Comparison of different standard settings in introduction to medicine and medical education course at Bisha Medical College

Mohammed Abbas (1,2),Abdulla M.Alshahrani(3),Karimeldim Salih(1,4,5), Masoud Ishag, (4,6),Mutasim E. Ibrahim(7) A.M.S. Eleragi (7), Mohamed Ohaj(8),Bahaeldin K.Elamin (7,9)

(1) Pediatric department, College of Medicine, Bisha University

(2) Pediatric Department, Faculty of Medicine, Kasala University.

(3) Family Medicine, College of Medicine, University of Bisha

(4) Medical Education Department, College of Medicine, Bisha University

(5) Pediatric department, Faculty of medicine, University of Bahri.

(6) Depart of Medicine, College of Medicine, University of Bisha

(7) Microbiology Department, College of Medicine, University of Bisha

(8) Biochemistry department, College of Medicine, University of Bisha

(9) Microbiology department, Faculty of Medical Laboratory Sciences, University of Khartoum

Corresponding Authors:

1. Mohammed Abbas Mohammed Abbas

   College of Medicine, University of Bisha

E-mail: moh.abs.moh@gmail.com, mamohamed@ub.edu.sa **Mobile:** 00966500827887

College of Medicine

University of Bisha

P. O. Box 1290 Bisha 61922, Kingdom of Saudi Arabia

2. Dr. Bahaeldin K. Elamin

E-mail: bkelamin@ub.edu.sa

College of Medicine, University of Bisha

University of Bisha

P. O. Box 1290 Bisha 61922, Kingdom of Saudi Arabia

Mobile: +966509952588

Co-authors email address:

Abdulla M.Alshahrani   : ab_alshahrani@ub.edu.sa

Karimeldim Salih: karimeldin_salih@hotmail.com

Masoud Ishag: mieadam@ub.edu.sa

Mutasim E. Ibrahim: meibrahim@ub.edu.sa

A.M.S. Eleragi: ameleragi@ub.edu.sa

Mohamed Ohaj: momohamed@ub.edu.sa

Abstract

Introduction: Assessment is the corner stone in medical education, and a valid standard setting is urgently needed to ensure competent graduate to serve the community. Despite the existence of many types of standard setting, however at the undergraduate level, Angoff method remains the main widely studied and applied type of standard setting.

Objectives: To develop standard criteria for the pass marks in our settings using modified Angoff method, secondly to compare between students' performance on the fixed pass mark, norm-referenced method versus the modified Angoff method.

Methodology: This experimental prospective study which was conducted at the College of Medicine, where 48 students were enrolled in the study. The exam items were MCQs type A (30 questions) and seven medical education expert raters who were engaged in teaching of the course, were nominated to determine the pass scores of modified Angoff's method. All students who passed successfully the first year and attended 75% or more of the course content were enrolled and recruited in the present study.

Results: Reliability between the different judges were within the normal ranges, difference between modified Angoff and other with significant value and less students failed on using modified Angoff standard setting compared to other methods .

Conclusion: Modified Angoff should be adopted regard less of type of educational instruction to ensure accountability

Key Word: Modified Angoff, fixed methods, assessment, Norm- referenced

Introduction

The outcome of learning assessments should be determined using standard setting which is defined as the point where clear differentiation between the best and less examinee scores is achieved. Therefore it is of crucial importance to use a systematic way to gather valuable judgments, reach consensus and express that consensus as a single score on a test [1-3] . Moreover; in line with the above mentioned Cusimano1996 stated that standard setting is the only process of deciding what is good enough in terms of assessment evaluation [4].There are two main types of standard-setting which were widely used; item centered like Angoff, Ebel and Nedelsky, and person-centered methods which are represented by Borderline and Contrasting Groups methods [5]. Despite the existence of many types of standard setting, however at the undergraduate level, Modified Angoff method remains the main widely studied and applied type of standard setting [6].It is worth saying that Angoff, method was performed through nominating a board of subject experts to judge about to which extent minimally competent student will answer the questions, then judges can access the rating of other experts, followed by discussion to reach a consensus, where the agreed upon average from all the judges will be the standard setting to be adopted for the exam under investigation [7].

Norm-referenced method is the process of evaluating (and grading) the learning of students by judging (and ranking) them against the performance of their peers. It is used to evaluate the effectiveness of teaching programs and helps determining students' preparedness for programs [8]. However; it was reported to be insensitive to instruction and does not provide an estimate of the absolute level of performance achieved [9].

Fixed pass mark is determined either through judges who suggested the percentage where qualified examinees can score or through a responsible authority. It is worth saying that this method has no scientific basis, since it is just a non-defendable figure [1, 10]. Such invalid and unreliable pass mark can result in allowing non-competent candidates to practice and unrealistically high pass mark will exclude competent candidates [11]. The main aim of the current study to used modified Angoff method in order to develop standard criteria for the pass

marks in our settings, secondly to compare between students' performance on the fixed pass mark, norm-referenced method versus the modified Angoff method.

Methodology

This an experimental prospective study which was conducted at the College of Medicine, University of Bisha, Saudi Arabia, where 48 students were enrolled in the study Introduction to medicine and medical education course was introduced to the medical students at the second year (phase one).This course has three credit hours including four themes (Doctor,Health System in Saudi Arabia& Learning medicine, Instructional & assessment methods, Health services & medical education facilities and Medical research, publications & evidence-based medicine).Course contents were taught mainly by medical educationists and those who have an excellent experience in medical education disciplines. Usually the coordinator collect the questions according to the exam blueprint from the concerned instructors, then revised by student's assessment committee (SAC) for possible corrections of flaws and subsequently delivered to the examination officer for fine adjustment and design and kept in secured pox.

All students who passed successfully the first year and attended 75% or more of the course content were enrolled and recruited in the present study after filling consent form. Exclusion criteria include those who were absent and represent two students. The exam items were Multiple Choice Questions (MCQs) type A (30 questions) and seven medical education expert raters who were engaged in teaching of the course, were nominated to determine the pass scores of modified Angoff's method. The judges after the first round showed their results concerning the minimally competent students and after a second round they came to a consensus about their marks and an average was taken as modified Angoff. Then the exam questions were distributed to the concerned students in a proper examination environment, according to rules, the time allotted to the 30 MCQs was 60 minutes (2minutes per question) then after the question papers and answer sheets were collected and submitted to the examination officer for optical marking and further item analysis.

Statistical Analysis

The data were analyzed using to Statistical Package for the Social Sciences (SPSS) version 16 where the intraclass Correlation Coefficient was used to measure inter-rater reliability of the judges.

Results

A total of 48 medical students at the College of Medicine, University of Bisha were participated in this study. The mean (standard deviation) of the test scores was 62.1 (16.2 SD).Therefore, the cut-off passing score by the norm reference method was calculated as 46.0.

The seven scores from the judges for the modified Angoff's method were as follows: 45.5, 39.5, 44.3, 47.3, 38.7, 44.8 and 41.8 with high rate of reliability Table 1. The mean passing score obtained by the seven judges was 43.1. The inter-rater reliability of the seven judges was 0.77 (95% CI 0.62-0.88).This indicated of high inter rater reliability.

The cut-off score (43.1) given by the judges in modified Angoff's method was used to calculate pass/fail rate of 48 students. As shown in Table 2, the pass rate of the exam was 89.6% by modified Angoff's method, 85.4% by traditional method and 56.2% by fixed mark method.

Table 1. Mean scores, standard deviation and interrater reliability of seven raters on MCQs items assessing students at the course of introduction of medicine and medical education

Rater	Mean	SD	Cronbach's Alpha
1	45.5	11.4	.76
2	39.5	19.7	.71
3	44.3	13.6	.72
4	47.3	18.2	.78
5	38.7	24.6	.76
6	44.8	14.6	.76
7	41.8	19.3	.70
Modified Angoff's	43.1	17.3	0.77

Table 2: The percentages of medical students those are pass the exam of introduction of medicine and medical education three methods

Method	%(n) of students who passed the exam	P value according to modified Angoff in comparison to other methods
Fixed pass mark (60%)	56.2% (27/48)	Angoff against fixed pass mark< 0.007
The norm – reference method (mean-1 SD)	85.4% (41/48)	Angoff against norm – reference method< 0.001
Modified Angoff Method	89.6% (43/48)	

Discussion

To the best of our knowledge no much work was done to determine standards for assessing the outcome for our undergraduate medical candidates in a multiple-choice testing environment.

The current study applied modified Angoff's methodology for standard setting and yielded acceptable reliability and significant differences in student performance compared to the fixed pass mark method. Although the required reliability for the inter rater reliability should be more than 0.8 but in the current investigation, the agreement between the different judges showed reliability approximately 0.8which is accepted as indicator of consistency which fair for different opinion [4]. The outcomes for modified Angoff's were 89.6% (43/48) in comparison to 85.4% (41/48), 56.2% (27/48) for norm reference, fixed mark methods respectively with a p value of 0.007 indicating significant differences in the outcome. This results was in favor of modified Angoff's, a finding which was in agreement with a similar study performed 2006 by   Sanju *et*

al., in Birmingham medical school for 4th year medical students. Sanju *et al.,* reported that all student pass when applying modified Angoff's while only 85% passed with the norm-reference method [3]. Although the agreement was very obvious between Bisha and Birmingham studies however those who pass the exam in Sanju *et al.,* study were 100% in contrast to this recent work where only 89.6% passed. The following factors might contribute to this difference regarding those who pass: firstly the students in Bisha were junior students (second year) and Birmingham student were 4th year medical students, finally variation in the languages between the two groups cannot be ignored. Elfaki and Salih (2015) recently published a similar work for final year medical students in King Khalid University, Kingdom of Saudi Arabia (KSA) which showed a pass score for norm reference as 35% and for Angoff of 48%, where the findings of current study were 43.1%, 46.6% for Angoff'sand norm reference respectively [12].

Regarding the reliability with modified Angoff's, some authors considered more than 0.8 is needed for high stakes examination, however in this work we had a reliability approaching 0.8(.78) which in alignment with other studies where had the test-retest reliability (0.59–0.74) and (0.81 – 0.82) on testing and retesting [13].

The selected raters in this study were familiar with Angoff's method, the students, the curriculum and the course being assessed [14, 15] .The raters in this study were seven but still no concrete agreement in literature about the exact number who should do judging. Some authors-were very demanding in raters' involvement and requiring 5-30 judges to perform the job, others-were very practical since they reported the number of the raters should be determined according to the facility resources [16-18].

The attempt to apply the modified Angoff which is known to be sound, reliable, valid and accountable method to set the pass mark (PM) in our institution is highly requested. Since the PM in educational testing is the standard criterion that determines whether a student passes or fails an examination. This will in turn determine whether the student is considered competent enough or not, to serve the community.

In conclusion; Apart from modified Angoff's for standard setting, other methods were arbitrary, subjective, in contradistinction to modified Angoff which is objective with high face validity and

reflect the real life situation of the minimally competent students, and however the room is opened for further modification.

Abbreviations

SAC: student's assessment committee

MCQs: Multiple Choice Questions

SPSS: Statistical Package for the Social Sciences

SD: Standard Division

KSA: Kingdom of Saudi Arabia

PM: Pass Mark

Declarations

Ethics approval and consent to participate

Ethics approval was obtained from the college of medicine (University of Bisha) Ethics Committee approval 28/11/2017/ 38/367. Consent to participate and for publication was obtained from participants.

Consent to publish

Not applicable.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

No potential conflict of interest was reported by the authors.

Funding

Authors' Contributions

MA, AMA, KS and BKE designed the study. All authors obtained and ordered data. MEI and AMS ran statistical analysis. All authors participated in the discussion. MA, AMA, MO and KS wrote the paper. MI supported and reviewed the work. All authors read and approved the final manuscript.

Acknowledgement

References

1.   Kane M: **Validating the performance standards associated with passing scores**. *Review of Educational Research* 1994, **64**(3):425-461.
2.   Norcini JJ: **Setting standards on educational tests**. *Medical education* 2003, **37**(5):464-469.
3.   George S, Haque MS, Oyebode F: **Standard setting: comparison of two methods**. *BMC medical education* 2006, **6**(1):46.
4.   Cusimano MD: **Standard setting in medical education**. *Academic Medicine* 1996, **71**(10):S112-120.
5.   Nedelsky L: **Absolute grading standards for objective tests**. *Educational and Psychological Measurement* 1954, **14**(1):3-19.
6.   Talente G, Haist SA, Wilson JF: **A model for setting performance standards for standardized patient examinations**. *Evaluation & the health professions* 2003, **26**(4):427-446.
7.   Livingston SA, Zieky MJ: **Passing Scores: A Manual for Setting Standards of Performance on Educational and Occupational Tests**. 1982.
8.   Hussain S, Tadesse T, Sajid S: **Norm-Referenced and Criterion-Referenced Test in EFL Classroom**.
9.   Turnbull JM: **What is… normative versus criterion-referenced assessment**. *Medical teacher* 1989, **11**(2):145-150.

10. Kaufman DM, Mann KV, Muijtjens AM, van der Vleuten CP: **A Comparison of Standard-setting Procedures for an OSCE in Undergraduate Medical Education**. *Academic Medicine* 2000, **75**(3):267-271.

11. Rezigalla A: **Fixed pass mark: Time for change**. *Sudan Medical Monitor* 2016, **11**(4):119-119.

12. Elfaki OA, Salih KM: **Comparison of Two Standard Setting Methods in a Medical Students MCQs Exam in Internal Medicine**. *American Journal of Medicine and Medical Sciences* 2015, **5**(4):164-167.

13. Impara JC, Plake BS: **Teachers' ability to estimate item difficulty: A test of the assumptions in the Angoff standard setting method**. *Journal of Educational Measurement* 1998, **35**(1):69-81.

14. Norcini JJ: **Research on standards for professional licensure and certification examinations**. *Evaluation & the Health Professions* 1994, **17**(2):160-177.

15. Norcini J, Shea J: **The reproducibility of standards over groups and occasions**. *Applied Measurement in Education* 1992, **5**(1):63-72.

16. Norcini JJ, Shea JA: **The credibility and comparability of standards**. *Applied Measurement in education* 1997, **10**(1):39-59.

17. Hurtz GM, Hertz NR: **How many raters should be used for establishing cutoff scores with the Angoff method? A generalizability theory study**. *Educational and Psychological Measurement* 1999, **59**(6):885-897.

18. Cizek GJ: **An NCME instructional module on: Setting passing scores**. *Educational Measurement: Issues and Practice* 1996, **15**(2):20-31.

IJSER