

## Mapping undergraduate exit-level assessment in a medical programme: A blueprint for clinical competence?

C P L Tan,<sup>1</sup> MBBS (Lond), MRCP (UK), FRCGP (UK), MPhilHSE; S C van Schalkwyk,<sup>1</sup> PhD; J Bezuidenhout,<sup>1</sup> PhD; F Cilliers,<sup>2</sup> MB ChB, Hons BSc (MedSc) (MedBiochem), MPhil (Higher Education), PhD

<sup>1</sup> Centre for Health Professions Education, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa

<sup>2</sup> Educational Development Unit, Faculty of Health Sciences, University of Cape Town, South Africa

Corresponding author: C P L Tan (tanplchristina@gmail.com)

**Background.** Assessment is an essential component of a medical curriculum. High-stakes exit-level assessment used for licensing and certification purposes needs to be sound. Even though criteria for evaluating assessment practices exist, an analysis of the nature of these practices is first required.

**Objective.** To map current exit-level assessment practices, as described in institutional documentation.

**Methods.** This descriptive interpretive study centred on the document analysis of final-phase study guides of the undergraduate medical programme at Stellenbosch University, Cape Town, South Africa.

**Results.** The key findings were: (i) there is a diversity of methods and approaches to assessment in the final-phase modules; (ii) modules using similar assessment methods applied different credit weightings; (iii) similar assessment methods were described differently across the study guides; and (iv) study guides varied in the amount of information provided about the assessment methods.

**Conclusion.** There is a diverse range of assessment practices at exit level of the MB,ChB programme at Stellenbosch University. This in-depth analysis of assessment methods has highlighted areas where current practice needs to be investigated in greater depth, and where shifts to a more coherent practice should be encouraged. Assessment mapping provides a useful reference for programme co-ordinators and is applicable to other programmes.

*Afr J Health Professions Educ* 2016;8(1):45-49. DOI:10.7196/AJHPE.2016.v8i1.546



Assessment is an essential component of a medical curriculum and is used to measure and manage student progress. Assessment further serves as an indicator of educational efficacy to institutions and teachers.<sup>[1]</sup> Exit-level assessment is also important for reasons of public accountability and in the interest of patient protection.<sup>[1]</sup> Medical schools are increasingly being challenged to provide evidence that the assessments used can discriminate between sufficiently and insufficiently competent students.<sup>[2,3]</sup> Where exit-level assessments are used for licensing and certification purposes, they are regarded as being 'high-stakes' and therefore have significant implications for the student, curriculum, institution and public.<sup>[4,5]</sup>

The assessment of clinical competence is one of the most important tasks facing medical teachers and is used to certify a level of achievement at the end of a programme.<sup>[6,7]</sup> A range of methods are available to assess clinical competence. These include oral examinations, traditional long and short clinical cases, objective structured clinical examinations (OSCEs), standardised patient-based assessments, and workplace-based assessments such as the mini clinical evaluation exercise (Mini-CEX) and direct observation of procedural skills (DOPS).<sup>[4,8,9]</sup>

To make meaningful decisions about competence, the assessment needs to be sound. Various standpoints have been put forward on how this soundness can be realised. For example, a programmatic approach to assessment has been advocated to achieve fitness for purpose with the assessments used.<sup>[10,11]</sup> Norcini *et al.*<sup>[1]</sup> suggest that validity, reproducibility, equivalence, feasibility and acceptability are essential criteria for good or sound assessment. Multiple methods, preferably in a variety of contexts to capture different aspects of performance,<sup>[7]</sup> also need to be considered.

Given the existence of established criteria to guide sound assessment practices, it would seem reasonable to assume that their application in

medical education programmes is a priority for medical schools that hold themselves publicly accountable to ensure that assessments are seen as credible for all stakeholders. However, there appear to be few studies that have looked at exit-level assessment practices against such criteria.<sup>[4,8]</sup>

An analysis of the assessment practices that are in place is a first step before investigating exit-level assessment against established criteria. There appear to be few studies in this area,<sup>[12,13]</sup> this study seeks to address the gap. As a starting point, the investigation concentrated on assessment in the final 18-month phase of the Bachelor of Medicine and Bachelor of Surgery (MB,ChB) programme at Stellenbosch University, Cape Town, South Africa. Currently, no overall map exists of assessments as practised during this period. Creating such a map would help to provide an overall picture of what assessment takes place. A preliminary literature search for 'mapping' revealed that this term is often associated with 'curriculum mapping', 'concept mapping' and 'mind maps', which make use of visual or diagrammatic pictures instead of written or verbal descriptions to illustrate the relationships and connections between different components of a curriculum or concepts.<sup>[14,15]</sup> Applying mapping to assessment practices or activities would appear to be a reasonable step forward.

One way of analysing assessment activities is by focusing on how these are described in official faculty documents and student module study guides. The objective of the study was therefore to map current exit-level assessment practices as described in the documentation relevant to the final phase of a medical programme. The research question was: 'What can be learned about the assessment of clinical competence at exit level of an MB,ChB programme from an analysis of how this is described in student study guides provided for each of the modules in the final phase?'

## Methods

This descriptive interpretive study centred on a process of document analysis of the 2012/2013 study guides. In the final 18-month phase (which runs from July of one year to November of the following year) of the 6-year MB,ChB programme at Stellenbosch University, students rotate through 11 clinical modules varying in length from 3 to 7 weeks. Ten of the modules represent one clinical discipline each, and the remaining module, Health, Disease and Disability in the Community, is shared by the divisions of Family Medicine and Community Medicine, and the Centre for Care and Rehabilitation (Table 1). Detailed information relating to each module, including teaching schedules, duty rosters, projects and assignments, assessment methods and resource materials, is made available in study guides that are provided to all students and relevant faculty. In each module, three components contribute to the students' final overall mark: in-module and end-of-module assessments and final module examinations are conducted in either April or November of the final year.

The analysis of the 11 final-phase module study guides was undertaken in two stages. In the first stage, any available information pertaining to assessment conducted during the module (either in-module or end-of-module) and in the final examinations was gathered from the study guides. This collection included varying combinations of information with regard to the assessment schedules, written descriptions of methods of assessment, assessment checklists and marking grids, logbooks, proportion of

marks allocated for each assessment method and weighting (relating to the calculation of students' final overall mark for that module). As the study guides were written in English and Afrikaans, the information provided in both language versions was compared to check whether it was the same (by investigator JB, who is fluent in both languages). The information was collated on an Excel spreadsheet and categorised by modules and assessment methods to generate an overview of exit-level assessment in the programme and facilitate comparison between the modules.

As this process proceeded, it became clear that there were some gaps in information in the study guides. In the second stage of data collection, all 11 module chairs (faculty who were in charge of organising and co-ordinating the individual modules) were invited to participate in clarificatory interviews to verify and add to the correctness of assessment-related information in the study guides. The module chairs were invited by letter and email, with a follow-up email being sent to non-responders 4 weeks after the initial invitation. At the time of the interview, informed consent was obtained from study participants. An interview schedule was drawn up to serve as a prompt during the interviews. Notes were taken during the interviews, with additional notes recorded afterwards from memory recall. Where necessary, the data on the spreadsheet were amended based on the additional information obtained from these interviews.

Ethical approval was obtained from the Stellenbosch University Health Research Ethics Committee (Ethics Committee Reference No. N13/01/009) and institutional permission from the Stellenbosch University Division of Institutional Research and Planning to conduct this study.

## Results

The information provided in both the English and Afrikaans versions of all 11 final-phase module study guides was confirmed to be the same. Nine of the 11 module chairs consented to participating in interviews, 1 declined and 1 was unavailable. Ultimately, 8 module chairs and 1 module team member were interviewed. Interviews, lasting between 20 and 100 minutes, were conducted over a period of 7 weeks by investigator CPLT.

Twenty-one different assessment methods were identified from the study guides. The results are summarised in Tables 2 and 3 to illustrate the differences in methods used during the modules (in-module and end-of-module) and

in the final module examinations. Assessment methods used were grouped together under three main categories, i.e. (i) written; (ii) performance-based; and (iii) other forms of assessment that did not fall under the previous two categories. In drawing up the groupings, it became evident that there was no uniformity in how assessments were described.

### Written assessments

The most common format of written assessments was multiple-choice questions (MCQs), used by six modules. 'Written' and 'slide' tests used in five modules signified some format of short-answer questions (SAQs), in which students were required to formulate responses to questions posed, based on a clinical scenario, clinical or laboratory investigations, or a photograph. In several instances, where information extracted from the study guides indicated similar terms being used by different modules, interviews revealed that the nature of the assessment was different. As an example, the slide test in Modules 2 and 9 referred to the projection of a PowerPoint presentation of clinical photographs on a screen while students were writing the test, whereas in Module 7, this referred to a written-format assessment which 'includes clinical material as well as special investigations' (Study guide 7), with 'questions based around clinical scenarios' (Module chair D). 'Other written' assessments were used in two modules. These included assignments that students were required to complete during the modules, such as an electronic literature search relating to a patient that the student had cared for during the module, and an evidence-based medicine presentation.

### Performance-based assessments

Performance-based assessment methods included an assessment of clinical skills in a controlled setting in the form of an OSCE and/or objective structured practical examination (OSPE), which was used in four modules. The number of stations was variable. The OSCE and OSPE used in the final summative examination for Module 4 comprised 16 active stations, each of 7 minutes' duration, whereas the OSCE for Module 8 had approximately 20 active stations, each of 5 minutes' duration. 'Unprepared OSCE questions' (Study guide 4) that were used as an in-module assessment method in Module 4 were described by Module chair E to be of a written format and were used 'to test knowledge'. The OSPE in-module assessment for Module 5 was described as including written clinical scenarios, use of videoclips

**Table 1. Modules in the final phase of the MB,ChB programme at Stellenbosch University**

Anaesthesiology
Health, Disease and Disability in the Community
Internal Medicine
Obstetrics and Gynaecology
Ophthalmology
Orthopaedic Surgery
Otorhinolaryngology and Head and Neck Surgery
Paediatrics and Child Health
Psychiatry
Surgery
Urology
Total: 11 modules

**Table 2. Range of assessment methods used during the modules**

Module	Duration (weeks)	Assessment methods (in-module and end-of-module assessments)							Contribution to final module mark, %
		Written			Performance based				
		MCQ	SAQ	Other written	OSCE/OSPE	Clinical cases	Diverse clinical	Other	
1	3		Written test (12.5)			Clinical examination (12.5)	Clinical case discussion (12.5); clinical examination method (12.5)	Continuous (P/F)	50
2	3		Slide test (25)	Clinical case studies (25)			Skills logbook (5); practical ability (40)	Dedication and enthusiasm (5)	50
3	5						General oral and simulated clinical oral (50)		50
4	6	(15)			OSCE (10)			Ward mark (25)	50
5	7	(10)			OSPE (20)		Portfolio (20)	Attitude (satisfactory/unsatisfactory)	50
6	7	(15)				Clinical (17.5)		Continual (17.5)	50
7	6			Electronic literature search (5)		Clinical long case (40)	Clinical procedures (completed: Yes/No); X-ray presentation (5)	Professional conduct (satisfactory/unsatisfactory)	50
8	5	(5)		EBM presentation (5); work rehab task (2.5); physical rehab task (2.5); community project (12.5)			Clinical portfolio (17.5)	Continuous tutor assessment (5)	50
9	3	(25)	Slide test (25)				Skills logbook (P/F)	Dedication and enthusiasm (satisfactory/unsatisfactory)	50
10	3	(17)			Skills (in skills lab) (2)	Clinical (20)		Attitude (1)	40
11	5		Written test (20)			'OSCE' (clinical and oral) (25)		Integrity assessment (5)	50

MCQ = multiple choice question; SAQ = short answer question; OSCE = objective structured clinical examination; OSPE = objective structured practical examination; EBM = evidence-based medicine; P/F = pass/fail. Figures in parentheses refer to the percentage contribution to the final module mark.

and interactive sessions with standardised patients (Module chair F). The inmodule OSCE in Module 11 was actually a combined oral and clinical case assessment.

The use of clinical cases (involving real patients) was employed in five modules, varying from 15 to 30 minutes per case. Module chairs pointed out that the number of cases used in the final examinations varied – from 1 (Module 3) to 2 (Module 7) and 3 (Modules 1 and 6). The number of clinical cases used in the same module also differed when used for in-rotation assessment (e.g. Modules 6 and 7 used 1 case each) compared with the final examinations (the same Modules 6 and 7 used 3 and 2 cases, respectively). In two other modules (Modules 9 and 11), there appeared to be some overlap between the use of clinical cases and oral assessment in the final examinations, as described by the respective module chairs.

A number of 'diverse clinical' assessment methods was described in the study guides, comprising skills logbooks, portfolios, assessment of 'practical ability' (based broadly on history-taking and examination technique, mastery

of skills prescribed in a logbook, ability to formulate and summarise clinical problems and develop a management plan); clinical examination method (based on specific physical examination techniques in that module); clinical case discussions and X-ray presentations to ward consultants; and oral assessment. This loose grouping was made by investigator CPLT in the initial mapping of all assessment methods extracted from the study guides, as these methods shared a common clinical thread but did not fit into the two previously described groups of performance-based assessment methods.

### Other assessments

The remaining category of assessment methods used in 10 of the final-phase modules, primarily as part of in-rotation assessment, is labelled 'other'. These methods dealt mainly with various aspects of professionalism. In four modules, although this assessment did not appear to carry an actual mark, the student was required to obtain a 'satisfactory' judgement.

**Table 3. Range of assessment methods used for the final module examinations**

Module	Duration (weeks)	Assessment methods (final module examinations)					Contribution to final module mark, %
		Written		Performance based			
		MCQ	SAQ	OSCE/OSPE	Clinical cases	Diverse clinical	
1	3		Written examination (12.5)		Clinical examination (37.5)		50
2	3						50
3	5				Clinical (17)	General oral and simulated clinical oral (33)	50
4	6			OSCE and OSPE (50)			50
5	7					Oral (50)	50
6	7	(20)			Clinical (30)		50
7	6		Slide: written (25)		Clinical (25)		50
8	5			OSCE (50)			50
9	3				Clinical oral examination (50)		50
10	3			OSCE (24)		Oral (36)	60
11	5				Clinical, oral and X-ray discussion (50)		50

Figures in parentheses refer to the percentage contribution to the final module mark.

Structured marking guidelines to assist the assessment of this component were provided in the study guides for Modules 7 and 8. For the remaining eight modules, module chairs confirmed that there were no guidelines and the allocation of marks was subjective.

## Summary of results

Ten of the modules used at least one written and one clinical assessment method during the module, whereas Module 3 relied on one method in the form of an oral assessment (Table 2). On overall review of the final examinations (Table 3), Modules 1, 6 and 7 used a written and clinical assessment method. Three modules (3, 10 and 11) used two clinical assessment methods. Two modules (4 and 8) used a multiple station OSCE and/or OSPE format, and two modules (5 and 9) used an oral assessment format alone. Information relating to the final examinations for Modules 6 and 9 was not described in the study guides; this additional information was obtained only at the time of interview. There was no information available regarding the final examination in the Module 2 study guide.

The students' final overall mark for each module was based on two components: the total marks awarded for the rotation (from in-rotation and end-of-rotation assessments) and those from the final examinations. In 10 of 11 modules, the weighting for these two components was equal. In the remaining Module 10, 40% of the final overall mark was derived from the rotation marks and 60% from the final examination marks. As indicated by the figures in parentheses in Tables 2 and 3, the weighting of individual assessment methods varies considerably between modules.

## Discussion

Four key findings emerged from this study. Firstly, there was a diversity of assessment methods and approaches in the final-phase modules. Secondly, modules using similar assessment methods applied different weightings. A third finding was that the information provided about similar assessment

methods was described differently in the various module study guides. These are not necessarily synonymous with what is described in the literature. And fourthly, study guides varied in the amount and detail of information provided about the assessment methods used in the respective modules.

## Range of methods used

The diversity of methods and approaches to assessment across the final-phase modules is similar to that reported in McCrorie and Boursicot's<sup>[12]</sup> UK study and by Ingham<sup>[13]</sup> in Australia. Conversely, a single assessment method was used in several modules. The question is whether the (mix of) methods are utilised in a way that is appropriate to exit-level assessment. Miller's<sup>[16]</sup> 'pyramid', often used to illustrate the multidimensional complexity of assessing clinical competence, moving upwards from reproduction or factual recall in the lower tiers of the 'pyramid' to demonstration and application at the summit, provides a useful framework for responding to this question. The study findings indicate that a substantial proportion of assessment still takes place at the 'lower' tiers of the pyramid. This finding raises questions about how this might influence the validity of decisions on the clinical competence of the student. Analysis of how assessment is described in the student study guides does not provide sufficient information to draw final conclusions, and further research is required in this area. Other questions deserving further study include whether the range of methods used is appropriate to the outcomes of the relevant exit-level modules and what the findings reveal about the validity of the opinions offered by external examiners.

## Weighting of assessment methods

Modules using similar assessment methods applied different weightings, suggesting that the emphasis placed on the assessment method varied across modules. Possible explanations include resource constraints

(such as available assessors and space to conduct assessment), and the opinion of assessment conveners about the perceived merits of the chosen methods. Wass *et al.*<sup>[8]</sup> have shown that weighting accorded to items per test or total test time can significantly affect reliability, but this has to be considered carefully with other established criteria for good or sound assessment in a high-stakes context. The reasons behind these decisions were beyond the ambit of this study, and these too warrant further investigation.

## Description of assessment methods

The study guides serve primarily as a reference for students and faculty to provide official information relating to each module. There was little uniformity in how assessments were described. The varying use of terms, such as OSCE and OSPE, suggests that faculty in different modules may have a different understanding of similar assessment methods, which could impact on reliability and fairness. The absence of clear descriptions of what individual assessment methods entail could potentially lead to confusion and incorrect assumptions by students. Defining and providing consistent and adequate information in the module study guides and official faculty documents regarding the assessment methods used would reduce any possible misunderstanding. Incorporating this detail into faculty development programmes would also promote consistency in the future practices of assessors.

Variable in-rotation assessment practice without any descriptions or guidelines of how the marks are determined was noted in several modules, which could result in subjective interpretation and impact on fairness. These in-rotation assessments dealt mainly with aspects of professionalism. The assessment of professionalism is equally complex and requires a multidimensional approach. While itemised checklists and rating scales may not necessarily be the best solution, the introduction of some form of global overall rating could be considered as an alternative and go some way to addressing the difficulties of assessing aspects of behaviour or professionalism during placements.<sup>[7]</sup> Ultimately, whether quantitative and qualitative measures are used, their utilisation in a defensible manner is key to making valid inferences.

## Level of detail provided

Study guides varied in the number of assessment methods used in the respective modules and amount of detailed information provided. There were instances where there was no information regarding the final examination or the assessment methods used. Study guides have the potential to help students to manage their own learning. One of their many uses as a management tool could be for examination preparation by providing information on the format and arrangements for assessment.<sup>[17]</sup> Although the broad outlines in the Stellenbosch University study guides are similar, a structured template could be used to provide guidance for uniformity in the writing detail.

## The way forward

Overall or central co-ordination of the assessment of the MB,ChB programme could address some of the issues highlighted, such as the uniformity of detail and consistency of description regarding assessment methods in all the final-phase modules. One next step could be to investigate what exit-level assessment actually takes place at Stellenbosch University, and how this relates to what is described in the final-phase module study guides. This could be further expanded to study the exit-level assessment taking place at other medical schools in a similar context, such as in sub-Saharan Africa. Exploring the reasons around choices of assessment methods, decisions on weighting, and clinical competencies considered appropriate for medical graduates could also be avenues for further research.

## Conclusion

This study provides an in-depth analysis of assessment methods across an undergraduate medical programme, highlighting the range and diversity of existing assessment practices at the exit-level phase of the MB,ChB programme at Stellenbosch University. A limitation of the research is that the findings reported are not necessarily generalisable to earlier phases of the MB,ChB programme at the university. In addition, actual assessment practices and content will require separate verification.

This study has highlighted potential areas where current practice needs to be investigated in greater depth, and where a shift to a more coherent practice should be encouraged. Assessment mapping provides a useful reference for programme co-ordinators and the tool has applicability for other programmes.

## References

- Norcini J, Anderson B, Bollela V, et al. Criteria for good assessment: Consensus statement and recommendations from the Ottawa 2010 Conference. *Med Teach* 2011;33(3):206-214. [<http://dx.doi.org/10.3109/0142159X.2011.551559>]
- Health Professions Council of South Africa. Health Professions Act 56 of 1974. Regulations relating to the registration of students, undergraduate curricula and professional examinations in medicine. *Government Gazette* 31886, 19 February 2009.
- General Medical Council. Assessment in Undergraduate Medical Education. Advice Supplementary to Tomorrow's Doctors. London: General Medical Council, 2011. [[http://www.gmc-uk.org/static/documents/content/Assessment\\_in\\_undergraduate-web.pdf](http://www.gmc-uk.org/static/documents/content/Assessment_in_undergraduate-web.pdf) (accessed 20 February 2013)].
- Roberts C, Newble D, Jolly B, Reed M, Hampton K. Assuring the quality of high stakes undergraduate assessments of clinical competence. *Med Teach* 2006;28(6):535-543. [<http://dx.doi.org/10.1080/01421590600711187>]
- Norcini JJ, Lipner RS, Grosso LJ. Assessment in the context of licensure and certification. *Teach Learn Med* 2013;25(S1):S62-S67. [<http://dx.doi.org/10.1080/10401334.2013.842909>]
- Wass V, van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. *Lancet* 2001;357:945-949.
- Epstein RM. Assessment in medical education. *N Engl J Med* 2007;356:387-396.
- Wass V, McGibbon D, van der Vleuten C. Composite undergraduate clinical examinations: How should the components be combined to maximise reliability? *Med Educ* 2001;35(4):326-330.
- Wilkinson TJ, Frampton CM. Comprehensive undergraduate medical assessments improve prediction of clinical performance. *Med Educ* 2004;38(10):1111-1116. [<http://dx.doi.org/10.1111/j.1365-2929.2004.01962.x>]
- Dijkstra J, van der Vleuten CPM, Schuwirth LWT. A new framework for designing programmes of assessment. *Adv Health Sci Educ Theory Pract* 2010;15(3):379-393. [<http://dx.doi.org/10.1007/s10459-009-9205-z>]
- Van der Vleuten CPM, Schuwirth LWT, Driessen EW, et al. A model for programmatic assessment fit for purpose. *Med Teach* 2012;34(3):205-214. [<http://dx.doi.org/10.3109/0142159X.2012.652239>]
- McCrorie P, Boursicot KAM. Variations in medical school graduating examinations in the United Kingdom: Are clinical competence standards comparable? *Med Teach* 2009;31(3):223-229. [<http://dx.doi.org/10.1080/01421590802574581>]
- Ingham AI. The great wall of medical school: A comparison of barrier examinations across Australian medical schools. *Australian Medical Student Journal* 2011;2(2):5-8.
- Harden RM. AMEE Guide no. 21. Curriculum mapping: A tool for transparent and authentic teaching and learning. *Med Teach* 2001;23(2):123-137. [<http://dx.doi.org/10.1080/01421590120036547>]
- Daley BJ, Torre DM. Concept maps in medical education: An analytical literature review. *Med Educ* 2010;44(5):440-448. [<http://dx.doi.org/10.1111/j.1365-2923.2010.03628.x>]
- Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;65(9 Suppl):S63-S67.
- Harden RM, Laidlaw JM, Hesketh EA. AMEE Medical Education Guide No. 16: Study guides – their use and preparation. *Med Teach* 1999;21(3):248-265. [<http://dx.doi.org/10.1080/01421599979491>]