acad Med 1990;65 9 Suppl: S39-40.
9. Wass V, McGibbon D, Van der Vleuten C. Composite undergraduate clinical examinations: How should the components be combined to maximize reliability? Med Educ 2001;35:326-30.
10. Adeniyi OS, Ogli SA, Ojabo CO, Musa DI. The impact of various assessment parameters on medical students' performance in first professional examination in physiology. Niger Med J 2013;54:302-5.
11. Rudner LM. Informed test component weighting. Educ Meas Issues Pract 2001;20:16-9.
12. Mujeeb AM, Pardeshi ML, Ghongane BB. Comparative assessment of multiple choice questions versus short essay questions in pharmacology examinations. Indian J Med Sci 2010;64:118-24.
13. Mahmood H. Correlation of MCQ and SEQ scores in written undergraduate ophthalmology assessment. J Coll Physicians Surg Pak 2015;25:185-8.
14. Hastings JK, Flowers SK, Pace AC, Spadaro D. An objective standardized clinical examination (OSCE) in an advanced nonprescription medicines course. Am J Pharm Educ 2010;74:98.
15. Rahayu GR, Suhoyo Y, Nurhidayah R, Hasdianda MA, Dewi SP, *et al*. Large-scale multi-site OSCEs for national competency examination of medical doctors in Indonesia. Med Teach 2016;38:801-7.
16. Wadi MM. Question vetting: Theory and practice. Educ Med J 2012;4(1):e1-4.