Caribbean Teaching Scholar
Vol. 5, No. 1, April 2015, 37–46

ERA
EDUCATIONAL RESEARCH ASSOCIATION

A retrospective co-relational analysis of third year MBBS students' performance in different modalities of assessment in haematology and final integrated examinations

Sehlule Vuma[a], Bidyadhar Sa[b] and Samuel Ramsewak[b]

[a]*Department of Para-clinical Sciences, The University of the West Indies, St Augustine, Trinidad and Tobago;* [b]*Centre for Medical Sciences Education, The University of the West Indies, St Augustine, Trinidad and Tobago*

This paper reports an analysis of correlations in students' performance in different modalities of assessment in haematology and multi-specialty (anatomical pathology, chemical pathology, haematology, immunology, microbiology and pharmacology) final integrated examinations. It is broadly agreed among medical educators that proper alignment between learning objectives, modes of delivery and assessment modalities is a key factor in shaping the desired outcomes. It is equally important that modalities of assessments are in concurrence among themselves within the assessment framework. A descriptive retrospective correlational analysis of 159 third-year Bachelor of Medicine and Bachelor of Surgery (MBBS) students' results in different assessment modalities in five courses covering Applied Para-clinical Sciences, Integrated Para-clinical Sciences and Basic Para-clinical Sciences was performed. Results show positive correlations amongst all haematology components as well as the final integrated examination and the continuous assessment element had the strongest correlations with the total haematology component. It was concluded that combinations of multiple modes of assessment are important for adequate and fair assessment of knowledge and skill and that continuous assessment encourages students to work consistently throughout the course.

Key words: Assessment, correlations, haematology, performance

Introduction

It is broadly believed that increasing the number of different forms of assessment helps student attainment, as students have different strengths in different assessment methods. Different medical schools use different combinations of multiple choice questions (MCQs), extended matching questions (EMQs), essays or free response short answer questions (SAQs), progressive disclosure questions (PDQs), observed structured clinical examinations (OSCE), observed structured practical examinations (OSPE) and oral examinations. Different forms of assessment may be better suited to assessing different objectives in the curriculum. Some forms of assessment have advantages over others. MCQs have several advantages over SAQs: they are easy to mark as marking is done by machine, and they avoid the marking discrepancies that might come with SAQs. Constructing MCQs, however, is more

time consuming for the teachers: though less time consuming for students to answer. MCQs are able to examine more content in the curriculum, whereas, SAQs present the problems of sampling. Whatever the chosen modality, assessments need to be reliable and valid. Epstein (2007) said using multiple methods of assessment can also assess students' learning needs and help to identify those that need help. However Epstein also reports that the question of summative versus formative assessments remains challenging. Different schools thus use different combinations of assessment to cover both the depth and breadth of their curriculum, both in terms of knowledge and skills depending on modes of delivery of the curriculum. von Bergmann, Dalrymple, Wong & Shuler (2007), for example, used MCQs and a computer-based objective test for assessment of their problem based learning content and they showed the combination to be significantly reliable.

Whatever the chosen combinations of assessment modalities, it is important that they are well aligned, hence correlations should be performed. Wass, McGibbon & Van der Vleuten (2001) showed that correlations differed between different forms of assessment. They showed the correlation between MCQs and EMQs was 0.43, between MCQs and SAQs was 0.46 and between EMQs and SAQs was 0.60. They reported that MCQs might be more suited for testing factual knowledge and SAQs more suitable for problem solving skills. Adeniyi, Olgi, Ojabo and Musa (2013) showed that in a physiology course continuous assessment (CA) (also known as in-course or formative assessment), had the best correlation (r = 0.801), while the oral examination had the least correlation (r = 0.277), with overall performance. They also showed that the essay component was the best predictor of overall performance (r = 0.421), followed by MCQ, while the practical component was the least reliable predictor of performance. Similar results were shown in a third year internal medicine course by Gupta et al (2007) who showed that in-course assessments correlated positively with end of semester assessments. As seen with essays and orals, correlational analysis is more varied in assessments that evaluate practical skills like the OSCE due to inter-observer subjectivity. Inter-rater reliability limits the ability to achieve high correlations (Campos-Outcalt, Watkins, Fulginiti, Kutob & Gordon, 1999), highlighting the challenge for examiners to increase objectivity in the OSCE. Pepple, Young and Carroll (2010), in a study among physiology students, showed that, in the Caribbean, for most students, the strong correlation between results in MCQs and essay questions was independent of the assessment modality.

Our study focuses on third year Bachelor of Medicine and Bachelor of Surgery (MBBS) students, at a major university in the Caribbean. No similar studies have been done in our setting before. The study provides useful information for the faculty as a whole. In this setting third year students take Applied and Integrated Para-clinical Sciences (pathology) courses which integrate the sub-specialties of anatomical pathology, chemical pathology, haematology, immunology, microbiology and pharmacology. Para-clinical Sciences bridge the gap between the pre-clinical and the clinical years. The study evaluates the students' performance in the haematology components against the final integrated examinations.

In the current system, first-year students take a six-week introductory integrated course, Basic Para-clinical Sciences (BPS) (see Table 1). In haematology, students are introduced to the basic concepts in their first year. In the second year there is no haematology teaching (there is some teaching in the other sub-specialties in the second year). In third year students discuss the pathophysiological basis of disease applying the basic knowledge from first year.

Table 1. Courses in the Para-clinical Sciences and details of haematology component of these courses

Year	Course	Semester	Haematology component: Course content	Course teaching/ delivery methods
1	Basic Para-clinical Sciences (BPS)	Semester 1	• Peripheral blood and blood components, haemopoiesis, introduction to haemostasis, introduction to basic laboratory procedures in haematology	• Problem based learning • Didactic lectures
3	Applied Para-clinical Sciences I (APS-I)	First half of Semester 1	• Anaemias	• Problem based learning • Didactic lectures
3	Applied Para-clinical Sciences II (APS-II)	Second half of Semester 1	• Haemostasis, thrombosis and transfusion medicine	• Problem based learning • Didactic lectures
3	Applied Para-clinical Sciences III (APS-III)	Semester 2	• Haematological malignancies and bone marrow failure syndromes	• Problem based learning • Didactic lectures
3	Integrated Para-clinical Sciences (IPS)	Semesters 1 & 2 (Rotations throughout the year)	• Integrates APS-I, II and III	• Clerkships • Clinical and practical application and exposure

In haematology students are exposed, for the first time in their medical training, to real live patients. In the other sub-specialties there is no live patient contact at this stage (except for isolated laboratory procedures). The live patient exposure in haematology is aimed at giving the students practical application of the material covered in lectures and the material introduced in problem based learning scenarios. Such exposure is designed to increase student interaction and understanding.

Medical schools are placing more and more emphasis on skills training (Wass, McGibbon & Van der Vleuten, 2001) as a means of developing competency. Tools for the assessment of skills include in-course assessments (known here as "continuous assessment") as well as final end of course examinations. Each sub-specialty contributes equally in the combined integrated examinations (see Table 2). Assessments need to assess factual knowledge, clinical and practical application of knowledge and need to be well correlated. Prior to 2010, SAQs were used. However because of issues raised by teachers regarding discrepancies in marking and the time needed to facilitate SAQs, they were replaced by EMQs and PDQs for third year courses in 2010. Similarly oral examinations were also removed from

the syllabus. Correlations between the old and new combinations of assessment modalities in the current system need to be analysed.

Table 2. First and third year course assessments in multi-specialty integrated course of Para-clinical Sciences

Year		Continuous assessment (CA)		End of course		Final score
1	Basic Para-clinical Sciences (BPS)	*/+Observed practical skills examination (OPSE)	*Problem-based learning	*Multiple choice questions (MCQs)	*Short answer questions (SAQ)	100%
3	Applied Para-clinical Sciences I (APS-I)	**Progressive disclosure questions (PDQ) (20%)	*Problem-based learning (5%)	*MCQ (50%)	**Extended matching questions (EMQ) (25%)	100%
3	Applied Para-clinical Sciences II (APS-II)	** PDQ (20%)	*Problem-based learning (5%) *	*MCQ (50%)	** EMQ (25%)	100%
3	Applied Para-clinical Sciences III (APS-III)	** PDQ (20%)	*Problem-based learning (5%)	* MCQ (50%)	** EMQ (25%)	100%
3	Integrated Para-clinical Sciences (IPS)	Varies by sub-specialty (25%) [Includes different combinations of case presentation, case write-up, MCQs and ***OSPE]		*OSPE (45%)		70% Clinical skills assessed (30%)

**Old assessment modality; **Newly introduced assessment modality; ***Haematology component of CA; +Not included in this analysis;*

Research question

What is the correlation between the different assessment modalities in haematology and the multi-specialty integrated examinations in the first year and third year Para-clinical Sciences courses?

Methodology

This study was conducted with 159 third year students during the academic year 2010-2011. Ethical approval was obtained from the Ethics Committee, and the Office of the Dean of the Faculty of Medical Sciences. This was a descriptive retrospective, correlational, cross-sectional analysis. Data was accessed from the Assessment Unit of the Faculty maintaining the anonymity of the students. The data included the results of the haematology components; the overall results of the integrated third year courses APS-I, APS-II, APS-III and IPS, and the results from the first year course, BPS. The focus on haematology in this integrated setting, was due to convenience, as the principal investigator is a lecturer in haematology. This study could be used as an initial base line for the rest of the department. Students who did not take all four third year courses, and students who did not start first year with this current group, were excluded from the study. For the IPS, only the

pathology clerkships were analysed and the results of the clinical skills component were not analysed. The CA results for the BPS course were not available for analysis. Correlations between the haematology components of the different assessments and the total haematology component, as well as the multi-specialty final integrated examinations results of the courses were analysed. Correlational analysis was performed by the Pearson product moment correlation method. Linear regression analysis was also performed between the total haematology component and the multi-specialty final integrated examination.

Results

Table 3 shows the breakdown of the different assessments in the department; the breakdown of the haematology components of these assessments, and the students' haematology mean scores in each assessment.

Table 3. Course assessments and scores in the haematology components of BPS, APS I, II & III and scores in multi-specialty final integrated examinations

	CA	**MCQ**	**SAQ**	**EMQ**	**OSPE**	**THC**	**FIE**	
First year - Basic Para-clinical Sciences (BPS) (CA-OSPE)								
Total score possible	NA	9	5			14	100%	
Mean		4.5 (50%)	2.89 (57.8%)			7.38 (52.7%)	63.98%	
Std deviation		1.618	.99			2.099	7.63	
Third year – Applied Para-clinical Sciences I (APS-I) (CA-PDQ)								
Total score possible	6	9		3		18	100%	
Mean	3.6 (60%)	4.15 (46.1%)		2.58 (86%)		10.29 (57.2%)	62.53%	
Std deviation	1.22	1.47		.741		2.50	9.35	
Third year – Applied Para-clinical Sciences II (APS-II) (CA-PDQ)								
Total score possible	15	8		4		27	100%	
Mean	11.42 (76.1%)	4.44 (55.5%)		2.58 (64.5%)		18.36 (68%)	66.65%	
Std deviation	2.71	1.42		1.18		4.09	9.02	
Third year – Applied Para-clinical Sciences (APS-III) (CA-PDQ)								
Total score possible	15	14		3		32	100%	
Mean	9.58 (63.9%)	9.73 (69.5%)		2.36 (78.7%)		21.61 (67.5%)	69.99%	
Std deviation	2.28	2.10		.71		3.97	8.55	
Third year - Integrated Para-clinical Sciences (IPS) (CA-OSPE)								
	OSPE						OSPE	FIE
Total score possible	8				7	15	45%	70%
Mean	4.80 (60%)				3.90 (55.7%)	8.69 (57.9%)	28.23%	45.97%
Std deviation	1.50				1.30	2.11	4.53	5.65

CA-continuous assessment, SAQ-short answer questions, PDQ-progressive disclosure questions, MCQ-multiple choice questions, EMQ-extended matching questions, OSPE-observed structured practical examination, THC-total haematology component, FIE-final integrated examination

For APS-I and APS-III the EMQs had the best mean scores. For APS-II the CA had the best mean scores. For the overall multi-specialty final integrated examination the mean scores for BPS, APS-I, and APS-III are higher than the total haematology component. Table 4 shows the correlations between the haematology components of different assessments, the total haematology component as well as the multi-specialty final integrated examination.

Table 4. Correlations between the haematology components of the different assessments, total haematology components, and multi-specialty final integrated examinations

	Haematology components					**All sub-specialties combined**	
	SAQ	**MCQ**	**EMQ**	**OSPE (Final)**	**THC**	**Final integrated examination**	
	First year - Basic Para-clinical Sciences (BPS)						
MCQ	.248**				.889**	.600**	
SAQ					.663**	.456**	
THC						.679**	
	Third year - Applied Para-clinical Sciences I (APS-I)						
PDQ (CA)		.225**	.208**		.695**	.480**	
MCQ			.292**		.802**	.592**	
EMQ					.559**	.503**	
THM						.750**	
	Third year - Applied Para-clinical Sciences II (APS-II)						
PDQ (CA)		.317**	.233**		.866**	.634**	
MCQ			.316**		.640**	.574**	
EMQ					.587**	.533**	
THC						.800**	
	Third year - Applied Para-clinical Sciences III (APS-III)						
PDQ (CA)		.318**	.255**		.794**	.565**	
MCQ			.354**		.794**	.679**	
EMQ					.519**	.527**	
THC						.784**	
	Third year - Integrated Para-clinical Sciences (IPS)						
						Final integrated OSPE	**Final integrated exam**
OSPE (CA)				.134	.794**	.233**	.314**
Final OSPE					.709**	.583**	.550**
THC						.523**	.561**
Final Integrated OSPE							.960**

*** Correlation is significant at the 0.01 level (2-tailed)*

*CA-continuous assessment; SAQ-short answer questions; PDQ-progressive disclosure questions; MCQ-multiple choice questions; EMQ-extended matching questions; OSPE-observed structured practical examination; THC-total haematology component** Correlation is significant at the 0.01 level (2-tailed)*

In BPS, APS-I, APS-II, and APS-III, between individual assessments (SAQ, PDQ, MCQ, and EMQ), correlations were weak to moderately strong (range r=.208 to .354). In IPS, between the haematology observed structured practical examination and the haematology component of the observed structured practical examination, correlations were negligible with r=.134. However between the individual components and the total haematology component and the final integrated examination, correlations were strong to very strong (range r=.456 to .889). The final integrated OSPE and the final integrated IPS examination had the highest correlations of r=.960. The MCQs had higher correlations than EMQs with the total haematology component and multi-specialty final integrated examination for all three courses APS-I, II and III. The CA for haematology in APS-II, APS-III and IPS had the strongest correlations with the total haematology component.

Discussion

There are positive correlations among all the different haematology components of the different assessments, as well as the final integrated multi-specialty examinations. This suggests that the different combinations of the current assessment forms are useful both in haematology alone and in the multi-specialty courses. Combinations of multiple assessment modalities are known to be better for the students than the individual assessments; since students may be stronger in some assessment modes than in others.

The introduction of EMQs and PDQs fits in well with the rest of the assessments and we are able to examine more depth and breadth of the curriculum in the third year. This combination eliminated some of the problems we had in the past with SAQs. For example, the problems of sampling with SAQs meant that less of the curriculum was examined. Hence, SAQs may help students attempt to 'spot' what might be in the examination and not study the whole course content. While no similar study has been performed involving correlations with SAQ used prior to 2010 for the third year courses, the advantages of not using this assessment modality are well noted. The correlations performed in the first year course give a control, though limited, for comparison. The correlations between the MCQ and SAQ in BPS was r=.248. Whereas, for the third year courses, the correlations between the MCQ and PDQ ranges between r=.225 and r=.318, and MCQ and EMQ correlations range between r=.292 and r=.354. This is higher than the SAQs, while reducing the time taken by students to answer SAQ, as well as by the staff to mark them, with fewer discrepancies in marking.

EMQs are a form of MCQs and have many of the advantages of MCQs (objectivity, computer marking) but transform the questions into items that help engage students in higher level mental tasks and ask students to solve problems rather than recall isolated pieces of information (which is in keeping with our problem based learning method of teaching). EMQs can also help to prevent students answering by elimination rather than actually knowing the answer.

Fenderson, Damjanov, Robeson, Veloski and Rubin (1997) report that EMQs help in reducing the effect of 'cueing' where students can answer a question by recognizing the correct option that they otherwise would not have answered without being given the options. Furthermore, item analysis can be performed to demonstrate reliability and validity. Wass, McGibbon and Van der Vleuten (2001) studied the construct validity of EMQs and found they could measure clinical problem solving, because they correlated highly with clinical tests and problem solving questions and moderately with a factual knowledge testing. In our study, EMQs had strong correlations (r= .503 to r=.533) with the practical related OSPE in the multispecialty final integrated examination.

In this study, the MCQs have higher correlations, than EMQs, with the total haematology component and the multi-specialty final integrated examination for all three courses: APS-I, II and III. This may be due to the fact that the MCQs make up a higher percentage of the multi-specialty final integrated examination than the EMQs, or it could be that the EMQs require higher order cognitive skills. From this, the recommendation would be to increase the EMQ content in the examination papers, as suggested also by Wass, MacGibbon and Van der Vleuten (2001). Furthermore, this may suggest that students need more practice in doing the EMQs (which are new to them) throughout the year so that they are as comfortable with them by the end of the course.

The PDQ features an evolving case scenario, and tests the candidates' problem solving and reasoning ability, rather than mere factual recall, which is in keeping with the PBL philosophy. In our study the PDQs had strong correlations (r=.480 to r=.634) with the practical related OSPE, which assess not just factual recall but also clinical and practical application of knowledge. PDQs are easier to set than MCQs (Palmer, Duggan, Devitt & Russell, 2010) however, compared to MCQs, they have issues regarding 'sampling' whereas with MCQs more content can be tested (depending on the way the case scenario evolves). Thus the combination of PDQ, MCQ and EMQ in the department allows us to test the depth and breadth of the curriculum. Epstein (2007) said using multiple methods of assessment can also assess students' learning needs and can help to identify those that need help. Wass, McGibbon and Van der Vleuten al (2001) made a similar point: the use of a variety of assessments improves content validity and reliability which can be further improved by balancing the of number of test items and length of test.

PDQs used as continuous assessment are easier to mark than SAQs and feedback is given to students quickly giving students the opportunity to correct themselves ahead of final examinations. Assessments should motivate students (O'Farrell, n.d.) and guide them and should be useful when it comes to evaluating teaching and learning effectiveness. Like Adeniyi, Ogli, Ojabo & Musa (2013), our department believes in maintaining the CA. This encourages students to work consistently throughout the course and reduces the tendency to memorize information at the last minute. Indeed the CA component of the examination of APS-I, II and III has since increased from 25% to 30%. The CA for haematology in APS-II, APS-III and IPS had the strongest correlations with the total haematology

component. The strong correlations with the final multi-specialty integrated examinations results also suggest that haematology assessments are well aligned with other specialties within the Para-clinical Sciences.

Limitations

1. The number of students in the study is small.
2. Correlations among the other sub-specialties in the Para-clinical Sciences have not been conducted.
3. The effect of combining multiple sub-specialties in one big integrated examination was not analysed.
4. The difficulty level of the examination questions themselves were not analysed in this paper.
5. Generalisability of correlations of assessment modalities: Wass, McGibbon and Van der Vleuten (2001) raised the question of overall reliability in correlating examinations that are so varied in terms of length of examination, content, structure of the examination, duration of examination and the actual weighting in the final combined result.

Conclusion

The new assessment methods are well aligned. The different combinations of different assessment modalities are useful and help the students. It is recommended that examiners increase the EMQ component of the final examinations and that the teachers give students practice throughout the year on assessment modalities such as EMQs, particularly in those areas students are not comfortable with. Assessments that have high inter-rater variability have low correlations. This may be increased if the objectivity of these assessments is increased, as highlighted by Campos-Outcalt et al (1999). Furthermore, the CA is important, and should be maintained, as it encourages continuous learning among students. Our correlational analysis provides the opportunity to evaluate the content composition of the examination. As stated before, so far the weighting of the CA has since been raised from 25 to 30%; therefore, the EMQ component could also be increased in terms of number of items and duration of the examination. For the authors as teachers, the challenge to improve objectivity in all assessments that have high subjectivity and inter-examiner reliability has been emphasized. MCQs and EMQs are highly objective. PDQs may show some subjectivity (although much less than seen with SAQs); however, this can be reduced with the use of very clear answer keys. Assessments do take up a significant amount of time and resources. Currently the PDQ is paper based, however going forward, the authors are investigating the possibility of a computer-based PDQ. This correlational analysis motivated the authors to analyse the EMQ, MCQ and PDQ examination papers and students' performance results. While the authors note the clear benefits of introducing the EMQ and PDQ as assessment methods, as this was a retrospective analysis, no formal study was conducted among the students to document their views on these new assessments.

However informal comments from some of them suggest that they appreciated the clinical application of knowledge required instead of just the basic recall of facts.

References

Adeniyi, O.S., Ogli, S.A., Ojabo, C.O. & Musa, D.I. (2013). The impact of various assessment parameters on medical students' performance in first examination in physiology. *Nigerian Medical Journal, 54*(5), 302-305.

Campos-Outcalt, D., Watkins, A., Fulginiti, J., Kutob, R. & Gordon, P. (1999). Correlations of family medicine clerkship evaluations and objective structured clinical examination scores and residency directors' ratings. *Educational Research and Methods (Family Medicine), 31*(2), 90-94.

Epstein, R.M. (2007). Assessment in medical education. *New England Journal of Medicine*, 356, 387-396.

Fenderson, B.A., Damjanov, I., Robeson, M.R., Veloski, J.J. & Rubin, E. (1997). The virtues of extended matching and uncued tests as alternatives to multiple choice questions. *Human Pathology, 28*, 526–532.

Gupta, E., Sze-Piaw, C., Supamiam, A., Khaw, L.B., Loh, R., Tong, K.S., Keng, C.S., Zaher, Z.M.M., Yek, T.T., Myint, K.Y., Menon, V., Chai, K.K., Hoe, W.M., Krishnan, C., Idris, N., Yin, L.K., Jutti, R.C. & Palayan, K. (2007). How closely do in course assessments correlate with end of semester exams? *International e-Journal of Science, Medicine & Education, Abstracts from the International medical education conference 2007*. Retrieved from: http://web.imu.edu.my/ejournal/approved/Abs_PA01-15.pdf

O'Farrell, C. (n.d.). *Enhancing student learning through assessment: A toolkit approach*. Retrieved from http://www.tcd.ie/teaching-learning/academic-development/assets/pdf/250309_assessment_toolkit.pdf

Palmer, E.J., Duggan, P., Devitt, P.G. & Russell, R. (2010). The modified essay question: Its exit from the exit examination? *Medical Teacher, 32*, 300-307.

Pepple, D.J., Young, L.E. & Carroll, R.G. (2010). A comparison of student performance in multiple-choice and long essay questions in the MBBS stage I physiology examination at The University of the West Indies (Mona campus). *Advances in Physiology Education, 34*, 86-89.

von Bergmann, H,. Dalrymple, K.R,. Wong, S. & Shuler, C.F. (2007). Investigating the relationship between PBL process grades and content acquisition performance in a PBL dental program. *Journal of Dental Education, 71*(9), 1160-1170.

Wass, V., McGibbon, D. & Van der Vleuten, C. (2001). Composite undergraduate clinical examinations: how should the components be combined to maximize reliability? *Medical Education, 35,* 326-330.