Niger. J. Physiol. Sci. 29(December2014) 103-106
www.njps.com.ng

Niger. J. Physiol. Sci.

Comparison of Outcome of Students' Performance Using the Standard Setting Method with the Absolute Grading Method in Preclinical Examination

Dagogo J. Pepple

Department of Basic Medical Sciences (Physiology Section), University of the West Indies, Mona Campus, Kingston 7, Jamaica, W.I.

Summary: This study compared the outcome of students' performance using the standard setting method with the equivalent outcome they would have obtained using the absolute grading method. It involved the comparison of fail, pass, honors and distinction grades in Digestive System, Endocrine System, Cardiovascular System and Health and the Environment courses in the MBBS Stage I examination. The performance in Cardiovascular System was significantly better with the standard setting method ($\chi^2 = 27.53$; $p < 0.01$), median score in the honors range compared with the absolute grading method where the median score was in the pass range. On the other hand, the performance in Endocrine System was significantly better using the absolute grading method ($\chi^2 = 27.30$; $p < 0.01$), with median score in the honors range compared with the standard setting method where the median score was in the pass range. There was no difference in the performance in Digestive System ($\chi^2 = 7.45$; $p = 0.06$), median score in the pass range and Health and the Environment ($\chi^2 = 6.34$; $p = 0.09$), median score in the honors range; between the standard setting and absolute grading methods (Wilcoxon's signed rank). The overall pass, honors, distinction and failure rates were also identical in both methods (Mann Whitney U test). This suggests that overall the outcome of the students' performance in the standard setting method compared with the absolute grading method were not significantly different.

Keywords*:* standard setting, absolute grading method, students' performance, preclinical education.

**Address for correspondence:* dagogo.pepple@uwimona.edu.jm Tel: 876-977-2560

Manuscript Accepted: November 2014

INTRODUCTION

The Faculty of Medical Sciences at the University of the West Indies (UWI) awards the Bachelor of Medicine and Bachelor of Surgery (MBBS) degree after five years of study. The MBBS Stage I examination is the final examination taken at the preclinical level before the students proceed to the last two years of their clinical training. It is equivalent in content and scope to the United States Medical Licensing Exam (USMLE) Step 1 examination.

At the Mona Campus, there has been significant curriculum reform, which includes a change from discipline-based to systems-based teaching in the preclinical years, with early clinical-preclinical integration. A part of this reform was a change in the marking system from the absolute grading to the standard setting method (Branday and Carpenter, 2008).

The aim of standard setting is to minimize errors in determining the pass/fail, honors and distinction cut off points while accounting for the varying difficulty of the examinations. Valid and reliable tools for assessing the quality of medical education are central to high quality medical care (Boursicot and Roberts, 2006). A standard may be absolute or criterion- referenced, where it is based on pre-determined criteria, irrespective of examinee performance or relative, that is, norm-referenced, where it is dependent on the performance of the particular group of examinees (Bandaranayake, 2008).

Standard setting is the process of deciding "what is good enough" and the process used to arrive at such decision is paramount (Cusimano, 1996). A standard is a special score that serves as a boundary between those who perform well enough and those who do not (Norcini, 2003). The standard setting grades define the boundary between passing and failing, that is, it separates candidates who are competent from those who are not. Additionally, the cutoff marks to achieve honors and distinction grades can also be decided by this process. The Angoff method of standard setting is currently being used at the Faculty of Medical Sciences, The University of the West Indies, Mona Campus. It is objective, has the advantages of being used in a range of licensing

and certifying examinations and well supported by research. However, it can be very labor intensive and time consuming (George et al, 2006).

The absolute grading system, on the other hand, is based on the idea that grades should reflect mastery of specific knowledge and skills. The teacher sets the criteria for each grade and all students who perform at a given level receive the same grade. This system is a "percent of total points possible" and it is assumed that a student who scores 85% knows 85% of the material. Traditionally, in the old curriculum at UWI using the absolute grading method, students achieving a score of 50% and above passed the examination while those scoring 65% and over were awarded honors and those with scores above 75% were awarded distinctions. One weak point associated with the absolute grading system is the rationale for the cut-off or pass/fail scores that is not based on any analysis (Grading Systems, 1991).

We hypothesized that there was no difference in the outcome of students' performance that is pass/fail, honors and distinction cut off marks, using the standard setting method when compared with the absolute grading method. Based on the foregoing therefore, the present study was designed to compare the outcome performance on the MBBS Stage I examination of preclinical students at the Mona Campus using modified Angoff method of standard setting compared with the equivalent outcome they would have obtained using the absolute grading method.

MATERIALS AND METHODS

In this retrospective study, we collected from the office of the MBBS Program Director the standard setting grades of all preclinical students who took the MBBS Stage I examination in December 2007. There were 131 students who sat for the Endocrine System and Health and Environment courses of the exam, and 130 students who sat for the Digestive System and Cardiovascular System courses. The examination was made up of 50 multiple choice questions (MCQ) in the Digestive System course with a duration of one hour; 75 MCQ in Health and Environment course with a duration of 1.5 hours and 100 MCQ in both Endocrine System and Cardiovascular System courses with a duration of 2 hours each. All scores were converted to a 100% scale score. They were compared with equivalent scores that the students would have obtained via the absolute grading method.

At the Mona campus, the modified Angoff method of standard setting is used.

In this method, a minimum of eight judges made up of experienced lecturers in the different departments who were involved in teaching a particular course formed the panel for that course. The different systems are "team taught" by lecturers from the respective disciplines. The judges had undergone a two day training session in standard setting method. They examine each multiple choice question with a four option choice to select the correct answer from and estimate the probability that the "minimally competent" or "borderline" candidate would answer the item correctly. The lecturers answer the questions within the time allotted for the examination and each judge's estimate scores on all items are added up and averaged and the test standard for that course is the average of these means for all the judges. Questions that were either ambiguous or poorly framed were corrected, while those that were either too difficult or too easy were replaced. The judges therefore determine the pass mark and also the marks required to achieve honors and distinction.

We analyzed the data using the SPSS statistical package for medians, range, Wilcoxon's signed rank test for two (paired) related data, the Mann Whitney U test and Chi squared test with $p < 0.05$ taken as the level of statistical significance. The study was approved by the Ethics Committee of the Faculty of Medical Sciences, The University of the West Indies, Mona Campus.

RESULTS

In the standard setting system, the median score of 60% for Digestive System and 66% for Endocrine System were in the pass range while it was in the honors range for Cardiovascular System (63%) and Health and Environment (67%). With the absolute grading method, the median score of 61% for Digestive System and 62% for Cardiovascular System were in the pass range while it was in the honors range for Endocrine System (65%) and Health and Environment (69%). The performance in Cardiovascular System was significantly better with the standard setting method ($\chi^2 = 27.53$; $p = 0.01$), median score being in the honors range; compared with the absolute grade method where the median score was in the pass range. On the other hand, the performance in Endocrine System was significantly better using the absolute grade method ($\chi^2 = 27.30$; $p = 0.01$), with median score in the honors range, compared with the standard setting method where the median score was in the pass range. There was no difference in the performance in Digestive System ($\chi^2 = 7.45$; $p = 0.06$) with the median in the pass range, and Health and Environment ($\chi^2 = 6.34$; $p = 0.09$), with the median in the honors range, between the standard setting and absolute grading methods (Wilcoxon's signed rank). The overall pass, honors, distinction and failure rates are also similar in both methods (Mann Whitney U test) (Table 1).

Table 1. Outcome of students' performance using the standard setting method and the absolute grading method

Courses	Standard setting	Absolute grading	Chi – square
Digestive System (N = 130)	F (0-46%) = 8 P (47-64%) = 77 H (65-78%) = 42 D (79-100%) = 3 Median = 60% Range: 34-84%	F (0-49%) = 16 P (50-64%) = 67 H (65-74%) = 37 D (75-100%) = 10 Median = 61%	$\chi^2 = 7.48$ p = 0.06
Endocrine System (N = 131)	F (0-45%) = 2 P (46-49%) = 92 H (70-78%) = 33 D (79-100%) =4 Median = 66% Range: 42-87%	F (0-49%) = 7 P (50-64%) = 53 H (65-74%) = 52* D (75-100%) = 19 Median = 65%	$\chi^2 = 27.30$ p = 0.01
Cardiovascular System (N = 130)	F (0-42%) = 3 P (43-56%) = 40 H (57-66%) = 46* Median = 63% Range: 49-79%	F (0-49%) = 1 P (50-64) = 50 H (65-74) = 55 Median = 62%	$\chi^2 = 27.53$ p = 0.01
Health & Environment (N = 131)	F (0-49%) = 1 P (50-67%) = 70 H (68-77) = 43 D (78-100%) = 17 Median: 67% Range: 49-79%	F (0-49%) = 1 P (50-64) = 50 H (65-74) = 55 D (75-100) = 25 Median: 69%	$\chi^2 = 6.33$ p = 0.09
Summary			
Failure rate	4%	10%	
Pass rate	47%	43%	
Honors rate	35%	32%	
Distinction rate	14%	15%	

F = Fail; P = Pass; H = Honors; D = Distinction *p < 0.01

DISCUSSION

The result of the present study showed that overall there were no significant differences in the outcome of the student' performance in all the four courses examined between the standard setting and absolute grading methods. This might also suggest that the two methods of assessment test identical knowledge base of the students. The outcome of the students' performance in Cardiovascular System was better with the standard setting method while in Endocrine System it was better with the absolute grading method. Their performance in Digestive System and Health and Environment were identical in both methods.

It is interesting to observe that the pass score mark set with the standard setting method in three of the four courses is less than the fifty percent pass mark used in the absolute grading system. However, the score required to achieve a distinction grade with the standard setting method is higher in the standard setting method than in the absolute grading method in three of the four courses examined. This might suggest that the standard setting method favors more students passing an examination but require a bit more from them to achieve distinction grade.

The overall pass rate (i.e. students who obtained pass, honors and distinction grades) of 96% in the standard setting method in the present study is similar to the 100% pass rate reported by George et al (2006) using the modified Angoff method. Also there is no difference between the 96% pass rate for the standard setting method and the 90% pass rate for the absolute grade method in the present study.

Impara and Plake (1998) had suggested that most of the judges often find it difficult to accurately conceptualize borderline candidates. The variability seen in the cut off pass/fail, honors and distinction marks of the respective courses could be attributed to the varied perceived level of difficulty by the different judges.

The process of standard setting, although very labor intensive and time consuming is justified because of its usefulness to the students, faculty and university at large as part of an internal quality assurance process intended to improve the validity of pass/fail, honors and distinction scores, bring teachers together to discuss course objectives, relevance and assessment. This is in contrast to the absolute grading method where the pass/fail, honors and distinction scores are pre-determined based on traditional or historic antecedents.

The examination format being used presently at the MBBS Stage I of the Faculty of Medical Sciences University of the West Indies, Mona Campus is the multiple choice. One set back of this examination format is that graduates produced mainly through the

multiple-choice standard setting format, without written or essay examinations may be deficient in communication, especially in writing (Bassaw and Pitt-Miller, 2007).

In conclusion, there is no significant difference in the outcome of the performance of the students between the absolute grading and standard setting methods suggesting that they might be testing identical knowledge base. The use of the standard setting method is still in its infancy in many medical schools and there are others yet to adopt this method of assessment. The adoption and continued usage of the standard setting method is supported because of its objectivity and validity in setting the pass/fail, honors and distinction scores and its usefulness in quality assurance and bringing teachers together to discuss course objectives and assessment.

Acknowledgements

Thanks to Dr Tomlin Paul, Program Director and Dr Anthony Frankson, Statistician, Faculty of Medical Sciences, The University of the West Indies, Mona Campus for assistance with data acquisition and statistical analysis respectively.

REFERENCES

Bandaranayake RC. (2008). Setting and maintaining standards in multiple choice examinations: AMEE Guide No. 37. *Med. Teach.* 30: 836-845.

Bassaw B, Pitt-Miller P. (2007). Modernizing medical education, perspective from a developing country. *West Indian Med. J.* 56: 80-85.

Boursicot K, Roberts T. (2006). Setting standards in a professional higher education course: Defining the concept of the minimally competent student in performance based assessment at the level of graduation from medical school. *Higher Educ. Quarterly* 60**:** 74-90.

Branday JM, Carpenter RA. (2008). The evolution of undergraduate medical training at the University of the West Indies, 1948-2008. *West Indian Med. J.* 57: 530-536.

Cusimano MD. (1996). Standard setting in medical education. *Acad. Med.* 71: (10 suppl): S112-120.

George A, Haque MS, Oyebode F. (2006). Standard setting: Comparison of two methods. *BMC Med. Educ.* 6: 46-53.

Grading Systems (1991). Center for Teaching and Learning, University of North Carolina.

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

Norcini JJ. (2003). Setting standards on educational tests. *Med. Educ.* 37: 464-469.