BEST PRACTICES FOR QUALITY ASSESSMENT IN THE CLINICAL PHASE OF UNDERGRADUATE MEDICAL TRAINING

by

Hanneke Brits

Thesis submitted in fulfilment of the requirements for the degree
Philosophiae Doctor in Health Professions Education
PhD. HPE

in the

Division Health Sciences Education
Faculty of Health Sciences
University of the Free State
Bloemfontein

August 2020

Promoter: Dr J Bezuidenhout
Co-Promoter: Dr LJ van der Merwe

DECLARATION

I hereby declare that the compilation of this thesis is the result of my own, independent investigation. I have endeavoured to use the research sources cited in the text in a responsible way and to give credit to the authors and compilers of the references for the information provided, as necessary. I have also acknowledged those persons who have assisted me in this endeavour. I further declare that this work is being submitted for the first time at this university and faculty for the purpose of obtaining the degree Philosophiae Doctor in Health Professions Education and that it has not previously been submitted to any other university or faculty for the purpose of obtaining a degree. I also declare that all information provided by study participants will be treated with the necessary confidentiality.

24 August 2020

Prof. Hanneke Brits **Date**

I hereby cede copyright of this product in favour of the University of the Free State.

24 August 2020

Prof. Hanneke Brits **Date**

ACKNOWLEDGEMENTS

I wish to convey my sincere thanks and appreciation to the following persons who assisted me to complete this study:

- My promoter, Dr Johan Bezuidenhout, Head of Department at the Division of Health Sciences Education, Faculty of Health Sciences, University of the Free State, for his guidance, support and expertise. Also for his willingness to allow me to do things my way, most of the time.
- My co-promoter, Dr Lynette van der Merwe, for her optimism, encouragement, unfailing patience, guidance, support and attention to the quality of the study.
- Prof. Gina Joubert of the Department of Biostatistics, UFS, for her support, effort and guidance to ensure the correct statistics and interpretation of results.
- Prof. Mathys Labuschagne, for conducting the focus group interview.
- Mr Johan Botes, for assisting with the formatting, finding articles and always being available to give practical advice and support.
- To my language editor, Ms Hettie Human, for her expertise and efficiency in editing my thesis and giving feedback with a short turnaround time.
- To the focus group participants, for their valuable input and time. It is highly appreciated.

TABLE OF CONTENTS

DECLARATION ii

ACKNOWLEDGEMENTS iii

TABLE OF CONTENTS iv

LIST OF ABBREVIATIONS AND ACRONYMS vii

LIST OF TABLES viii

LIST OF FIGURES ix

CONCEPT CLARIFICATION x

CHAPTER 1: ORIENTATION TO THE STUDY 1

1.1 INTRODUCTION 1

1.2 BACKGROUND TO THE RESEARCH PROBLEM 5

1.2.1 Assessment in the clinical phase of the MBChB programme at the UFS 5

1.2.2 Introduction to assessment 6

1.2.3 Components of assessment 7

1.2.4 Changes in assessment 7

1.2.5 Quality assessment 8

1.3 PROBLEM STATEMENT 9

1.4 THE RATIONALE AND OVERALL GOAL OF THE STUDY 10

1.5 AIM OF THE STUDY 10

1.6 RESEARCH QUESTIONS 10

1.7 OBJECTIVES OF THE STUDY 10

1.8 RESEARCH DESIGN 11

1.8.1 Research paradigm 11

1.8.2 Theoretical perspective 13

1.8.3 Research methodology 14

1.9 RESEARCH METHODS 15

1.9.1 Part 1 – Rapid review 15

1.9.2 Part 2 – Review of current assessment practices 15

1.9.2.1 Review of student opinions 16

1.9.2.2 Evaluation of assessment practices 16

1.9.2.3 Predictability and evaluation of students' assessment marks 16

1.9.3 Part 3 – Focus group interview 16
1.9.4 Contribution – Best-practice recommendations 17
1.10 TRUSTWORTHINESS 17
1.10.1 Credibility 17
1.10.2 Transferability 18
1.10.3 Dependability 18
1.10.4 Confirmability 18
1.10.5 Internal and external validity 18
1.10.6 Reliability 19
1.10.7 Objectivity 19
1.11 ETHICAL CONSIDERATIONS FOR THE STUDY 19
1.12 DEMARCATION AND SCOPE OF THE STUDY 20
1.13 VALUE, SIGNIFICANCE AND CONTRIBUTION OF THE STUDY 20
1.13.1 Value 20
1.13.2 Significance 20
1.13.3 Contribution 20
1.14 SCHEMATIC OVERVIEW OF THE STUDY 21
1.15 SCHEMATIC CHAPTER LAYOUT OF THE STUDY 22
1.16 CHAPTER SUMMARY 23

CHAPTER 2: ARTICLE 1 – A FRAMEWORK TO BENCHMARK THE QUALITY OF CLINICAL ASSESSMENT IN A SOUTH AFRICAN UNDERGRADUATE MEDICAL PROGRAMME 24

CHAPTER 3: ARTICLE 2 – STUDENTS' VOICES: ASSESSMENT IN UNDERGRADUATE CLINICAL MEDICINE 36

CHAPTER 4: ARTICLE 3 – ASSESSMENT PRACTICES IN UNDERGRADUATE CLINICAL MEDICINE TRAINING: WHAT DO WE DO AND HOW WE CAN IMPROVE? 48

CHAPTER 5: ARTICLE 4 – EVALUATION OF ASSESSMENTS MARKS OF AN UNDERGRADUATE MEDICAL TRAINING PROGRAMME: WHERE ARE WE AND HOW CAN WE IMPROVE? 59

CHAPTER 6: ARTICLE 5 – QUALITY ASSESSMENT IN UNDERGRADUATE MEDICAL TRAINING: HOW TO BRIDGE THE GAP BETWEEN WHAT WE DO AND WHAT WE SHOULD DO 72

CHAPTER 7: CONTRIBUTION OF THE STUDY 86

CHAPTER 8: CONCLUSIONS, IMPLICATIONS AND RECOMMENDATIONS 110

8.1 INTRODUCTION 110

8.2 OVERVIEW OF THE STUDY 110

8.2.1 Research question 1: What are the current policies and guidelines to benchmark quality assessment in the clinical phase of the MBChB programme? 112

8.2.2 Research question 2: What are the current assessment practices employed in the MBChB programme at the UFS and how valid and reliable are these assessments? 112

8.2.3 Research question 3: What best practices can be implemented to enhance quality assessment in the clinical phase of the MBChB programme? 114

8.3 IMPLICATIONS OF THE STUDY 116

8.3.1 Goal and contribution 116

8.3.2 Final conclusion 117

8.3.3 Study limitations and critique 117

8.3.4 Recommendations 118

8.3.5 Trustworthiness of the research 118

8.3.6 Personal experience 119

8.4 FINAL REMARK 119

REFERENCES 120

APPENDICES 128

Appendix A: Ethics approval letter 129

Appendix B: Gatekeepers' approval 130

Appendix C: Language editor declaration 131

Appendix D: Turnitin report 132

LIST OF ABBREVIATIONS AND ACRONYMS

AMEE	–	Association of Medical Education in Europe
HPA	–	Health Professions Act
HPCSA	–	Health Professions Council of South Africa
MBChB	–	Bachelor of Medicine and Bachelor of Surgery
MCQ	–	Multiple-choice questions
OSCE	–	Objective structured clinical examinations
PhD	–	Philosophiae Doctor
QUAL	–	Qualitative
QUAN	–	Quantitative
SAQA	–	South African Qualifications Authority
T&L	–	Teaching and learning
UFS	–	University of the Free State
WBA	–	Workplace-based assessment

LIST OF TABLES

Table 1.1: Research design that was followed in this thesis 12

LIST OF FIGURES

Figure 1.1: A schematic summary of the research process (Compiled by the researcher Brits 2020) 4
Figure 1.2: Schematic flow of the abductive approach that was applied (Adapted from the work by Kovács and Spens by the researcher Brits 2020) 13
Figure 1.3: Schematic display of the triangulated mixed methods research model used for this study (Compiled by the researcher Brits 2020) 14
Figure 1.4: A schematic overview of the study (Compiled by the researcher Brits 2020) . 21
Figure 1.5: A schematic chapter layout of the study (Compiled by the researcher Brits 2020) 22
Figure 8.1: A brief overview of the study (Compiled by researcher Brits 2020) 116
Figure 8.1: A brief overview of the study (Compiled by researcher Brits 2020) 116

CONCEPT CLARIFICATION

Assessment is the process whereby people (learners) are tested on specific knowledge or skills (Leinster 2013:9).

Best practice is defined as, "a procedure that has shown by research and experience to produce optimal results and that is established or proposed as a standard suitable for widespread adaption" (Merriam-Webster 2019:online).

Competency is defined as the ability to perform a task at a set standard and includes attributes such as knowledge, skills and attitudes (Moynihan et al. 2015:online).

Critical review refers to the careful judgement or evaluation of literature or processes with the intention of making decisions or changes, if necessary (Merriam-Webster 2019:online).

Evaluation refers to the systematic collection, interpretation and analysis of processes or action to determine the current condition (Leinster 2013:21; Merriam-Webster 2019:online).

Thesis by publication is defined as follows by the University of the Free State in its Policy Document for Masters and Doctoral studies:

> A doctoral thesis is the sole research component of a doctorate. It must demonstrate that the candidate has made a specific contribution to the enhancement of knowledge in the chosen field, while providing evidence of independent critical ability. A *PhD* by publication requires three publishable articles *(*UFS 2016:3).

SUMMARY

Medical universities have a responsibility to ensure quality assessment of clinical competence when they certify that they produce competent medical practitioners who can integrate knowledge, skills and attitudes. The assessment of clinical competence is complex, and can be characterised by tension between validity, reliability and fairness, due to the assessment on the "does" level.

The defined problem that was addressed is that assessment in the clinical phase of the undergraduate medical programme (MBChB) at the University of the Free State has not been reviewed critically or benchmarked against local and international standards. This thesis intended to benchmark clinical assessment practices against an assessment framework and then propose an action plan on how to bridge the gap between theory and practice when assessing clinical competence.

A pragmatic approach was followed to address the practical problems of uncertainty regarding the quality of assessment. From a theoretical perspective, an abductive approach was used to achieve inference. An explanatory sequential mixed method research design was used. During triangulation, alignment of and gaps between theory and practice were identified and solutions recommended. A proposal with an action plan was drafted to enhance the quality of clinical assessment in the undergraduate medical programme.

Firstly, an assessment framework to benchmark clinical assessment in undergraduate medical training was compiled. A rapid literature review of local, national and international official regulations and policies, supported by best evidence practices, were used to compile this assessment framework. In this framework, the three components of quality assessment, namely, accreditation, assessment and quality assurance, were addressed.

In the second part of the study, current assessment practices were reviewed through data collected from three sources, namely, students, lecturers and student marks, to ensure that different aspects were included in the review. A questionnaire with open and closed-ended questions was completed by clinical students in the undergraduate medical programme, to get the students' perspectives on assessment. More than half the students were of the opinion that current assessments were not fair, and >90% complained about the lack of formal feedback after assessments. Secondly, the teaching and learning coordinators and

module leaders of all the clinical departments involved in undergraduate medical training completed questionnaires on the assessment methods used in their departments. They also made recommendations for ways to improve current assessment practices. Using multiple choice questions and objective structured clinical evaluations were standard practice in most disciplines. Workplace-based assessment (WBA) was not well established and was only used in 30.1% of disciplines. The overemphasis on summative assessment was identified as an area for improvement. Thirdly, current assessment practices were evaluated for reliability. The decision reliability between end-of-block assessment and summative assessment was excellent, with a G-index of agreement of between 0.86 and 0.98. Using unobserved long cases during summative assessment was shown to be unreliable and questionable. During a formal focus group interview, answers were sought on how to bridge the gap between theoretical principles of quality assessment and current assessment practices.

Finally, the researcher compiled a proposal with an action plan on how to enhance quality assessment in the clinical phase of the undergraduate medical programme. Most of the practices that compromise the quality of assessment can be addressed on an operational level, and will not be costly to implement. This includes training of assessors, implementation of WBA, effective feedback to students and blueprinting and moderating all assessments. Assessor training will improve the quality of assessments, and will also contribute to the professional development of assessors. Continuous WBA will have the ultimate effect of improving validity and reliability, which will benefit all stakeholders.

Key words: assessment, assessment framework, assessment practices, assessors, blueprinting, clinical competence, continuous assessment, moderation, summative assessment, workplace-based assessment

BEST PRACTICES FOR QUALITY ASSESSMENT IN THE CLINICAL PHASE OF UNDERGRADUATE MEDICAL TRAINING

CHAPTER 1:
ORIENTATION TO THE STUDY

1.1 INTRODUCTION

The researcher approached this thesis by submitting articles for publication, to recommend best practices for enhancing quality assessment in the clinical phase of undergraduate medical training. A pragmatic paradigm was used as a point of departure to evaluate current assessment practices and make best practice recommendations (cf. 1.8.1). Miller (1990:S67) emphasises the importance of quality and responsibility relating to assessment:

> If we are to be faithful to the charge placed upon us by society to certify the adequacy of clinical performance ... then we can no longer evade the responsibility for finding a method that will allow us to do so.

Although the duration and training involved in undergraduate medical programmes are different at different universities around the world, passing the final assessment of the undergraduate medical programme at any institution enables a student to graduate and qualify as a medical doctor.

In South Africa, undergraduate medical programmes are presented at nine accredited universities (HPCSA 2019: online). The programme presented is the Bachelor of Medicine and Bachelor of Surgery (MBChB). For this study, the term MBChB is used to refer to the undergraduate medical programme. The duration of training is between five and six years and it is presented in three phases, namely, orientation, pre-clinical and clinical training, as explained by the Health Professions Council of South Africa (HPCSA 2014:online) and faculty yearbooks and brochures of the various institutions. The MBChB programme at the University of the Free State (UFS) is a five-year (10 semester) outcomes-based education programme, divided into three phases. In an outcomes-based curriculum, the outcomes should be stated clearly in relation to the requirements of regulatory bodies. Assessment should be aligned with these outcomes, according to set criteria, and effective quality assurance measures should be in place (Spady 1994:6–7).

At the UFS, Phase I is covered in the first six months of training and provides an introduction to the foundational aspects of training required by future medical professionals. Phase II takes place over 24 months and concludes the pre-clinical phase. Phase II includes system-based and clinical skills training modules. The clinical phase (Phase III) takes place over the last two and a half years of study (semesters 6 to 10), and entails rotations through various clinical departments at multiple health facilities. Clinical training takes place according to a fixed programme and for fixed periods. Formative assessment takes place throughout the rotation and includes an end-of-block assessment. Together, these assessments contribute to a block mark (module mark), which gives the student access to the final, summative, end-of-year assessment.

The purpose of this study is to review current assessment practices of the clinical phase of the MBChB programme at the UFS critically and to recommend best practices to enhance the quality of these assessments. These recommendations should help to ensure that the assessment principles and practices used in the clinical phase of the MBChB programme are aligned with best practices, and implemented according to the assessment regulations and policies of various regulatory bodies. The UFS MBChB programme is accredited and/or registered with the following bodies: UFS (BC834100), the Council on Higher Education, the South African Qualifications Authority (SAQA 9633) and the HPCSA (HPCSA 8371), and is presented at UFS and Free State Department of Health facilities.

In Part 1 of this study, a rapid review of assessment policies, guidelines and principles that are applicable to the clinical phase of the MBChB programme, and stipulated by the different accrediting bodies, was performed. This served as a theoretical guide to benchmark assessment. In Part 2, current clinical assessment and moderation practices were reviewed critically to determine components of validity and reliability. This was done by gathering information via questionnaires from students and lecturers in the clinical phase of the MBChB programme at the UFS. Student marks (assessment results) were also analysed to evaluate pass/fail decisions, the test-retest reliability, predictability and stability of assessments. During Part 3, a focus group interview was held with the clinical teaching and learning (T&L) coordinators in the clinical phase of the MBChB programme. The purpose of the focus group interview was to align and/or maintain the theoretical principles with current assessment practices, to enhance best practice and quality assessment. Finally, information gathered was used to identify and recommend the best assessment practices, to inform assessment practices and to enhance the quality of assessment in the clinical

phase of the MBChB programme. These practices included audit measures to enable regular review of assessment practices.

Figure 1.1 provides a summary of the research process.

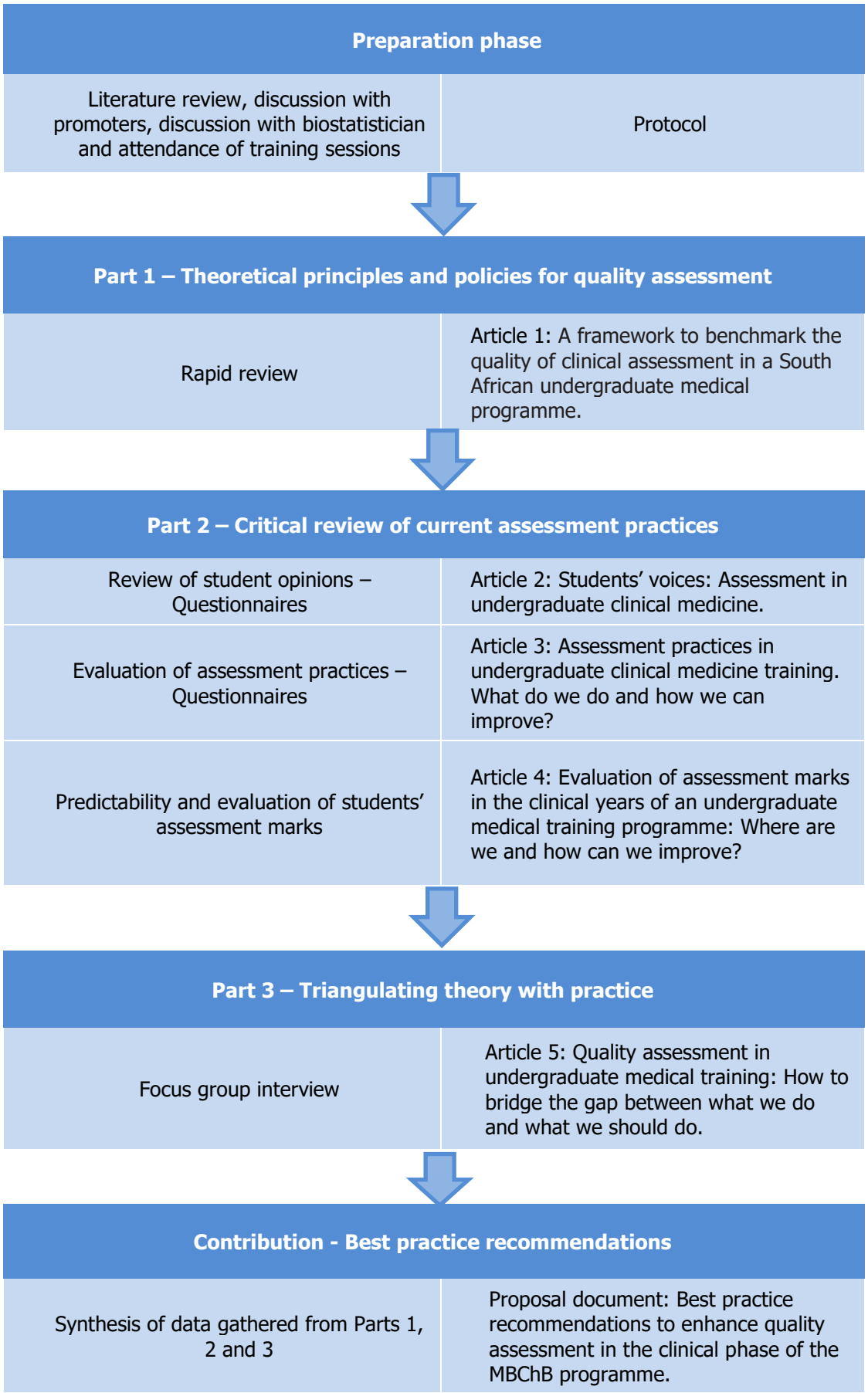


Figure 1.1: A schematic summary of the research process (Compiled by the researcher Brits 2020)

1.2 BACKGROUND TO THE RESEARCH PROBLEM

This study was motivated by the perception that not all clinical assessments in the MBChB programmes are of a high quality. A number of clinical students who participated in the 2016, 2017 and 2018 end-of-year assessments were not satisfied with their marks in/outcome of the summative assessment. Personal communication with students and confidential examination reports indicate that some of them only passed after reassessment, while others, despite good block marks, failed, and had to repeat the year or block. The phenomenon of student dissatisfaction with their marks, e.g., complaints that the questions do not cover all the work, were too difficult, or that the content had not been taught, is not unique to the UFS, and is well described in the literature (Patil *et al*. 2015:s76). However, self-assessment by students is usually accurate, and a study performed at Utah University found that medical students could accurately assess their knowledge, skills and behaviour on assessment, with less than 10% of students overrating their knowledge (Han & Riazi 2018: 386; Panadero, Brown & Strijbos 2016: 811; Torres & Cochran 2016:1154). It is, therefore, important to investigate the opinions and perceptions of students regarding the quality of assessment.

In turn, some assessors at the UFS were of the opinion that certain students were fortunate or "lucky" to have passed the final assessment, and that they may not be competent to qualify as doctors. The phenomenon of "failure to fail" non-competent students is referenced in various caring professions (Gainsbury 2010:2; Hughes, Mitchell & Johnston 2016:54; Shapton 2006:40). Studies report that between 18% and 46% of nurse assessors agreed that some students pass despite unsatisfactory clinical performance (Brown *et al.* 2012:16; Watson 1999:218).

1.2.1 Assessment in the clinical phase of the MBChB programme at the UFS

At the end of the fourth and fifth (final) years of study in the UFS MBChB programme, students who comply with the regulations regarding examination admission (including attendance regulations and a block mark of above 45%) may participate in the end-of-year summative assessment. To progress to the following year of study (fourth year), or to graduate from the programme (final year), a student must pass all the modules, including clinical and theory components, individually. In some instances, students who fail the assessment may qualify for reassessment, as stipulated in the Faculty of Health Sciences Rule Book (UFS 2019:14–15). Students are assessed in the following 12 clinical disciplines:

Anaesthesiology (5th year), Family Medicine (5th year), Internal Medicine (4th and 5th years), Obstetrics and Gynaecology (4th and 5th year), Oncology (4th year), Ophthalmology (4th year), Orthopaedics (4th and 5th years), Otorhinolaryngology (4th and 5th years), Paediatrics and Child Health (4th and 5th years), Psychiatry (4th and 5th years), Surgery (4th and 5th years), and Urology (5th year).

At the time of the study, the researcher was the chairperson of the Phase III working group that oversees the training and assessment of clinical students in the MBChB programme. All the clinical T&L coordinators and module leaders in Phase III are members of this working group. The role of the Phase III chair includes regular engagement with clinical departments and individual T&L coordinators and module leaders regarding teaching and learning activities, assessments and student outcomes. As part of quality assurance in the MBChB programme, and to meet accreditation requirements, all teaching and learning activities and assessment strategies must be reviewed and improved regularly.

1.2.2 Introduction to assessment

To assess the quality of an assessment, it must be evaluated using appropriate criteria or tools (Whiting, Wolff & Mallett 2017:online). A literature search was done for policies and criteria on quality assessment in the clinical phase of an undergraduate medical programme. The terms undergraduate, medical, clinical assessment, assessment policy and quality assurance were used for searches, and the following databases were consulted: National Electronic Theses and Dissertation (ETD) Portal, UFS website, Scopus, NEXUS, World Federation on Medical Education (WFME) and Association of Medical Education in Europe (AMEE).

Although the UFS's assessment policy (UFS 2017:6) sets minimum criteria for the assessment of undergraduate students, and the HPCSA prescribes core competencies for undergraduate medical students (HPCSA 2014:1), strategies to ensure and assess the implementation of these policies are lacking. No UFS-based or South African study could be found on the quality of assessment of medical students in the clinical phase of training. Internationally, AMEE guidelines provide frameworks for clinical assessment (Pangaro & Ten Cate 2013:e1201). A United Nations Educational, Scientific and Cultural Organization (UNESCO) document addressing the United Nations' Sustainable Millennium Goals provides a 14-step cycle for quality assessment (UNESCO 2015:9), though it relates more to the evaluation of national education programmes. These guides provide valuable information

and principles, but are not directly transferable to the undergraduate South African and UFS medical context. These guidelines focus on general programme or medical assessment practices, rather than on assessment in the clinical phase of an undergraduate medical programme.

1.2.3 Components of assessment

Three components of quality in assessment are described in the literature, namely, accreditation, assessment and audit (Kis 2005:5; Van der Bank & Van der Bank 2014:398). The five-year MBChB programme at the UFS is fully accredited by relevant authorities; therefore, the focus of this study was not on accreditation. Instead, the quality of assessment was evaluated by investigating the possible gap between current assessment practices and regulations prescribed by the accreditation authorities. As part of the audit component, current moderation principles and practices were evaluated and adaptations suggested, to enhance quality assessment.

1.2.4 Changes in assessment

Assessment of clinical training in higher education continues to change as new assessment methods are introduced, for example, using simulated patients and simulation units, continuous assessment, as well as the innovative assessment of new competencies and attitudes. At the same time, traditional forms of assessment, for example, unobserved long cases, have become less popular (Wass & Van der Vleuten 2004:1178). Miller (1990:S63) states that "no single assessment method can provide all the data required for judgement of anything so complex as the delivery of professional services by a successful physician".

Assessment should be an integral part of curriculum design, and not an afterthought when marks must be submitted. Biggs (1996:147) describes the term constructive alignment as outcomes, teaching and training activities and assessment being planned to complement and support each other. In an outcomes-based curriculum, "everything" should be built on clearly defined exit outcomes – assessment should be aligned with outcomes and assessment should be criterion-referenced (Spady 1994:6). Achieving this alignment requires that the assessment of clinical training must be varied and updated and refined constantly.

1.2.5 Quality assessment

Originally, the quality of assessment was measured against validity, reliability and fairness (Van der Vleuten 1996:49–51). Lately, the following important components or sub-components to ensure quality assessment were added to the cornerstones of validity, reliability and fairness, namely, feasibility, educational effect, acceptability, assessment of higher cognitive skills and benchmarking (Cruess, Cruess & Steinert 2016:182; Darling-Hammond *et al*. 2013:1; Norcini & Friedman Ben David 2013:288, Norcini *et al*. 2011:209; Van der Vleuten & Schuwirth 2005:310–314).

Assessment during the clinical phase of the MBChB programme at the UFS comprises written papers, which may include multiple-choice questions (MCQs) and short questions, clinical case presentations, objective structured clinical examinations (OSCEs), objective structured practical examinations (OSPEs), evaluation of logbooks, and assignments. Each clinical department responsible for modules in the clinical phase of training of the MBChB programme has a unique assessment strategy. These different forms of assessment may increase the validity of the assessment, as it is possible to assess theory as well as practice, and to assess skills and attitudes in real-life (authentic) situations (Epstein 2007:388; Van der Vleuten & Schuwirth 2005:312).

Poor quality assessment presents many problems. Boud (1998:1) states that, "students can, with difficulty, escape from the effects of poor teaching, they cannot escape the effects of poor assessment". Clinical competence is specific, and not generic, and, therefore, competence or non-competence in one skill does not imply competence or non-competence in related skills. The number of assessments used is very important, especially in high-stakes assessments. Researchers warn against using one assessment when pass/fail decisions have serious implications, such as for registration or licencing (Van der Vleuten *et al*. 2017:611). With one chance of clinical assessment in the final examination, a clinically competent student could fail that component on the day and will need to repeat the assessment and/or undergo further training, according to Rule M3.8 of the UFS School of Medicine Undergraduate Rule Book 2018 (UFS 2018:14). Failure or repeating a module or year has financial and emotional implications for a student; trust may be lost in the institution and the throughput rate and reputation of the faculty or university may be affected (Najimi *et al.* 2013:1). On the other hand, if an incompetent student passes, patients and the healthcare system may be affected negatively, causing loss of life and

avoidable expenses, and it could even lead to misconduct claims (Bates & Slight 2014:1027; Swaminath & Raguram 2010:111).

Certain competencies are difficult to assess during summative assessment. An example is the competencies in the original CANMEDS document (Seely *et al.* 1996:1) that were adapted to AFRIMEDS and included in the HPCSA document on required competencies for undergraduate students. The competencies for a healthcare practitioner include being a professional, communicator, collaborator, leader and manager, health advocate and scholar (HPCSA 2014:1). These competencies could be assessed well as part of continuous assessment and feedback.

Furthermore, assessment results do not guarantee competence in the workplace. In health professions education, Chen, Van der Broek and Ten Cate (2015:433) propose including the term entrustable professional activities in undergraduate medical education, as a measure to ensure workplace competence. Entrustable professional activities can be described as professional competencies, rather than isolated skills, that a trainee should be trusted to perform without supervision due to the trainee having demonstrated competence in that activity (Ten Cate 2013:157). Entrustable professional activities are assessed using a statement by the supervisor that the trainee had been supervised and that it can be predicted that the trainee will perform the entrusted activity adequately without supervision (Chen *et al*. 2015:434).

The purpose of quality assurance and audit is to maintain and enhance agreed upon or set standards and to be confident that the requirements that are set, are met and reported on (Manghani 2011:34). Moderation is an integral part of the current MBChB programme, and is used as the audit tool for ongoing evaluation of the quality of assessment in the clinical phase of the MBChB programme at the UFS. Maxwell (2010:457) describes moderation of assessment as the process of ensuring consistency and fairness of professional judgements relating to student performance.

1.3 PROBLEM STATEMENT

It is evident from the literature review that there is a paucity of literature on clinical assessment practices for undergraduate medical students, particularly in the South African context, where no studies have been published. Medical programmes are accredited and registered with various regulatory bodies, which set specific requirements for education

programmes, as well as assessment. The problem that was addressed by this study is that, despite guidelines and policies, assessment practices in the clinical phase of the MBChB programme at the UFS have not been reviewed critically or benchmarked against local and international policies and guidelines. This lack of benchmarking may pose problems for future accreditation, as well as for quality assurance.

1.4 THE RATIONALE AND OVERALL GOAL OF THE STUDY

The overall goal of the study was to contribute to high quality assessment practices in the clinical phase of the MBChB programme by recommending assessment practices of which the outcome can be confidently defended and justified. This study could lead to other institutions adopting similar assessment practices; in addition, it will be possible to compare assessment practices locally and internationally.

1.5 AIM OF THE STUDY

The aim of this study was to critically review current assessment practices to propose best-practice recommendations to enhance the quality of assessment in the clinical phase of the MBChB programme.

1.6 RESEARCH QUESTIONS

To achieve the goal of the study, three research questions were formulated:

1. What are the current policies and guidelines to benchmark quality assessment in the clinical phase of the MBChB programme?
2. What are the current assessment practices employed in the MBChB programme at the UFS and how valid and reliable are these assessments?
3. What best practices can be implemented to enhance quality assessment in the clinical phase of the MBChB programme?

1.7 OBJECTIVES OF THE STUDY

Specific objectives were set to answer the research questions. These objectives were:

i) To compile a theoretical guide to benchmark current assessment practices based on published guidelines, policies and principles of assessment in the clinical phase of the MBChB programme;

ii) To obtain opinions from the clinical-phase students regarding current assessment practices in the clinical phase of the MBChB programme at the UFS;

iii) To obtain accurate information regarding current assessment and moderation practices from T&L coordinators and/or module leaders in the clinical phase of the MBChB programme at the UFS;

iv) To evaluate if current module assessment marks are reliable predictors of final assessment marks in the clinical phase of the MBChB programme at the UFS;

v) To get a better understanding of and clarity about the current assessment practices from the T&L coordinators in the clinical phase of the MBChB programme at the UFS during the focus group interview;

vi) To align, through triangulation, the current assessment practices with the policies, guidelines and principles of assessment, to enhance best-practice assessment and moderation in the clinical phase of the MBChB programme at the UFS; and

vii) To propose best-practice recommendations to guide the quality of assessment in the clinical phase of the MBChB programme at the UFS.

1.8 RESEARCH DESIGN

1.8.1 Research paradigm

Creswell (2014:6) describes a research paradigm as a "knowledge claim" about the research, and a foundation from which the research will be approached. Crotty (1998:5) developed a diagram to assist with research design and planning. An adaptation of this diagram is depicted in Table 1.1, to summarise the research design that was followed in this research.

Table 1.1: Research design that was followed in this thesis (Adapted from Crotty 1998:5 by researcher Brits 2020)

PARADIGM	THEORETICAL PERSPECTIVE	METHODOLOGY	METHODS
What will be my knowledge claim?	Which approach will I use?	How will I go about finding out?	Which techniques will I use?
Pragmatism was the paradigm that was used to address the practical problem of uncertainty regarding the quality of assessment in the clinical phase of the MBChB programme.	An **abductive approach** was used to achieve inference.	**Mixed methods** methodology was implemented by using both quantitative and qualitative methods to collect data. During triangulation, alignment of and gaps between theory and practice were identified and addressed.	Part 1: Rapid review Part 2: Questionnaires Quantitative comparison of marks Part 3: Focus group interview

The paradigm that was chosen, therefore, informed the research perspective, methodology and methods. Pragmatism was the philosophical foundation of this research project. The term pragmatism is described as a practical approach to problems and affairs (Morgan 2014:1045). Although there is no consensus about the philosophical foundation of mixed method research, the main advocates for mixed method research (Johnson & Onwuegbuzie 2004:24; Tashakkori & Teddlie 2010:22) express the view that, in a mixed method approach, the research tool and methods are more important than a specific research philosophy. However, the philosophical foundation is still important for understanding how the researcher approached and reports on the research (Creswell 2014:4).

1.8.2 Theoretical perspective

Charles Sanders Peirce is considered to be the father of the abductive approach; however, he considered Aristotle to be the founder of the abductive, or what was then known as the reductive, approach (Danermark *et al*. 2001:89; Kovács & Spens 2005:135). Peirce divided inference into three kinds of thinking: induction, deduction and abduction. Induction is described as a process that involves a generalisation about knowledge or observation. Deduction starts with a theory and then bases an inference on facts. In abduction, an inference is based on the best available or most likely explanation (Jiang 2001:94). An abductive approach usually starts with an observation that does not match a known theory, or an unexpected observation. A cyclical process of interaction between theory and practice then takes place, until a "new theory" can be suggested. This theory can then be applied to practice (Kovács & Spens 2005:135). Figure 1.2 is adapted from the work published by Kovács & Spens (2005:139), and demonstrates the abductive approach that was used in this project.

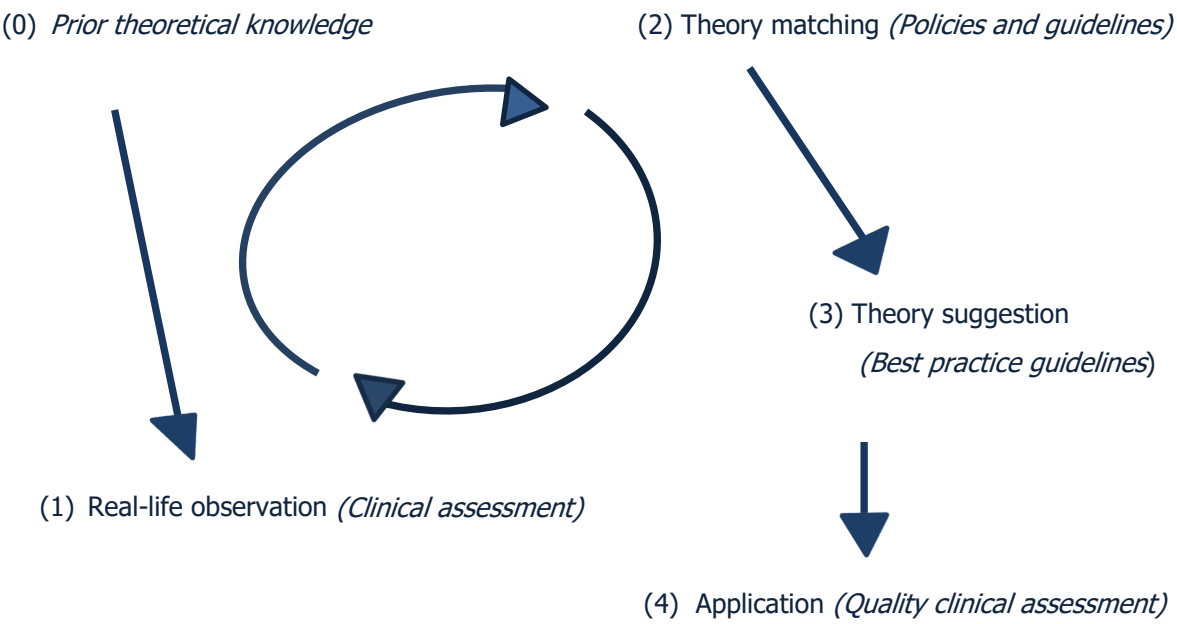


Figure 1.2: Schematic flow of the abductive approach that was applied (Adapted from the work by Kovács and Spens by the researcher Brits 2020)

1.8.3 Research methodology

A mixed methods research methodology, which is "problem-driven" and uses both qualitative and quantitative approaches to complement or support each other within the same research project (Denscombe 2014:153; Fouché & Delport 2011:66), was used. In quantitative research, the researcher collects information that can be counted or which is numerical, to test a relationship between variables. Closed-ended questions are asked and data is collected using valid and reliable tools. Data is then coded and entered into a computer program for analysis. The researcher cannot merely display the results, instead, he/she needs to interpret the results to make sense of it (Ivankova, Creswell & Plano Clark 2011:257). In qualitative research, the researcher wants to improve understanding and/or obtain explanations for behaviour or views. Data is collected in the form of words and then grouped as themes. Data can be collected from interviews, observations, documents, audio-visual material, and even artefacts (Ivankova *et al*. 2011:259).

Figure 1.3 displays the triangulated mixed methods model that was used in this research.

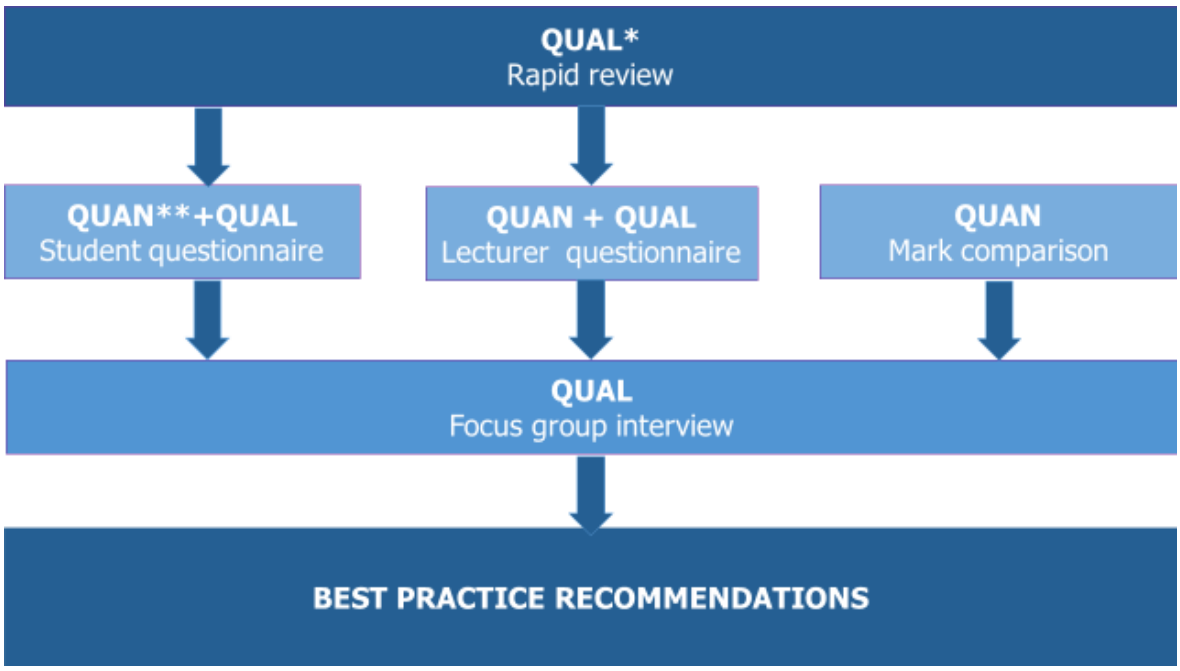


Figure 1.3: Schematic display of the triangulated mixed methods research model used for this study (Compiled by the researcher Brits 2020)
***QUAL = qualitative **QUAN = quantitative**

A triangulated mixed methods design, as described by Creswell (2014:40), was used for this project. The triangulated mixed methods design is described as a method that collects and interprets qualitative and quantitative data on the same topic from different sources separately. This data is then interpreted separately and compared, to validate, confirm or strengthen findings (Creswell & Plano Clark 2011:62). Qualitative data was collected from the rapid review, while quantitative data was collected with the aid of student and lecturer questionnaires and student marks. After separate data interpretation, the theory and practice results were used together in the focus group interview to, first, clarify concepts and to find areas of similarities and differences. Solutions were sought to bridge the gap between theory and practice.

1.9 RESEARCH METHODS

1.9.1 Part 1 – Rapid review

In Part 1, qualitative data was gathered on current assessment policies and best assessment practices using a rapid review. Researchers hold different opinions about the use of literature or documents as a research method, however, researchers, such as Bowen (2009:27) and Creswell and Poth (2018:43), include document reviews, specifically, as a qualitative research method. The challenges of presenting an undergraduate medical programme in the context of a unique environment (clinical training area), and fulfilling requirements of various regulatory bodies were taken into consideration. As described by Creswell and Plano Clark (2011:8), qualitative data that was collected was used to formulate a theoretical guide to measure and inform quality in the assessment of students in the clinical phase of the MBChB programme. Information gathered during Part 1 of the study also guided the researcher to adjust the draft questionnaire on current assessment practices that was administered in Part 2, and with the interpretation of results.

Article 1: A framework to benchmark the quality of clinical assessment in a South African undergraduate medical programme.

1.9.2 Part 2 – Review of current assessment practices

In Part 2 of the study, mainly quantitative data was collected from three sources. Firstly, the opinions of clinical students regarding assessment practices in the clinical phase of the MBChB programme were gathered through anonymous questionnaires. Secondly,

information on current assessment practices in the clinical phase of the MBChB programme was gathered using a questionnaire survey completed by academic staff (T&L coordinators and/or module leaders in the Phase III workgroup). Thirdly, students' assessment results (marks obtained during the block and end-of-year assessments) were compared for reliability, predictability and stability. The purpose of the mark comparison and analysis was to evaluate aspects of the reliability of current assessment results. Reliability is prerequisite for valid assessments (Thatcher 2010:125; Twycross & Shields 2004:36).

Quantitative and some qualitative data collection was done in three parts, through two sets of questionnaires and the analysis of students' assessment results. These steps provided evidence of the quality of assessment in the clinical phase of MBChB training. The data from Parts 1 and 2 were used for triangulation in Part 3.

1.9.2.1 *Review of student opinions*

Article 2: Students' voices: Assessment in undergraduate clinical medicine.

1.9.2.2 *Evaluation of assessment practices*

Article 3: Assessment practices in undergraduate clinical medicine training. What do we do and how we can improve?

1.9.2.3 *Predictability and evaluation of students' assessment marks*

Article 4: Evaluation of assessment marks in the clinical years of an undergraduate medical training programme: Where are we and how can we improve?

1.9.3 Part 3 – Focus group interview

In Part 3 of the study, the results of the quantitative and qualitative data collected during Parts 1 and 2 were used to inform questions for a focus group interview with T&L coordinators involved in the clinical phase of the MBChB programme. The focus group interview aimed to obtain qualitative data that contributed to identifying best assessment practice recommendations, to address the gap between current assessment practices and theoretical policies and principles that inform assessment.

Article 5: Quality assessment in undergraduate medical training: How to bridge the gap between what we do and what we should do.

1.9.4 Contribution – Best-practice recommendations

Finally, the results of Parts 1 to 3 were synthesised to propose best-practice recommendations prepared for the Executive Committee of the School of Clinical Medicine to enhance quality assessment in the clinical phase of the MBChB programme. As part of quality assurance, a moderation recommendation was included to ensure regular reassessment of the assessment practices.

Proposal: Best practice recommendations to enhance quality assessment in the clinical phase of the undergraduate medical programme at the University of the Free State.

1.10 TRUSTWORTHINESS

Guba (1981:79) states that, for research to be trusted, it must have truth value, it must be applicable, it must be consistent and the results must not be influenced by outside factors, such as bias or opinion. In mixed method research, both qualitative and quantitative data is gathered. It was, therefore, important that aspects that ensure the integrity and quality of both types of data were addressed.

For qualitative research, the criteria to ensure the rigour of the findings are credibility, transferability, dependability and confirmability (Guba 1981:80; Schwandt, Lincoln & Guba 2007:11).

1.10.1 Credibility

Credibility is described as the confidence that researchers have in the truth of their findings and the context in which data was gathered (Guba 1981:79). Credibility can be established through triangulation of results by using different methods (method triangulation) to gather data, and by using different people (source triangulation) to gather data. Good interview technique and checking and clarifying responses with participants will also improve credibility (Onwuegbuzie & Leech 2007:239). In this study, credibility was ensured using method and source triangulation of results.

1.10.2 Transferability

Transferability is the degree to which the data or results that were gathered will be applicable or can be extrapolated to another setting or participants (Guba 1981:79–80). By describing the setting, method and the sampling in this study well, other researchers can decide on transferability (Li 2004:305). Including local, national and international guidelines in the rapid review, means the results of this study should be transferable to other settings.

1.10.3 Dependability

Dependability describes the degree to which the data is constant or repeatable if collected under similar circumstances (Guba 1981:80). To increase dependability of this study's data, all processes followed were described accurately, as explained in the different methods in the papers.

1.10.4 Confirmability

Confirmability is "concerned with establishing that data and interpretations of the findings are not figments of the inquirer's imagination, but are clearly derived from the data" (Tobin & Begley 2004:392). To ensure confirmability, the data interpretation was discussed and clarified with the promoters.

The following criteria ensured the vigour of quantitative research: internal and external validity, reliability and objectivity (Guba 1981:80; Schwandt *et al.* 2007:11).

1.10.5 Internal and external validity

For the research to be valid, it should measure what it intends to measure (Katzenellenbogen & Joubert 2014:123; Kelley 1927:14). The validity of research can be increased by using different instruments to collect data, and then triangulating the data (Creswell & Poth 2018:260). In this research, attention was given to both internal and external validity. For internal validity, the face validity of the questionnaires was tested in a pilot study, and the focus group interview questions were reviewed by experts in the field before it they were used. The questions in the questionnaires derived from the literature review, to ensure construct and content validity. External validity was addressed by

consulting local, national and international guidelines on policies and quality assessment, for inclusion in the data collection instruments. This increased generalisability of the results.

1.10.6 Reliability

Reliability is defined as the extent to which an assessment tool produces consistent, stable and repeatable results (Heale & Twycross 2015:66). Only piloted questionnaires were used for data collection. The accuracy of the data from the questionnaires entered on the Excel data sheet was confirmed by double typing. Student marks were not retyped; instead, the original data sheets that were provided were used for data analysis. Spot checks were done to ensure the accuracy of the data used. Triangulation between the qualitative and quantitative data was used to identify inconsistencies.

1.10.7 Objectivity

Lather (1990:319) states that, "objectivity means being aware and honest about how one's own beliefs, values, and biases affect the research process". To increase the objectivity of this study, the researcher took note of factors that may undermine the objectivity, and minimised the effect of these factors as far as possible. Results were discussed with an expert in biostatistics, to ensure the correctness of interpretations.

1.11 ETHICAL CONSIDERATIONS FOR THE STUDY

Ethical approval for the study was requested and granted (UFS-HSD 2019/0001/2304) through the Health Sciences Research and Ethics Committee (HSREC) via the Research Information Management System of the UFS. Approval to include medical students and UFS personnel was obtained from UFS authorities (head of School: Clinical Medicine, the dean of the Faculty of Health Sciences, the vice-rector: Research and the dean of Student Affairs) through the completion of a HSREC 17 form. Participation by students was voluntary, with implied consent being given by completing the questionnaires. Informed consent was signed by all T&L coordinators and module leaders who completed questionnaires and/or participated in the focus group interviews. Participants did not receive any compensation for participating, neither were they penalised for not participating in the study. All questionnaires were anonymous. Although it was possible to identify specific departments from the questionnaires and student marks, no departments are mentioned in the results. All data was managed confidentially.

1.12 DEMARCATION AND SCOPE OF THE STUDY

This study was done in the field of health professions education and lies in the domain of assessment of undergraduate medical training in the clinical phase of the medical programme. The study is interprofessional, as it reached across health professions education and medicine.

1.13 VALUE, SIGNIFICANCE AND CONTRIBUTION OF THE STUDY

1.13.1 Value

The value of this research study will be to enhance assessment and quality assurance practices in the clinical phase of the MBChB programme in the School of Medicine, Faculty of Health Sciences at the UFS.

1.13.2 Significance

The study may be used to enhance quality assessments in the clinical phase of the MBChB programme at the UFS. Quality assessment is necessary if the Faculty is to state with confidence that students who pass the examination are competent to work as medical doctors, and that students who fail need more training. Quality assessment will increase the credibility of the training programme and the university.

1.13.3 Contribution

This study provided best practice recommendations to enhance the quality of assessment of MBChB students in the clinical phase of training. Implementing these recommendations will help to ensure accurate judgement on student performance and to bridge the gap that exists between current assessment practices and prescribed policies and guidelines on quality assessment. Results from this study were published in peer-reviewed journals for others to apply in clinical training facilities and programmes.

1.14 SCHEMATIC OVERVIEW OF THE STUDY

Figure 1.4: A schematic overview of the study (Compiled by the researcher Brits 2020)

1.15 SCHEMATIC CHAPTER LAYOUT OF THE STUDY

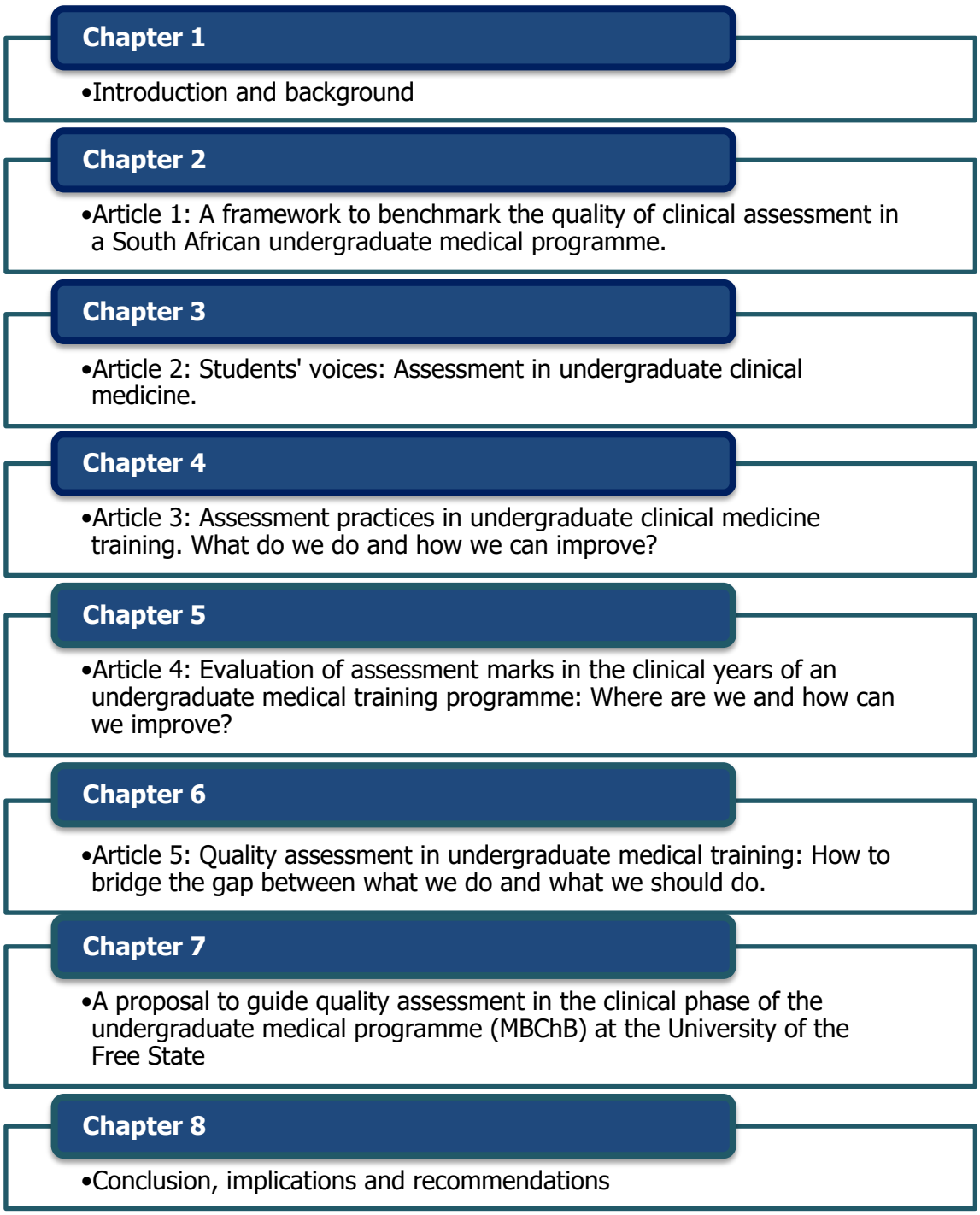


Figure 1.5: A schematic chapter layout of the study (Compiled by the researcher Brits 2020)

1.16 CHAPTER SUMMARY

Chapter 1 oriented the reader by describing the background, context and content of the study. This chapter addressed assessment practices, assessment in the undergraduate clinical medical programme, and possible gaps in the assessment, and then described methods followed to identify and address these gaps.

The research design of the study was described, with a justification for the methodology used to address the research questions. The ethical implications and approval, as well as the quality and rigour of the study, were discussed. The chapter concluded with a schematic overview and chapter layout of this thesis.

Chapters 2 to 6 will provide the results of the first five research objectives, in the form of articles, and Chapter 7 will offer a proposal to guide quality assessment in the clinical phase of the undergraduate medical programme (MBChB) at the UFS.

CHAPTER 2:

ARTICLE 1 – A FRAMEWORK TO BENCHMARK THE QUALITY OF CLINICAL ASSESSMENT IN A SOUTH AFRICAN UNDERGRADUATE MEDICAL PROGRAMME

This article was prepared according to the submission guidelines of the *South African Family Practice* journal.

Journal:	South African Family Practice
Publisher:	AOSIS Publishing – African Online Scientific Information Systems
Country of publication:	South Africa
ISSN:	1025-9848 (PRINT) 2071-9736 (ONLINE)
DOI:	10.4102/safp.v62i1.5030

Citation:
Brits, H., Bezuidenhout, J. & Van der Merwe, L.J. 2020. A framework to benchmark the quality of clinical assessment in a South African undergraduate medical programme. *South African Family Practice*. 62(1), a5030. DOI: 10.4102/safp.v62i1.5030

Submitted:	20 September 2019
Accepted:	18 November 2019
Published:	4 February 2020


SAFP
SOUTH AFRICAN
FAMILY PRACTICE
SOUTH AFRICAN ACADEMY OF
FAMILY PHYSICIANS
ISSN: 2078-6190
2020
AOSIS
Open access at safpj.co.za
Vol. 62 No. 1 Part 1

Overview

ISSN

1025-9848 (PRINT)
2071-9736 (ONLINE)

Focus and scope

The *South African Family Practice* (SAFP) journal is the official journal of the South African Academy of Family Physicians (SAAFP), and is aimed at all SAAFP members (including family physicians, registrars, associate members, students), working within primary care (both private and public health sectors, as well as urban and rural practice settings) within South Africa and the wider region. SAFP is a peer-reviewed scientific journal, which strives to provide primary care physicians (and their teams), as well as researchers, with a broad range of scholarly work in the disciplines of family medicine, primary health care, rural medicine, district health and other related fields. SAFP publishes original research, clinical reviews, and pertinent commentary that advance the knowledge base of these disciplines. The content of SAFP is designed to reflect and support further development of the broad basis of these disciplines through original research and critical review of evidence in important clinical areas; as well as to provide practitioners with continuing professional development material. SAFP adheres to the international acceptable editorial standards, as published by The International Committee of Medical Journal Editors (ICMJE).The journal's editors are supported by an editorial board, which consists of South African members representing the nine academic training programmes as well as a representative from RuDASA (Rural Doctors Association of South Africa), and key members from the international family medicine and primary care community.

DHET Accreditation

The journal is DHET accredited because it is listed on the following approved indexing services:

- SCOPUS

Editor-in-Chief

Assistant Editor

South African Family Practice
ISSN: (Online) 2078-6204, (Print) 2078-6190

AOSIS

Page 1 of 9 Review Articles

A framework to benchmark the quality of clinical assessment in a South African undergraduate medical programme

Authors:
Hanneke Brits[1]
Johan Bezuidenhout[2]
Lynette J. van der Merwe[3]

Affiliations:
[1]Department of Family Medicine, Faculty of Health Sciences, Free State University, Bloemfontein, South Africa

[2]Department of Health Professions Education, Faculty of Health Sciences, Free State University, Bloemfontein, South Africa

[3]Department of Undergraduate Programme Management, Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa

Corresponding author:
Hanneke Brits,
britsh@ufs.ac.za

Dates:
Received: 20 Sept. 2019
Accepted: 18 Nov. 2019
Published: 04 Feb. 2020

How to cite this article:
Brits H, Bezuidenhout J, Van der Merwe LJ. A framework to benchmark the quality of clinical assessment in a South African undergraduate medical programme. S Afr Fam Pract. 2020;62(1), a5030. https://doi.org/10.4102/safp.v62i1.5030

Background: The outcome of the undergraduate medical programme is to produce clinically competent health care providers relevant for the South African context. Educational institutions find it hard to ensure the quality of assessments where competency must be assessed. This study aimed to compile an assessment framework that can be used to benchmark current assessment practices in the clinical phase of the undergraduate medical programme where competency must be certified.

Methods: In this observational, descriptive study, qualitative data were gathered using the steps described by the World Health Organization for rapid reviews. Literature was searched, screened and selected before data were analysed and a framework was constructed.

Results: Twenty-five official documents were included in the study. The framework addressed the three components of quality assessment, namely, accreditation, assessment and quality assurance. Assessors should attend to the principles of assessment, namely, validity, reliability, fairness, feasibility, educational effect and acceptability, but realise that no assessment meets all these criteria. The first step to ensure quality assessment is to identify a clear outcome. Assessment should be planned and aligned with this outcome.

Conclusion: It is clear that clinical assessment is multidimensional and that no assessment is perfect. Programme accreditation, assessment practices and psychometrics can assist to improve the quality of assessment but cannot judge clinical competence. Using experienced assessors with a variety of assessment methods on a continuous basis is the proposed way to assess clinical competence. An assessment framework can assist to improve assessment, but it cannot guarantee quality assessment.

Keywords: accreditation; assessment policies; assessment guidelines; clinical assessment; quality assurance in assessment; principles of quality assessment; undergraduate assessment.

Introduction

In South Africa, undergraduate medical training programmes are offered at nine accredited universities.[1] The formal undergraduate medical training programme offered is the Bachelor of Medicine and Bachelor of Surgery (MBChB). Passing the final assessment of the MBChB programme enables a student to graduate and qualify as an entry-level medical practitioner.[2] The main outcome of the MBChB programme is to produce clinically competent health care providers relevant for the South African context.[3] This context is determined by the quadruple burden of disease (e.g. trauma, gastro-enteritis with dehydration, malnutrition, maternal and labour complications, human immunodeficiency virus and tuberculosis as well as lifestyle diseases),[4] which, in turn, determines the required competencies to manage these conditions. Clinical assessment is unique because of integration of competencies and more than one possible correct approach to a problem.[5] If clinical competence is assessed on the 'does' level of Miller's pyramid,[6] there will always be a compromise on reliability.[7] Locally, the educational institution finds it hard to defend the quality of high-stakes competency assessments against validity, reliability and fairness. This study forms part of a larger study addressing quality of assessment in the clinical phase of the undergraduate medical programme.

Three components of quality in assessment have been described in the literature, namely, accreditation, assessment and audit.[8,9] In this article, registration is added to accreditation, as the MBChB degree is a professional qualification that must be accredited and registered with various authorities. The term audit is replaced by quality assurance, as both these terms serve the purpose

of improving quality; however, the term audit is usually used in the financial context, and quality assurance in education and other fields.[10]

To assess the quality of an assessment, it must be benchmarked against appropriate criteria. Benchmarking is described as the process of comparing standards with external criteria, with the aim of improvement.[11] Clinical assessment should be benchmarked against best-practice evidence to ensure global relevance.[12] Pangaro and Ten Cate recommended an assessment framework to benchmark assessment and competence against.[13] In developing a framework, it is essential to clarify and/or define the terms or concepts that form the basis of the framework.[14] McCall states that a good definition contains previously defined words, classifies and quantifies and has no counterexamples.[15] The following terms and processes are used in this article.

Framework

An assessment framework can be described as a 'common language and mental model' that guides assessors on what to look for in student assessment to maximise the reliability of the assessment. This framework also informs students and leadership on what to expect during assessment.[13] It is important to realise that not all aspects of a framework necessarily apply to all assessments.[16] A synthetic framework that integrates the domains of knowledge, skills and attitudes to ensure competence in real-life situations[13] has been proposed for this study. Furthermore, for a framework to be of practical value, it should be simple enough to understand, remember and implement, while training and monitoring should form part of the implementation process.[13]

Accreditation

Accreditation entails certification, which confirms that a programme and/or training facility is capable of fulfilling required specifications for a specific period. For instance, the South African Qualifications Authority (SAQA) accredits the providers who offer outcomes-based learning programmes that are aligned with registered unit standards and qualifications of the National Qualifications Framework (NQF).[17]

Assessment

South African Qualifications Authority defines assessment as 'a process used to identify, gather and interpret information and evidence against required competencies' in order to make a judgement about a learner's achievement.[18] The University of the Free State (UFS) describes assessment as 'the process of determining the value, significance, or extent of what students know, understand and can do with their knowledge as a result of their educational experience'.[19] Assessment is therefore a comprehensive process, includes a variety of measurements for judging performance. The content and standard of assessment, types of assessment, assessment methods and principles of assessment are included under this concept.

Quality assurance

Quality, standards and relevance are key elements of quality assurance.[9] As far back as 1980, Donabedian defined quality in order to measure it.[20] He concluded that quality is not one-dimensional but includes various aspects that should be agreed upon before the measurement. By applying the criteria of validity, reliability and defined concepts, quality has been described as 'a situation when a set of inherent characteristics consistently fulfil the requirements of the organization's ... stakeholders'.[21]

The context of undergraduate medical training and assessment

The duration of undergraduate medical training is between 5 and 6 years, and it is offered in three phases, namely, orientation, pre-clinical and clinical training.[1] The MBChB programme at the UFS is a 5-year (10-semester) outcomes-based programme presented at the UFS and the accredited training platforms of the Free State Department of Health. The clinical phase of the MBChB programme comprises the second half of the third year, the fourth year and the fifth year (semesters 6–10) of undergraduate medical training. During this phase, students rotate through the different clinical disciplines and receive clinical exposure to patients, as well as theoretical training. Formative assessment takes place during rotations and summative assessment at the end of the academic year. To progress to the next year, a student must pass assessments in all disciplines, and both the clinical and theoretical components separately.[22]

A preliminary literature review was performed to identify frameworks, policies and criteria on quality assessment in the clinical phase of an undergraduate medical programme. At the institutional level, the assessment policy of the UFS sets minimum criteria for assessment of undergraduate students, which requires alignment with national policies and acts.[23] The Health Professions Council of South Africa (HPCSA) also prescribes core competencies (adapted from CanMEDS) for undergraduate medical students.[24] Internationally, the guidelines of the Association for Medical Education in Europe (AMEE) describe the importance of frameworks for clinical assessment,[13] and the World Federation of Medical Education (WFME) has published basic standards for assessment in undergraduate training.[25] Although these guides provide valuable information and principles, they are not directly transferable to the undergraduate South African and UFS medical training context where major emphasis is placed on clinical skills and clinical competence in the workplace. These guidelines focus on general programmes or assessment practices, and they do not apply specifically to assessment in the clinical phase of an undergraduate medical programme.

Research question

What are the current regulations and policies as well as best evidence practices that inform quality assessment in the clinical phase of an MBChB programme?

Objective

To compile a framework that can be used to benchmark current assessment practices based on official regulations and policies, and supported by best evidence practices to ensure defendable assessment in the clinical phase of the MBChB programme in South Africa.

In this article, a rapid review of the regulations and policy documents of the bodies that regulate the assessment and training of the MBChB programme at the UFS was used to formulate a framework for clinical assessment. This framework may be helpful to benchmark the quality of assessment in the current MBChB programmes in South Africa and beyond its borders.

Methods

Qualitative data were gathered using a rapid review. No formal definition or uniform method is described to conduct a rapid review although a rapid review can be described as a simplified systematic review.[26,27] Rapid reviews are typically used to inform decisions and compile guidelines.[28,29] The single research question, narrow time frame, limited data sources, use of a single reviewer and minimal data synthesis[26] justify the use of a rapid review in this article.

The World Health Organization (WHO) proposed seven steps to follow for a systematic review, which may be adjusted according to the specific needs for a rapid review.[29] These steps are:

- needs assessment and topic selection
- protocol development with or without registration
- literature search
- screening and selection of the literature
- data extraction
- risk-of-bias and quality assessment of data
- knowledge synthesis.

In rapid reviews, some of these steps are commonly simplified or omitted, but the description of the method should not be compromised.[26] The components most commonly adjusted are the use of one reviewer instead of two reviewers, not conducting quality assessments of included data and not using grey data.[27]

The following steps were used in this rapid review.

Topic selection

The difficulty to defend the quality of clinical assessment in an undergraduate medical programme was identified as an area to investigate.

Protocol development and registration

A protocol was developed before the study commenced. The protocol limited the inclusion of primary source documents to the following:

- official regulations and policy documents of the regulatory bodies responsible for assessment in undergraduate medical education at the UFS and South Africa
- international guidelines on clinical assessment issued by the WFME
- the AMEE guidelines on frameworks for clinical assessment.

The primary outcomes to investigate were:

- accreditation and registration
- assessment
- quality assurance.

Systematic document search

The following search strategies were followed:

- No date limitations were placed on the documents included in the review.
- The search process was conducted from May to July 2019.
- An Internet data search was conducted.
- The *Health Professions Act, 56* of 1974[2] was used as the original document source for national policies and the UFS assessment policy[23] for institutional policies.
- These databases identified from the literature study were consulted: UFS official website, HPCSA, SAQA, Council on Higher Education (CHE), WFME and AMEE websites.
- The following words and/or phrases were searched: accreditation, assessment policies, assessment guidelines, clinical assessment, quality assurance in assessment, principles of quality assessment and undergraduate assessment. Searches were conducted with single words and phrases and the inclusion and exclusion of 'AND' and 'OR'.
- Backward searching was performed using references and cross-references to related policies and regulations of the identified regulatory bodies.
- Forward searches of the literature entailed the search for related and updated information from the same documents or topics to ensure that all relevant information was identified.

Screening and selection of articles

- Documents were screened for selection by a single review author. Documents initially found not to meet the outcome of the study were not included but saved separately. These documents were screened a second time to ensure that relevant data were not omitted. When in doubt, the study leader could assist with selection decisions.

Data extraction

Documents were grouped according to the primary outcomes that were accreditation and registration,

assessment and audit. The assessment category was subdivided into the following subcategories:

- assessment content and standards
- assessment types
- assessment methods
- principles of quality assessment.

A table displayed the specific outcomes that were addressed by each document included in the study.

Limiting the risk-for-bias

This was omitted in this review, although care was taken to include all relevant documents by following the prescribed protocol. Document quality was not assessed as only official policies and regulations were included.

Knowledge synthesis

For each category, the results of the review were summarised and discussed. This was supplemented by a secondary literature search to clarify concepts. The guidelines for framework development described by Pangaro and Ten Cate[13] were then followed to display the results visually. Finally, recommendations were made for the implementation of the framework and research limitations were discussed.

Quality and rigour of the data collection

To ensure the credibility of the data collected and to ensure that relevant documents were included in the document review, the protocol was strictly followed. National and international guidelines were added to enable the transferability of results to other institutions. The steps followed in the rapid review were described clearly to assess the dependability of the results.

Ethical considerations

The study was registered and approved by the Health Sciences Research and Ethics Committee (HSREC) at the University of Free State (UFS-HSD 2019/0001/2304). As only documents in the public domain were used for this literature review and analysis, no permission was necessary.

Results

The MBChB programme is offered under the legislation of the Department of Health and the Department of Education (previously the Ministry of Education). The *Health Professions Act, 56* of 1974[2] was used as the original document source for national policies, the UFS assessment policy[23] for local policies and the AMEE (https://amee.org/home) and WFME (https://wfme.org/) websites to benchmark against international assessment principles. Twenty-five documents were included in the rapid review. Table 1 displays the documents used in this rapid review.

Accreditation and registration

According to the *Higher Education Act*, the Ministry of Education must oversee and take responsibility for norms and standards in higher education.[30] To assist with this task, the Minister of Education and Training established SAQA as a juristic person who must implement the objectives of the NQF.[18] The NQF was established under the *SAQA Act* to classify, register, publish and articulate approved national qualifications.[33] Medical training is addressed under the sub-framework for higher education. The CHE, as the Quality Committee for Higher Education as provided for in the *Higher Education Act*, oversees the quality of training and assessment in higher education. The quality committee must register appropriate professional bodies (in this case, the HPCSA) to ensure that qualifications meet the requirements for professional registration. The quality committees make recommendations to SAQA to register higher education qualifications.[38]

The *Health Professions Act* makes provision for appointing professional councils (in this case, the HPCSA) to establish professional boards. The Medical and Dental Board is responsible for overseeing undergraduate medical training, as well as registering health professionals under this act.[31] The HPCSA is responsible for accrediting universities and health care training in South Africa.[38]

Because of globalisation and the increased demand for accountability in health care, the WHO and the WFME worked together on documents for the accreditation of health training institutions worldwide. The WFME gives 'recognition status' to an accrediting agency that meets international standards.[43,45]

Assessment

Four components of assessment were identified, namely, assessment content and standards, assessment types, assessment methods and principles of assessment.

Assessment content and standards

An assessment to ensure a competent practitioner must include elements of knowledge, skills and values.[2,18,23,27] South African Qualifications Authority describes knowledge as foundational competence, skills as practical competence and values as reflective competence. It also emphasises the importance of assessing prior learning, and that assessment must include content to identify and stimulate further learning.[34] Assessment of values, also described as core competencies, soft skills or critical cross-field competencies, has been prescribed as components of assessment in different documents.[18,19,24,34] Critical cross-field competencies identified were problem-solving, critical thinking, teamwork, responsibility, data management, effective communication and effective use of resources.[34] The core competencies for a health care practitioner include being a professional, a communicator, collaborator, leader and manager, health advocate and scholar.[24] Assessment

TABLE 1: Primary documents used in document review.

Document	Accreditation or registration	Assessment				Quality assurance
		Assessment content and standard	Assessment types	Assessment methods	Principles of quality assessment	
South Africa. Council on Higher Education. *Higher Education Act 101* of 1997.[30]	√	-	-	-	-	-
South Africa. *Health Professions Act 56* of 1974 (Amended 2007). Education training and registration.[31]	√	-	-	-	-	-
South Africa. *Health Professions Act 56* of 1974 (Amended 2009). Regulations relating to the registration of students, undergraduate curricula and professional examinations in medicine.[2]	-	√	√	-	√	√
Health Professions Council of South Africa. Core competencies for undergraduate students in clinical associate, dentistry and medical teaching and learning programmes in South Africa 2014.[24]	-	√	-	-	-	-
Health Professions Council of South Africa. Accredited facilities. 2019.[1]	√	-	-	-	-	-
Health Professions Council of South Africa. Professional Boards. 2019.[32]	√	-	-	-	-	-
South African Qualifications Authority. *South African Qualifications Authority Act 58* of 1995.[33]	√	-	-	-	-	-
South African Qualifications Authority. The National Qualifications Framework Curriculum Development. 2000.[34]	-	-	√	-	√	-
South African Qualifications Authority. National Qualifications Framework and the Standards setting. 2003.[35]	-	-	-	-	√	√
South African Qualifications Authority. Criteria and Guidelines for Assessment of NQR Registered Unit standards and Qualifications. 2001.[36]	-	√		√	-	-
South African Qualifications Authority. Guidelines for integrated assessment. 2005.[37]	-	√	√	-	√	-
South African Qualifications Authority. National Policy and Criteria for Designing and Implementing Assessment for NQF Qualifications and Part-Qualifications and Professional Designations in South Africa. 2014.[18]	-	-	√	√	-	-
South Africa. National Qualifications Framework. *National Qualifications Framework Act 67* of 2008.[38]	√	-	-	-	-	-
University of the Free State. Teaching-Learning Policy. 2008.[39]	-	-	√	√	√	-
University of the Free State. Quality assurance policy. 2009.[40]	-	-	-	-	-	√
University of the Free State. Guidelines for the implementation of external moderation. 2009.[41]	-	-	-	-	-	√
University of the Free State. Assessment policy on the UFS coursework learning programme. 2016.[23]	-	-	√	√	√	√
University of the Free State. General rules for undergraduate qualifications, postgraduate diplomas, Bachelor Honours degrees, Master's degrees, Doctoral degrees, Higher Doctorates, Honorary degrees and the Convocation. 2019.[29]	-	-	√	-	√	-
University of the Free State. Faculty of Health Sciences. Rule book. School of Clinical Medicine. Undergraduate Qualifications. 2019.[22]	-	√	√	-	√	-
University of the Free State. School of Clinical Medicine. Undergraduate programme management. 2019. SOP Quality assurance.[42]	-	-	-	-	-	√
World Federation for Medical Education. 2015. Standards. Basic Medical Education.[25]	-	-	√	√	√	√
World Federation for Medical Education. Accreditation. 2017.[43]	√	-	-	-	-	-
Pangaro L, Ten Cate O. AMEE Guide No. 78.[13]	-	√	√	√	√	-
Tavakol M, Dennick R. AMEE Guide 119.[44]	-	-	-	-	√	-

NQR, National Qualification Register; NQF, National Qualification Framework; UFS, University of the Free State; SOP, Standard Operating Procedures; AMEE, Association for Medical Education in Europe.

standards are the minimum criteria that must be achieved to pass an assessment. These standards include criteria for content and difficulty, and should be reasonable, defensible and fair. Students and assessors must know all the required standards before the assessment.[35] The MBChB programme is registered on NQF level 8.[35] Although there is no perfect passing score, the UFS sets the pass mark at 50%.[23] In Clinical Medicine, students must pass both theoretical and clinical assessments separately in order to progress.[2,22] None of the documents addressed specific standard setting methods or processes.

Assessment types

Different types of assessment applicable to medical assessment were identified from the document review, namely, formative assessment, integrated assessment and summative assessment.[2,18,19,37,39] Some of these types may overlap or be inclusive of each other; for example, integrated assessment may take place during formative and/or summative assessment.

Formative assessment is described as a series of assessments that occur during the learning and training process.[18,23] The purpose of formative assessment is to support learning, identify learning needs and accumulate marks.[23]

Summative assessment is the assessment that takes place after learning. The aim of summative assessment is to award grades and to validate performance and competence.[18,23] Integrated assessment is described as 'assessment that permits the learner to demonstrate applied competence' using different methods of assessment.[18,34] Integrated assessment may occur at any time during the learning process.

Assessment methods

Theoretical, practical and integrated assessment methods were described, and they relate to the aim or outcome of the assessment. Theoretical assessments include multiple-choice questions, modified essay questions or short-answer questions, as well as long questions. Oral examinations

can be used to test knowledge or to combine knowledge with communication skills. Clinical assessments include unobserved long cases, mini clinical evaluation exercises (mini-CEX), objective structured clinical examinations (OSCEs) and direct observation of clinical practice (DOPS). Integrated assessment methods include portfolios, logbooks, elective reports and workplace-based assessments, as well as feedback from stakeholders.[19,23,25]

Principles of quality assessment

From the UFS general rules[19] and assessment policies,[23] which are aligned with the *Higher Education Acts*, the *NQF Act* and, by implication, the *Health Act*, the following principles were identified:

- Assessment should be an integral part of curriculum planning and must be aligned with outcomes.
- Assessment should be performed on the appropriate NQF level in accordance with programme registration.
- All assessments should be planned to cover all assessment domains.
- Assessment takes place in a system and must be planned accordingly.
- In order to be a quality assessment, each of these assessments should fulfil criteria for validity, reliability, transparency, fairness and practicability.
- Moderation should form part of overall, as well as individual, assessments.
- There should be accountability for each assessment, with evidence that the assessment was moderated.

An assessment can be considered credible if the criteria for fairness, validity, reliability and practicability have been met.[36]

Quality assurance

Quality assurance policies are essential for ensuring that specifications and standards are maintained.[23] This article focusses on quality assurance applicable to assessment and addresses moderation, benchmarking and security of the assessment process. Moderation is guided by moderation policies.[41] It is a process that involves a professional judgement of the validity, reliability and fairness of the assessment and involves students, assessors and external stakeholders. The WFME sets global standards for assessment that serve as benchmarks against which those responsible for medical education can evaluate their activities.[25]

Discussion

All the primary documents necessary for the rapid review were available in the public domain on the identified websites. Information in these documents was aligned with each other. Many cross-references to other documents were found in source documents. By comparing the information in the respective documents, it was found that there was no contradiction in the documents. The data included in the rapid review can therefore be considered representative and appropriate for the purpose of this study.

The three components of quality assessment, namely, accreditation and registration, assessment and quality assurance, should from part of an assessment framework to benchmark current assessment. The inclusion of best-practice evidence in the framework will make the framework globally relevant.[12]

Accreditation and registration

Accreditation and registration is usually not a problem for training facilities in South Africa as the HPCSA conducts regular site visits and requires annual progress reports from training facilities to ensure compliance with accreditation and registration requirements.[46] For the MBChB programme, the following must be in place:

- Accreditation of the training provider and the qualification by the HPCSA.
- Training may take place only at a university registered with the Department of Education.
- The qualification must be registered with SAQA.
- All students in the MBChB programme must be registered with the HPCSA.

A recommendation of the 2010 Ottawa Conference was to develop criteria for accreditation of international medical educational programmes.[47] In response, the WHO and WFME developed international accreditation criteria. The WFME awarded 'recognition status' to the HPCSA as the accrediting body in South Africa; all training programmes accredited by the HPCSA will therefore have internationally accredited status.[43]

Assessment

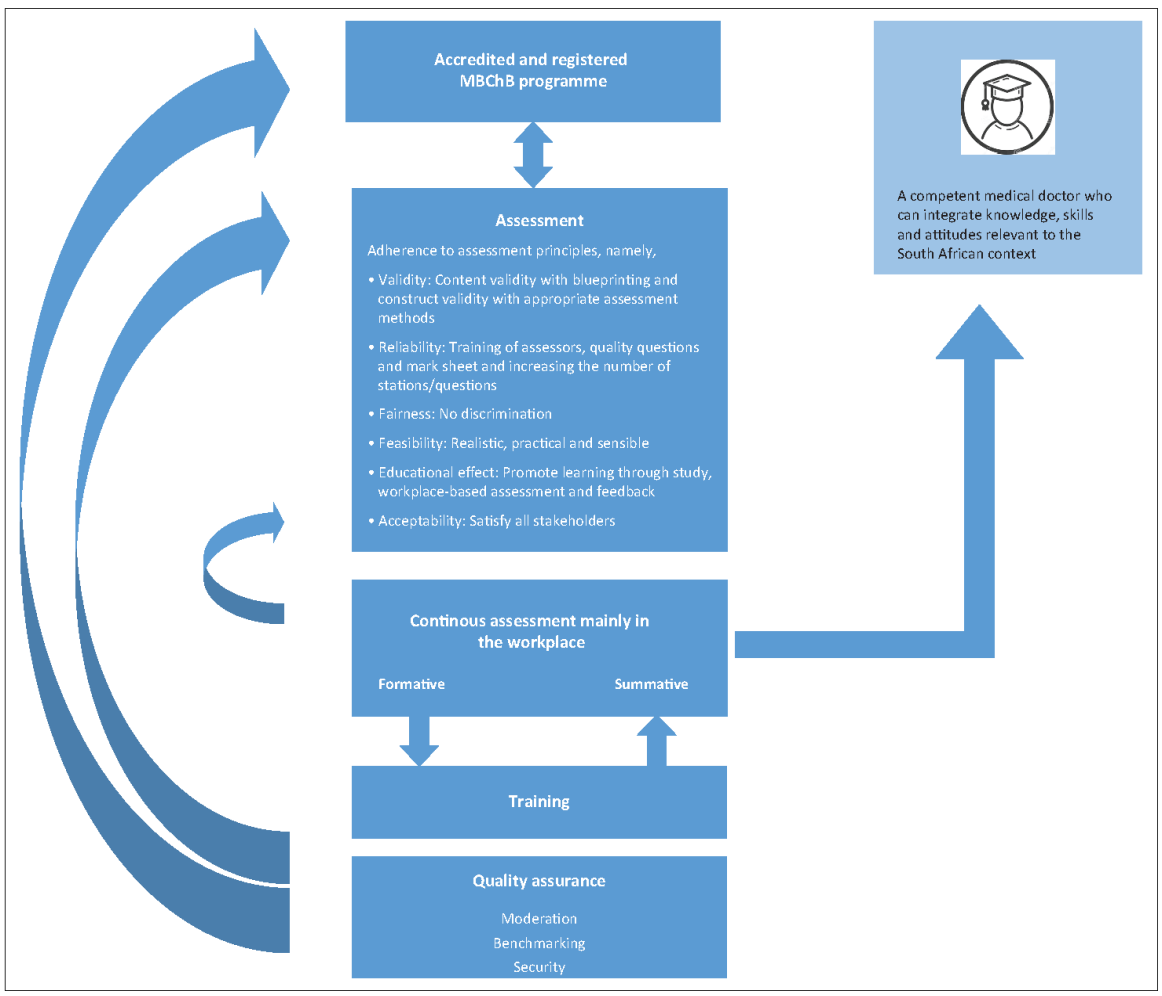
Assessment in medical education is complex and includes various stakeholders, each with their own expectations. These stakeholders include students, teachers, lecturers, educational institutions, health care systems, regulatory bodies and patients.[47] A competent health care practitioner who can integrate knowledge, skills and attitudes relevant to the South African context is the ultimate outcome of the outcomes-based medical curriculum. This competency must be observable and measurable to certify the student as competent. Competency is best assessed on the 'Does' level according to Miller's pyramid.[6] The overarching term of workplace-based assessment may be a solution to assess knowledge skills, behaviour, attitude and self-reflection in real-life situations.[48] In spite of the advantages of workplace-based assessment, Miller states that, 'no single assessment method can provide all the data required for judgement of anything as complex as the delivery of professional services by a successful physician'.[6] This is echoed when researchers warned against the use of a single assessment when pass or fail decisions have serious implications, such as for registration or licencing.[49] Assessment should be a continuous process with many data points that can be taken into consideration to make an informed judgement on competence.

The quality of clinical assessment can be improved if attention is given to the following assessment principles:

- Validity: Content validity can be improved with blueprinting of individual as well as overall assessments, and construct validity with appropriate assessment methods.
- Reliability: Reliability can be improved by training assessors, enhancing the quality of questions and mark sheets, and by increasing the number of stations or questions per assessment.
- Fairness: Although all assessment cannot be equal, there should be no discrimination against any student, assessor or patient. It is also important that assessment should be conducted according to the NQF level that the programme is registered for.
- Feasibility: All assessments should be realistic, practical and sensible in the context where they take place. This can be achieved by careful planning and consideration of all resources.
- Educational effect: Assessment should promote learning through study for assessments, or making use of workplace-based assessment and constructive feedback.
- Acceptability: All stakeholders, including patients, students and the educational institution, should be satisfied with the assessment. This can be achieved through transparency and keeping all stakeholders informed.

Quality assurance

Moderation is a quality assurance process that confirms that the assessment is valid and reliable and meets minimum standards.[50] Moderation should form part of the overall assessment in the MBChB programme, as well as of each assessment. Moderation can be conducted internally and/or externally, and should take place before and after assessments. An external moderator should moderate all high-stakes examinations.[40,41] The aim of moderation is to check consistency and standards.[12] Benchmarking is also part of the moderation process as the aim is quality improvement.



MBChB, Bachelor of Medicine and Bachelor of Surgery.

FIGURE 1: Schematic display of the framework to measure the quality of assessment in the clinical phase of an undergraduate medical programme.

Limitations of the study: Although the rapid review was performed according to the protocol, the risk-for-bias and quality of documents were not evaluated by a second reviewer. These results may not be 100% transferable to all MBChB programmes as different universities have different assessment policies and methods.

The complexity of clinical assessment warrants that assessment be 'evaluated on a programmatic level' rather than on individual assessment level, as no individual assessment meets all the criteria of validity, reliability, educational impact, acceptability and cost.[7]

Conclusion

This rapid review attempted to develop a framework to benchmark the quality of assessment in the clinical phase of an undergraduate medical programme. As a first step, all stakeholders should be aware of the outcome of the programme. All assessment and training in the MBChB programme must be aligned with the outcome of the programme, namely, to produce a competent medical practitioner who can integrate knowledge, skills and attitudes relevant to the South African context.

Open-mindedness is essential during the assessment process. Programme accreditation, assessment practices and psychometrics can assist to improve the quality of assessment but cannot judge clinical competence. Experienced assessors, using a variety of assessment methods on a continuous basis, is the proposed way to go. By implementing quality assurance processes, institutions can ensure that specifications and standards are maintained and improved, and that they are globally competitive. It is clear that clinical assessment is multidimensional and that no assessment is perfect. An assessment framework can assist to improve assessment, but it cannot guarantee quality assessment.

Figure 1 is a schematic display of the framework for measuring the quality of assessment in the clinical phase of an undergraduate medical programme.

Acknowledgements

Competing interests

The authors have declared that no competing interests exist.

Author's contributions

H.B. was in charge of conceptualisation of the study, protocol development, data collection and writing of the article. J.B. and L.J.V.d.M. assisted with conceptualisation and planning of the study, as well as critical evaluation and the final approval of the manuscript.

Funding information

This research received no specific grant from any funding agency in the public, commercial or non-for-profit sectors.

Data availability statement

Data sharing is not applicable to this article as no new data were created or analysed in this study.

Disclaimer

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of any affiliated agency of the authors.

References

1. Health Professions Council of South Africa. Accredited facilities [homepage on the Internet]. 2019 [cited 2019 Jul 24]. Available from: https://www.hpcsa.co.za/PBMedicalDental/AccreditedFacilities.
2. South Africa. Health Professions Act 56 of 1974 (Amended 2009). Regulations relating to the registration of students, undergraduate curricula and professional examinations in medicine [homepage on the Internet]. [cited 2019 May 11]. Available from: https://www.hpcsa.co.za/uploads/editor/UserFiles/5.%20REGULATIONS%20RELATING%20TO%20THE%20REGISTRATION%20OF%20STUDENTS,%20UNDERGRADUATE%20CURRICULA%20AND%20PROFESSIONAL%20EXAMINATIONS%20IN%20MEDICINE.pdf.
3. University of the Free State. School of Clinical Medicine Vision and Mission statement [homepage on the Internet]. [cited 2018 Apr 13]. Available from: https://www.ufs.ac.za/health/departments-and-divisions/school-of-medicine-home
4. Basu D. Diseases of public health importance in South Africa. SAJPH. 2018;2(3);48. https://doi.org/10.7196/SHS.2018.v2.i3.72
5. Van der Vleuten CP, Schuwirth LW, Scheele F, Driessen EW, Hodges B. The assessment of professional competence: Building blocks for theory development. Best Pract Res Clin Obstet Gynaecol. 2010;24(6):703–719. https://doi.org/10.1016/j.bpobgyn.2010.04.001
6. Miller GE. The assessment of clinical skills/competence/performance. Acad Med. 1990;65(Suppl 9):S63–S67. https://doi.org/10.1097/00001888-199009000-00045
7. Van der Vleuten CPM. Revisiting 'Assessing professional competence: From methods to programmes'. Med Educ. 2016;50(9):885–888. https://doi.org/10.1111/medu.12632
8. Kis V. Quality assurance in tertiary education: Current practices in OECD countries and a literature review on potential effects [homepage on the Internet]. Organisation for Economic Co-operation and Development. 2005 [cited 2018 Sep 11]. Available from: https://www.oecd.org/education/skills-beyond-school/38006910.pdf.
9. Van der Bank CM, Van der Bank M. Quality assurance in higher education: A case study of the Vaal University of Technology. J Educ Soc Res. 2014;4(1):395–406. https://doi.org/10.5901/jesr.2014.v4n1p395
10. Ingram D. Quality assurance vs. audit [homepage on the Internet]. Small Business – Chron.com. 2019 [2019 Jun 13]. Available from: https://smallbusiness.chron.com/quality-assurance-vs-audit-17455.html.
11. Stapenhurst T. The benchmarking book: A how-to-guide to best practice for managers and practitioners. 2nd ed. Oxford: Elsevier; 2011.
12. Scott R. Benchmarking: A literature review [homepage on the Internet]. Edith Cowan University; 2011 [cited 2019 Aug 12]. Available from: https://intranet.ecu.edu.au/__data/assets/pdf_file/0010/357193.
13. Pangaro L, Ten Cate O. Frameworks for learner assessment in medicine: AMEE Guide No. 78. Med Teach. 2013;35(6):e1197–e1210. https://doi.org/10.3109/0142159X.2013.788789
14. Van Mil JW, Henman M. Terminology, the importance of defining. Int J Clin Pharm. 2016;38(3):709–713. https://doi.org/10.1007/s11096-016-0294-5
15. Mc Call B. How to write a good definition [homepage on the Internet]. 2016 [2019 Apr 15]. Available from: https://www.brightstorm.com/math/geometry/geometry-building-blocks/writing-a-good-definition/.
16. Norcini J, Anderson MB, Bollela V, et al. Draft 2018 Consensus framework for good assessment. Med Teach. 2018;40(11):1102–1109. https://doi.org/10.1080/0142159X.2018.1500016
17. Service Seta. Accreditation [homepage on the Internet]. 2017 [cited 2019 May 7]. Available from: http://www.serviceseta.org.za/index.php/quality-assurance/accreditation.
18. South African Qualifications Authority. National policy and criteria for designing and implementing assessment for NQF qualifications and part-qualifications and professional designations in South Africa [homepage on the Internet]. 2014 [cited 2019 May 15]. Available from: http://www.saqa.org.za/docs/pol/2014/Nationa%20pol%20crite.pdf.
19. University of the Free State. General rules for undergraduate qualifications, postgraduate diplomas, Bachelor Honours degrees, Master's degrees, Doctoral degrees, higher doctorates, honorary degrees and the convocation [homepage on the Internet]. 2019 [2019 May 15]. Available from: https://www.ufs.ac.za/docs/default-source/policy-institutional-documents/2019-general-rules.pdf?sfvrsn=97dcb421_0.

20. Donabedian A. The quality of care. JAMA. 1988;260(12):1743–1748. https://doi.org/10.1001/jama.1988.03410120089033

21. Elshaer I. What is the meaning of quality? [homepage on the Internet]. MPRA Paper No. 57345. 2012 [2019 May 15]. Available from: https://mpra.ub.uni-muenchen.de/57345.

22. University of the Free State. Faculty of Health Sciences. Rule book. School of Clinical Medicine. Undergraduate qualifications [homepage on the Internet]. 2019 [cited 2019 Jun 4]. Available from: https://apps.ufs.ac.za/dl/yearbooks/335_yearbook_eng.pdf.

23. University of the Free State. Assessment policy on the UFS coursework learning programme [homepage on the Internet]. 2016 [cited 2019 May 11]. Available from: https://www.ufs.ac.za/docs/default-source/all-documents/assessment-policy-on-the-ufs-coursework-learningprogrammes.pdf?sfvrsn=3716c321_0.

24. Health Professions Council of South Africa. Core competencies for undergraduate students in clinical associate, dentistry and medical teaching and learning programmes in South Africa [homepage on the Internet]. 2014 [cited 2018 Mar 12]. Available from: http://www.hpcsa.co.za/uploads/editor/UserFiles/downloads/medical_dental/MDB%20Core%20Competencies%20-%20ENGLISH%20-%20FINAL%202014.pdf.

25. World Federation for Medical Education. Standards. Basic medical education [homepage on the Internet]. 2015 [cited 2019 May 11]. Available from: https://wfme.org/standards/bme/.

26. Haby MM, Chapman E, Clark R, Barreto J, Reveiz L, Lavis JN. What are the best methodologies for rapid reviews of the research evidence for evidence-informed decision making in health policy and practice: A rapid review. Health Res Policy Syst. 2016;14(1):83. https://doi.org/10.1186/s12961-016-0155-7

27. Tricco AC, Antony J, Zarin W, et al. A scoping review of rapid review methods. BMC Med. 2015 Dec;13(1):224. https://doi.org/10.1186/s12916-015-0465-6

28. Khangura S, Konnyu K, Cushman R, Grimshaw J, Moher D. Evidence summaries: The evolution of a rapid review approach. Syst Rev. 2012;1(1):10. https://doi.org/10.1186/2046-4053-1-10

29. Tricco AC, Langlois EV, Straus SE, editors. Rapid reviews to strengthen health policy and systems: A practical guide. Geneva: World Health Organization; 2017. Licence: CC BY-NC-SA 3.0 IGO.

30. South Africa. Council on Higher Education. Higher Education Act 101 of 1997 [homepage on the Internet]. [cited 2019 Jun 4]. Available from: https://www.gov.za/sites/default/files/gcis_document/201409/a101-97.pdf.

31. South Africa. Health Professions Act 56 of 1974 (Amended 2007). Education, training and registration [homepage on the Internet]. [cited 2019 Apr 15]. Available from: https://www.gov.za/sites/default/files/gcis_document/201409/a29-07.pdf.

32. Health Professions Council of South Africa. Professional boards [homepage on the Internet]. 2019 [cited 2019 Jun 4]. Available from: https://www.hpcsa.co.za/PBMedicalDental.

33. South African Qualifications Authority. SAQA Act 58 of 1995 [homepage on the Internet]. [cited 2019 May 15]. Available from: http://www.saqa.org.za/show.php?id=5469.

34. South African Qualifications Authority. The National Qualifications Framework curriculum development [homepage on the Internet]. 2000 [cited 2019 May 15]. Available from: http://www.saqa.org.za/docs/pol/2000/curriculum_dev.pdf.

35. SAQA. National qualifications framework and the standards setting [homepage on the Internet]. 2003 [cited 2019 June 04] Available from: http://saqa.org.za/docs/pol/2003/standard_setting.pdf.

36. South African Qualifications Authority. Criteria and guidelines for assessment of NQR registered unit standards and qualifications [homepage on the Internet]. 2001 [cited 2019 May 15]. Available from: http://cdn.lgseta.co.za/resources/guidelines/2.4.1%20SAQA%20Criteria%20and%20Guidelines%20for%20Assessment.pdf.

37. South African Qualifications Authority. Guidelines for integrated assessment [homepage on the Internet]. 2005 [cited 2019 May 15]. Available from: http://www.saqa.org.za/docs/guide/2005/intassessment.pdf.

38. South Africa. National Qualifications Framework. National Qualifications Framework Act 67 of 2008 [homepage on the Internet]. [cited 2019 May 15]. Available from: https://www.gov.za/documents/national-qualifications-framework-act.

39. University of the Free State. Teaching-learning policy [homepage on the Internet]. 2008 [cited 2019 May 15]. Available from: https://www.ufs.ac.za/docs/default-source/all-documents/teaching-learning-policy-120eng.pdf?sfvrsn=b22de421_0.

40. University of the Free State. Quality assurance policy [homepage on the Internet]. 2009 [cited 2019 May 15]. Available from: https://www.ufs.ac.za/about-the-ufs/governance/policy-documents.

41. University of the Free State. Guidelines for the implementation of external moderation [homepage on the Internet]. 2009 [cited 2019 May 15]. Available from: https://www.ufs.ac.za/docs/default-source/all-documents/guidelines-for-the-implementation-of-external-moderation-404-eng.pdf?sfvrsn=2e2be421_0.

42. University of the Free State. School of Clinical Medicine. Undergraduate programme management. SOP quality assurance. Bloemfontein: UFS print; 2019.

43. World Federation for Medical Education. Accreditation [homepage on the Internet]. 2017 [cited 2019 May 15]. Available from: https://wfme.org/accreditation/.

44. Tavakol M, Dennick R. The foundations for measurement of assessment in Medical Education. AMEE Guide 119. Med Teach. 2017;39(10):1010–1015. https://doi.org/10.1080/0142159X.2017.1359521

45. World Health Organization. Transforming and scaling up health professional education and training [homepage on the Internet]. Policy Brief on Accreditation of Institutions for Health Professional Education. World Health Organization; 2013 [cited 2019 May 15]. Available from: https://whoeducationguidelines.org/sites/default/files/uploads/whoeduguidelines_PolicyBrief_Accreditation.pdf.

46. Heath Professions Council of South Africa. Regulations relating to the inspection of the evaluation and the teaching of medical and dental students in the republic of South Africa – GN R2270/76 [homepage on the Internet]. 2019 [2019 May 15]. Available from: https://hpcsa.co.za/Uploads/MDB/Rules%20%26%20Regulations/regulations_gnr_2270_76.pdf.

47. 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. https://doi.org/10.3109/0142159X.2011.551559

48. Govaerts M, van der Vleuten CP. Validity in work-based assessment: Expanding our horizons. Med Educ. 2013 Dec;47(12):1164–1174. https://doi.org/10.1111/medu.12289

49. Van der Vleuten C, Sluijsmans D, Joosten-ten Brinke D. Competence assessment as learner support in education. In: Mulder M, editor. Competence-based vocational and professional education. Springer International Publishing, Cham, 2017; p. 607–630.

CHAPTER 3:

ARTICLE 2 – STUDENTS' VOICES: ASSESSMENT IN UNDERGRADUATE CLINICAL MEDICINE

The article was prepared according to the submission guidelines of the *Pan African Medical Journal*.

Journal:	Pan African Medical Journal
Publisher:	The Manuscript Hut
Country of publication:	Kenya
ISSN:	1936-8688 (ONLINE)
DOI:	10.11604/pamj.2020.36.130.22168

Citation:

Brits, H., Bezuidenhout, J., Van der Merwe L.J. & Joubert, G. Students' voices: assessment in undergraduate clinical medicine. Pan African Medical Journal. 2020;36:130. DOI: 10.11604/pamj.2020.36.130.22168

Submitted:	4 March 2020
Accepted:	12 May 2020
Published:	25 June 2020


The Pan African Medical Journal
May - August, 2020
Volume 36
ISSN: 1937 – 8688
An Open Access Journal
PAMJ.ORG
Unleashing African Research
PanAfrican Medical Journal

Overview

ISSN

1937-8688 (ONLINE)

Focus and scope

We believe that scientific work done in Africa should be rapidly and freely made available to all researchers worldwide. Aim: To create, stimulate and perpetuate a culture of information sharing and publishing amongst researchers and other health actors of the African health scene in ways that will contribute to availability of health information, better understanding of Africa specificities and overall, to improve health outcomes of the people on the continent. Scope: We publish original scientific research on clinical, public health, social, political, economic and all other factors affecting the health of populations in Africa.

Currently Tracked

- DOAJ
- AIM
- Google Scholar
- AJOL
- EBSCO
- Scopus
- Embase
- IC
- HINARI
- Global Health
- PubMed Central
- PubMed/Medline
- ESCI

Managing Editors

Article

PanAfrican Medical Journal

Research

Students' voices: assessment in undergraduate clinical medicine

Hanneke Brits, Johan Bezuidenhout, Lynette Jean van der Merwe, Gina Joubert

Corresponding author: Hanneke Brits, Department of Family Medicine, School of Clinical Medicine, University of the Free State, Bloemfontein, South Africa. britsh@ufs.ac.za

Received: 04 Mar 2020 - **Accepted:** 12 May 2020 - **Published:** 25 Jun 2020

Keywords: Quality assessment, student opinions, clinical medicine, undergraduate medicine

Cite this article: Hanneke Brits et al. Students' voices: assessment in undergraduate clinical medicine. Pan African Medical Journal. 2020;36(130). 10.11604/pamj.2020.36.130.22168

Available online at: https://www.panafrican-med-journal.com//content/article/36/130/full

Students' voices: assessment in undergraduate clinical medicine

Hanneke Brits[1,&], Johan Bezuidenhout[2], Lynette Jean van der Merwe[3], Gina Joubert[4]

[1]Department of Family Medicine, School of Clinical Medicine, University of the Free State, Bloemfontein, South Africa, [2]Health Sciences Education, Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa, [3]Undergraduate Programme Management, School of Clinical Medicine, University of the Free State, Bloemfontein, South Africa, [4]Department of Biostatistics, Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa

[&]**Corresponding author**

Hanneke Brits, Department of Family Medicine, School of Clinical Medicine, University of the Free State, Bloemfontein, South Africa

Article

PanAfrican Medical Journal

Abstract

Introduction: *the perception exists among students that not all clinical assessments in undergraduate medical programmes are of high quality. 'Student voice' is a term used to describe how students feel about and experience their education in a safe and controlled environment. This study aimed to investigate the opinions and experiences of medical students at the University of the Free State on the quality of assessment in the clinical phase of medicine.*

Methods: *a cross-sectional study design was used. Quantitative data were collected with space to clarify opinions and make recommendations. The study population consisted of the clinical medical students in 2019 who had completed at least one module and one end-of-year assessment. Self-administered, anonymous questionnaires were distributed to obtain opinions and experiences regarding assessment. Questions in the questionnaire derived from an assessment framework for clinical medicine to ensure construct and content validity.*

Results: *one hundred and ninety-two (192) students completed questionnaires (84.6% response rate). Less than half of the students were of the opinion that the assessments were fair, with lack of blueprinting and incorrect level of assessment major contributors to this opinion. Two thirds believed that the assessment was aligned with outcomes, however training was not aligned with the assessment. More than 90% of students reported on the lack of feedback after assessment. Valuable suggestions from the students included ways of assessing professionalism, timing of assessments and training of assessors.*

Conclusion: *majority of students were of the opinion that there is room for improvement in the quality of assessment.*

Introduction

The cornerstones of good quality assessment are validity, reliability, fairness, feasibility, educational effect, acceptability, assessment of higher cognitive skills and benchmarking [1, 2]. It is important to be able to defend the quality of assessment in certifying courses e.g. medicine. In addition, students can give valuable information about their assessment experiences that may contribute to better assessment [3]. 'Student voice' is a term that describes the way students express how they feel and experience different aspects of education in a safe and controlled environment [4]. Youens and Hall [5] state that, 'instead of treating students as voices crying in the wilderness, we would be far better served if we asked the voices' owners what they think and listened actively to the answers'. Student opinions can be gathered through open forums, feedback slips, surveys and formal questionnaires [6]. Students' opinions should be seen as viewpoints while landmark events, relationships and students' gut reactions must be taken into consideration when these opinions are interpreted [7]. One experience (landmark event) may influence the way that a student responds to questions. For example if a student did not have enough time to complete the task at one of the 10 objective structured clinical examination (OSCE) stations the student passed through, the student may respond by stating that the time permitted for each OSCE station is insufficient [7]. Relationships, whether power relationships or personal relationships, between students and assessors may have a positive or negative effect on the way students express their opinions. Gee (2017) reports a student commenting: 'I try to put what I feel is true and needs to be said, rather than worrying about what the people will think when they read it' [7]. However, sometimes participants do not give any thought to their answers to questions and just complete the questionnaire as quickly as possible [7].

Several University of the Free State (UFS) undergraduate medical students who participated in the 2017 and 2018 end-of-year assessment expressed dissatisfaction with the marks they had obtained in clinical medicine (Personal communication with students). The phenomenon of student dissatisfaction with marks is not unique to the UFS, and is well described in the literature [8]. The motivation for this study was the perception that not all assessments in the clinical phase of the undergraduate medical programme at the UFS are of high quality. The clinical phase (Phase III) takes place over the last two and a half years of study, and entails rotations through various clinical departments in multiple health facilities. Clinical training takes place according to a fixed programme and for fixed periods. Formative assessment takes place throughout the clinical rotation and includes an end-of-block assessment. Together, these assessments contribute to a block mark (module mark), which gives students access to the final, summative, end-of-year assessment. Assessment in the clinical phase consists of theoretical assessment (multiple-choice questions and written questions) as well as practical assessments (long and short clinical cases and objective structured clinical assessments). This study aimed to investigate the opinions and experiences of students regarding the quality of the assessment they experienced in the clinical phase of the undergraduate medical programme.

Methods

A cross-sectional study design was used. Quantitative data were collected with space to clarify opinions. The study population consisted of all the students who had completed at least one module and one end-of-year assessment in the clinical phase of the undergraduate medical programme. All 227 students in the fourth and fifth years of the undergraduate medical programme at the UFS during 2019 were included in the study sample. The steps of questionnaire development, as described by Katzenellenbogen and Joubert, [9] were used to design the questionnaire. Questions included in the questionnaire derived from a framework to benchmark the quality of assessment in undergraduate clinical medicine [10]. Questions were grouped under headings, to structure the responses. These headings were principles of quality assessment (fairness, validity, reliability and educational effect), assessment methods, assessment of soft skills and recommendations on how to improve assessments. Responses were mainly yes or no followed by a question to justify or expand on the responses. A question on how to improve assessment was also included. The input of the co-authors was used to improve the validity and reliability of the questionnaire. For internal validity, the face validity of the questionnaires was tested in a pilot study conducted on five interns who had completed their studies at the UFS in 2018. No items were changed, but the layout was adjusted to improve flow. The questions in the questionnaires derived from the literature, to ensure construct and content validity. The researcher tried to be as objective as possible during data interpretation and representation. Numbers were double checked and responses were quoted verbatim under each heading.

McMillan and Schumacher [11] describe questionnaires as research instruments that can be used to gather information on the current statuses of a situation from a specific population. A self-administered, anonymous questionnaire was used for data collection, as it posed less of a threat of exposure to participants [12]. After the purpose of the questionnaire had been explained to the participants, student group leaders distributed the printed questionnaires and information leaflets to the students. After completion the questionnaires were returned to the group leaders and then to the researcher. Data was transferred to Excel data sheets by the first author twice, to check for integrity. Percentages were calculated for different responses per year group. Responses and the justification for answers were grouped by the first author according to question headings.

Ethical considerations: ethical approval for the study was obtained from the Health Sciences Research Ethics Committee, UFS (UFS-HSD 2019/0001/2304). UFS authorities approved the inclusion of medical students. Participation was voluntary, with implied consent being given by completing the anonymous questionnaires. Participants did not receive any compensation for completing the questionnaires and there was no penalty for not participating in the study.

Results

A total of 108 out of 119 fourth-year students (90.6%), and 84 out of 106 fifth-year students (79.2%) returned completed questionnaires. More than 80% of students gave justifications for responses and 75.5% made recommendations to improve assessment practices.

Quality of assessment: only 43.5% of fourth-year and 44.1% of fifth-year medical students felt that the current assessment is fair. In Table 1 the percentages of students who agreed with the statements are displayed. Open responses were not split for different year groups. About half the students clarified their responses in relation to the alignment of questions with outcomes in tests and exams; the following are some of the comments students made: 'The outcome is in the book, but they don't train that and then they ask that.' 'Because we spend too little time with patients, we don't see all the things that they ask in clinical cases, although it is mentioned in the module guides.' 'We're not assessed on what we see commonly.' 'We know the outcomes from our module guides.' Many students had opinions regarding the spread of questions and the following comments were made: 'You know who compiled the questions, because they only concentrate on their own field of expertise.' 'We need more assessments, to cover more work.' 'Some subjects ask 10 to 15 marks from 20 chapters, it's not fair.' 'With MCQs (*multiple choice questions*) you get a good spread of questions, but they are not necessarily important.' 'It is dependent on the subject, some do it better than others. OSCEs help.' More than 80% of students commented that the level of assessment was on specialist level, rather than on general practitioner level. The following were comments regarding the difficulty of questions: 'We're not specialists, don't assess us as specialists.' 'No need for specialist special investigations to be tested.' 'They don't assess hard work and professionalism.' 'They expose us to specialist work and then they ask that, we're not specialists.' 'Sometimes you're lucky and get common things.'

Assessment methods: regarding the appropriateness of assessment methods, most students commented negatively about multiple-choice questions as an assessment method. 'They (*multiple-choice questions*) don't test knowledge.' 'Questions not up to standard, they use the same old questions and things changed.' 'MCQs can't test arguments.' 'You get your marks quickly, but don't know the correct answers.' A third of students made recommendations about assessment methods. The following recommendations were made: 'Improvement is necessary, but I don't know how.' 'Please include short questions, so that I can explain.' 'We should get exposure to the exam assessments during rotations and block assessments.' 'Written OSCES are not clear and not clinical, don't use them.' Ten responses were received regarding assessors and they were all negative. Some of these were: 'Train the examiners.' 'Your marks are dependent on the examiner, they like you or they don't. It's not fair if they don't like you.' 'Some people give very poor marks, they think that 65% is good.' According to the students, feedback is almost non-existent. Table 2 indicates percentages of students of the two years who agreed with certain statements. More than 75% of students commented on feedback, or the lack of feedback. Most students asked for constructive feedback to assist with learning. Some comments were the following: 'Formal feedback sessions should be scheduled like in Phases 1 and 2'. 'Immediate feedback after clinical cases will help a lot. Then you know how to improve.' 'By the time that you get your marks,

you can't even remember what they asked and what you answered.' 'Please fill in an assessment form for clinical assessments and give a copy to the student after assessment. Then the student will see how to improve (*and*) the lecturer how to give marks.

Assessment of soft skills: almost all students, 85.2% of fourth years and 92.9% of fifth years, agreed that soft skills, such as professionalism, should be assessed. Practical suggestions for assessment of these skills included using cellular phone applications to assess group members weekly, asking patients and other health care professionals for feedback, and that the responsible registrar provides feedback after the rotation. Twenty percent of students commented on the positive or negative effects that role models could have on professional behaviour. Most value good role models, but they could also learn from bad role models, as indicated by this response: 'At least we can see how not to behave, if we look at some people.'

Competence and assessment: most students thought that assessment results were not an indication of competence. Table 3 displays these results. Only 48.1% of fourth years, compared to 78.6% of fifth-year students, felt confident about their skills even if they had passed an assessment. Many students mentioned the lack of feedback after assessment as a reason for their lack of confidence. Another reason mentioned was that the marks that they had obtained were not necessarily an indication of their competence. Some comments in this regard were: 'The doctor said that I sucked and then gave me 65%. For me that is good.' 'They said that I did well, but I only just passed.' Comments in relation to this topic included: 'You should know from your block assessments if you will pass, but it's not so.' 'Some students are good with exams, but not competent.' 'If I pass I'm good, not sure about others.' 'If I pass they (*the assessments*) are good, if I fail they are not.'

Recommendations to improve clinical assessment: students made recommendations to improve the type of assessments, examiners, the examination process and the content of the assessment, as well as general recommendations.

The type of assessment: students recommended continuous clinical assessment, next to patients or in clinical areas, with immediate feedback. More than 70% of students indicated that multiple-choice questions alone are not enough to test knowledge, and they wanted short questions, or short questions in addition to multiple-choice questions. Students felt strongly that end-of-block assessments should be sufficient to test competence, and that they do not need to do a final assessment at the end of the year, on the same subjects again. Students also recommended that assessment during the clinical blocks should promote them to an integrated assessment at the end of the final year. 'Give us one MIMA (*integrated medical assessment*) OSCE at the end of the year, like in second and third year.'

Examiners: only a few students made recommendations regarding examiners, among which, 'Train the examiners, some are clueless.' 'Some examiners should not examine, first examine them.' 'Use external examiners from other departments.' 'We need better role models as examiners.'

The examination process: many students recommended that continuous assessment takes place while patients are presented in clinical areas, and written feedback given. Another recommendation was to record presentations, which may assist students and examiners to clarify areas of disagreement or improvement. Students requested exposure to mock assessments, to prepare them for and to assist them in their preparation for final assessments. The timing of assessments is important: it should not be after long calls, or late in the day, when students and examiners are tired.

Content of assessments: about 30% of students believed that their training was not aligned with the assessment. 'We must spend more time in the clinics and wards, then we will learn more and do better in exams.' In contrast, students also commented as follows: 'They ask what is easy to mark, not what they teach us.' 'They use the same patient, and it's not fair if you are the first or last student.' 'Some patients are good to use in exams, but it's not really what we need to know.' Most students indicated that assessment should cover general conditions, and not specialist or super-specialist conditions. Some students also provided general recommendations about how better training may affect assessment. 'Everybody must train the same facts from prescribed references.' 'Different disciplines do things differently, they should have a uniform format, for instance when presenting patients.' 'Tutorials and discussions are much better than lectures. We can read better than some lecturers.' 'Use the good role models to train other lecturers.'

Discussion

The response rate of a survey matters, and a good response rate indicates that the researcher can generalize the results for the population under investigation [3]. The response rate of 85% in this survey makes the results obtained generalizable, as they represent the opinions of current clinical medical students at the UFS. It may also indicate that students were eager to voice their opinions. Less than half the students believed that current assessments were fair. They thought that the questions were too difficult for undergraduate students, and they were not satisfied with the spread of questions (blueprinting). This opinion of the students may be because specialists and super specialists conduct training in the undergraduate and postgraduate medical programme, and may find it difficult to accept that undergraduate students need only limited knowledge of 'their' subject. McConlogue [13], Price *et al.* [14] and Yorke [15] describe how complex it is to set exams when different dimensions must be considered. In summative assessment, blueprinting is problematic, because of the large volumes of work that must be covered by a single assessment; this is even more complex when clinical assessment involves patients and different assessment sites. Sites may differ regarding resources and disease profiles, and the conditions of patients may also change during or between assessments. Although >90% of students reported not receiving feedback after assessments, they value feedback and gave good suggestions to improve the lack of feedback. The value of formative feedback is that it assists students to measure themselves and 'identify the gap between their current and desired performance' [16]. To get the maximum advantage of feedback, it should be of good quality, specific and on time [17].

The Health Professions Council of South Africa describes core competencies for undergraduate medical students, including professionalism [18]. Most students wanted a formal assessment of 'soft skills', and suggested peer and patient assessment. It is concerning that they do not consider all lecturers and doctors to be good role models of professional behaviour. Less than half of fourth years, and more than three quarters of fifth-year students were confident about their competence after passing an assessment. In the fifth year group, this confidence may be the result of more experience with clinical assessment, and being better prepared for assessment. Most students expressed that assessment without feedback does not provide a good indication of competence or incompetence. An interesting finding is that students were satisfied with the assessment if they passed, but not when they had failed. The students also expressed different views on assessments in which they were involved and assessments involving others. A study in the United States found that medical students were able assess their own knowledge, skills and behaviour on assessment accurately, with fewer than 10% of students overrating their knowledge [19]. However, students could not accurately evaluate others, and, therefore, student assessments lack validity and reliability

[20-22]. Students provided a number of recommendations on how the current assessment could be improved. A limitation of this study is that it reports opinions and experiences, rather than facts which are noted. It is also possible that despite the pilot study some students did not understand the terms used in the questionnaire, nevertheless, they answered all the questions.

Conclusion

The students provided valuable feedback on their experiences of the current assessment in clinical medicine. Fifth-year students were more satisfied with their assessment than fourth years. Students had different and sometimes contrasting opinions on assessment of themselves and others. These results and the recommendations made by students will be discussed at appropriate forums, with the aim of improving the quality of assessment. Students will also receive feedback on this research to encourage transparency of the process. Majority of students were of the opinion that there is room for improvement in the quality of assessment.

What is known about this topic

- *The quality of clinical assessment should be established and defendable;*
- *Student opinions are valuable and reliable sources of information.*

What this study adds

- *Misalignment between outcomes, training and assessment;*
- *Recommendations on assessment methodology e.g. integrated assessments;*
- *Students value feedback.*

Competing interests

The authors declare no competing interests.

Authors' contributions

HB: conceptualization of study, protocol development, data collection and writing of paper, JB and LJvdM: promotors who assisted with conceptualization and planning of the study, as well as critical evaluation and final approval of the manuscript. GJ assisted with concept and methodology, performed data analysis and assisted with interpretation and write-up. All the authors read and approved the final version of the manuscript.

Tables

Table 1: percentages of fourth and fifth year students who answered yes to some of the questions on the quality of assessments
Table 2: percentage of fourth and fifth year students who answered yes to questions on feedback
Table 3: percentage of fourth and fifth year students who answered yes to questions on competence and assessment

References

1. Cruess RL, Cruess SR, Steinert Y. Amending Miller's pyramid to include professional identity formation. Academic Medicine. 2016;91: 180-185. **PubMed| Google Scholar**
2. Norcini J, Friedman Ben-David M. Concepts in Assessment. In: Dent JA, Harden RM, eds. A Practical Guide for Medical Teachers. London: Elsevier. 2013; 4th ed: 285-291.
3. Stark PB, Freishtat R. An evaluation of course evaluations. Science Open Research. 2014. doi: 10.14293/S2199-1006.1.SOR-EDU.AOFRQA.v1.
4. Cook-Sather E. Sound presence and power: 'Students voice' in the educational research and reform. Curriculum Inquiry. 2006;36(4): 359-390. **Google Scholar**

5. Youens B, Hall C. Incorporating pupil perspectives in initial teacher education-lessons from the Pupil Mentoring Project. Teacher Development. 2006;10(2): 197-206. **Google Scholar**
6. Education Council of Aotearoa New Zealand. Using student voice in teacher appraisal. 2018. Accessed on 1st September 2019.
7. Gee N. A study of student completion strategies in a Likert-type course evaluation survey. Journal of Higher Education. 2017;41(3): 340-350. **Google Scholar**
8. Patil SY, Gosavi M, Bannur HB, Ratnakar A. Blueprinting in assessment: A tool to increase the validity of undergraduate written examinations in pathology. International Journal of Applied and Basic Medical Research. 2015;5(S1): S76-S79. **PubMed| Google Scholar**
9. Katzenellenbogen J, Joubert G. Data Collection and Measurement. In: Ehrlich R, Joubert G, eds. Epidemiology - A Research Manual for South Africa. Cape Town: Oxford University Press. 2014;3rd ed: 111-116.
10. Brits H, Bezuidenhout J, Van der Merwe LJ. A framework to benchmark the quality of clinical assessment in a South African undergraduate medical programme. S Afr Fam Pract. 2020;62(1): a5030. **PubMed| Google Scholar**
11. McMillan JH, Schumacher S. Research in education: a conceptual introduction. New York: Open University Press. 2001; 5th ed.
12. Cohen L, Manion K, Morrison K. Research methods in education. New York: Taylor and Francis Group, Routledge. 2007; 6th ed.
13. McConlogue T. 'But is it fair?' Developing students' understanding of grading complex written work through peer assessment. Assessment and Evaluation in Higher Education. 2012;37(1): 113-123.
14. Price M, Handley K, Millar J, O'Donovan B. Feedback: All that effort, but what is the effect? Assessment and Evaluation in Higher Education. 2010;35(3): 277-289. **Google Scholar**
15. Yorke, M. Formative assessment in higher education: Moves towards theory and the enhancement of pedagogical practice. Higher Education. 2003;45(4): 477-501. **Google Scholar**
16. Wingate U. The impact of formative feedback on the development of academic writing. Assessment and Evaluation in Higher Education. 2010;35(5): 519-533. **Google Scholar**
17. McSweeney F. Students' views on assessment. Other resources. 2014;12. **Google Scholar**
18. Health Professions Council of South Africa. Core competencies for undergraduate students in clinical associate, dentistry and medical teaching and learning programmes in South Africa. Pretoria. 2014.
19. Torres MB, Cochran A. Accuracy and content of medical student mid-clerkship self-evaluations. Am J Surg. 2016;211(6): 1153-7. **PubMed| Google Scholar**
20. Hornstein HA. Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. Cogent Education. 2017;4: 1-8. **Google Scholar**
21. Boring A, Ottoboni K, Stark PB. Student evaluations of teaching (mostly) do not measure teaching effectiveness. Science Open Research. 2016. **Google Scholar**
22. Clayson DE. Student evaluation of teaching and matters of reliability. Assessment & Evaluation in Higher Education. 2018;43(4): 666-681. **Google Scholar**

Table 1: percentages of fourth and fifth year students who answered yes to some of the questions on the quality of assessments		
STATEMENT	**4th year**	**5th year**
Do you think that the questions in tests/exams are aligned with the outcomes of the programme?	63.0%	67.9%
Do you think that the questions are spread to cover all work? (Blueprinting)	35.2%	45.2%
Do you think that there is a good spread between easy and difficult questions? (Bloom´s taxonomy)	52.8%	67.9%
Do you think that appropriate assessment methods are used in tests/exams?	66.7%	67.9%

Table 2: percentage of fourth and fifth year students who answered yes to questions on feedback		
STATEMENT	**4th year**	**5th year**
Assessment results are available within 2 weeks of the assessment	35.2%	24.1%
Memorandums are available for assessments	2.8%	7.1%
We receive formal feedback after assessments	7.4%	9.5%

Table 3: percentage of fourth and fifth year students who answered yes to questions on competence and assessment		
STATEMENT	**4th year**	**5th year**
My end-of-block assessment mark is a good predictor of my end-of-year assessment mark.	26.9%	42.9%
Students who pass the final assessment in 5th year are competent to become interns.	38.9%	41.7%
Students who fail the final assessment in 5th year are not competent to become interns and need more training.	15.7%	8.3%

Student Questionnaire

Assessment in the clinical phase of the MBChB programme

1. Year of study 4th 5th

2. Regarding assessment in the clinical phase of the MBChB programme

2.1 Do you think that the current assessment is fair?	Y	N
2.2 Do you think that questions are spread to cover all work? (Blueprinting)	Y	N
2.3 Do you think that the questions in the tests/exams are aligned with outcomes of the programme?	Y	N
2.4 Do you think that there is a good spread between easy and difficult questions?	Y	N
Any further comments regarding the above		

3. Regarding the assessment methods used

3.1 Do you think that appropriate assessment methods are used in tests/exams?	Y	N
3.2 Do you think that different assessment methods should be used?	Y	N
Any further comments regarding the above		

4. Prediction of marks

4.1 Are you usually able to predict the marks you will obtain after an assessment?	Y	N
4.2 Are your end of block assessment marks a good predictor of end of year marks?	Y	N
Any further comments regarding the above		

5. Timelines following assessment opportunities:

5.1 Are assessment results available within two weeks after the assessment opportunity?	Y	N
5.2 Are memorandums for assessments made available?	Y	N
5.3 Do you receive any formal feedback after assessments ?	Y	N
5.4 Do you think that you learn from assessments?	Y	N
Any further comments regarding the above		

6. Competencies/ Soft skills

6.1 After you have passed an assessment, do you feel competent in that specific area/skill?	Y	N
6.2 Do you feel competent in related skills not tested?	Y	N
6.3 Do you think that Professionalism is an important skill to be assessed?	Y	N
Any further comments regarding the above		

7. How do you think clinical assessment can be improved?

.....

.....

.....

.....

8. How would you suggest that "soft skills" (e.g. communication, teamwork, professionalism) should be assessed?

.....

.....

.....

.....

9. Do you think that all students who pass the exams are competent to become interns?	Y	N
10. Do you think that all students who fail are not competent to be an intern and need further training?	Y	N
Any further comments regarding the above		

Any other comments regarding assessment

Thank you!!

Please note: By completing this questionnaire you give consent that the provided information may be used for the research project.

CHAPTER 4:

ARTICLE 3 – ASSESSMENT PRACTICES IN UNDERGRADUATE CLINICAL MEDICINE TRAINING: WHAT DO WE DO AND HOW WE CAN IMPROVE?

The article was prepared according to the submission guidelines of the *African Journal of Primary Health Care and Family Medicine*.

Journal:	African Journal of Primary Health Care and Family Medicine
Publisher:	AOSIS Publishing – African Online Scientific Information Systems
Country of publication:	South Africa
ISSN:	2071-2928 (PRINT), 2071-2936 (ONLINE)
DOI:	10.4102/phcfm.v12i1.2341

Citation:
Brits H, Bezuidenhout J, Van der Merwe LJ, Joubert G. Assessment practices in undergraduate clinical medicine training: What do we do and how can we improve? Afr J Prm Health Care Fam Med. 2020;12(1), a2341. DOI: 10.4102/phcfm.v12i1.2341

Submitted:	6 January 2020
Accepted:	18 May 2020
Published:	10 July 2020

AOSIS Publishing

PHCFM | AFRICAN JOURNAL OF PRIMARY HEALTHCARE & FAMILY MEDICINE

ISSN: 2071-2928

2020

AOSIS

Open access at phcfm.org

Vol. 12 No. 1

Overview

ISSN

2071-2928 (PRINT)
2071-2936 (ONLINE)

Focus and scope

The *African Journal of Primary Health Care & Family Medicine* (PHCFM), launched in 2008, is the official journal of WONCA (World Organization of Family Doctors) Africa Region. It provides a platform for scholarly exchange between family medicine and primary health care researchers and practitioners across Africa. It provides a contextual and holistic view of family medicine and primary health care as practised across the continent. The journal is indispensable to primary health care practitioners, family medicine specialists and academics with an interest in the African continent, and offers an engaging insight into the growth of these disciplines from a distinctly African perspective. PHCFM seeks to publish innovative research, reviews, country profiles, editorials and opinion pieces on all aspects of primary health care and family medicine in the African context. The research includes primary care, family practice, sexual health, oral health, emergency care, district health services, rural health care, health promotion, prevention of disease and disability, pain management and palliative care, community-orientated primary care, and the education and training of professionals and health workers in family medicine and primary health care.

DHET Accreditation

The journal is DHET accredited because it is listed on the following approved indexing services:

- SciELO SA
- SCOPUS

Editor-in-Chief

Assistant Editors

African Journal of Primary Health Care & Family Medicine
ISSN: (Online) 2071-2936, (Print) 2071-2928

AOSIS

Page 1 of 7 Original Research

Assessment practices in undergraduate clinical medicine training: What do we do and how can we improve?

Authors:
Hanneke Brits[1]
Johan Bezuidenhout[2]
Lynette J. van der Merwe[3]
Gina Joubert[4]

Affiliations:
[1]Department of Family Medicine, School of Clinical Medicine, Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa

[2]Division of Health Sciences Education, Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa

[3]Undergraduate Programme Management, School of Clinical Medicine, Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa

[4]Department of Biostatistics, Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa

Corresponding author:
Hanneke Brits,
britsh@ufs.ac.za

Dates:
Received: 06 Jan. 2020
Accepted: 18 May 2020
Published: 06 July 2020

How to cite this article:
Brits H, Bezuidenhout J, Van der Merwe LJ, Joubert G. Assessment practices in undergraduate clinical medicine training: What do we do and how can we improve?. Afr J Prm Health Care Fam Med. 2020;12(1), a2341. https://doi.org/10.4102/phcfm.v12i1.2341

Background: Assessment should form an integral part of curriculum design in higher education and should be robust enough to ensure clinical competence.

Aim: This article reports on current assessment practices and makes recommendations to improve clinical assessment in the undergraduate medical programme at the University of the Free State.

Methods: A descriptive cross-sectional study design was used. Qualitative and quantitative data were gathered by means of open- and closed-ended questions in a self-administered questionnaire, which was completed by teaching and learning coordinators in 13 disciplines.

Results: All disciplines in the undergraduate medical programme are represented. They used different assessment methods to assess the competencies required of entry-level healthcare professionals. Workplace-based assessment was performed by 30.1% of disciplines, while multiple-choice questions (MCQs) (76.9%) and objective structured clinical examinations (OSCEs) (53.6%) were the main methods used during formative assessment. Not all assessors were well prepared for assessment, with 38.5% never having received any formal training on assessment. Few disciplines (15.4%) made use of post-assessment moderation as a standard practice, and few disciplines always gave feedback after assessments.

Conclusion: The current assessment practices for clinical students in the undergraduate medical programme at the University of the Free State cover the spectrum that is necessary to assess all the different competencies required. Multiple-choice questions and OSCEs, which are valid and reliable assessment methods, are used frequently. Poor feedback and moderation practices should be addressed. More formative assessments, and less emphasis on summative assessment, should be considered. Workplace-based and continuous assessments may be good ways to assess clinical competence.

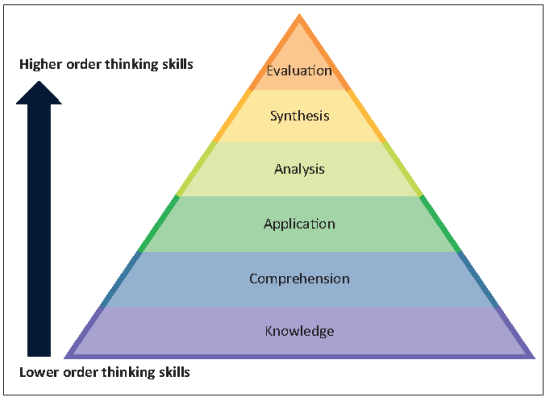
Keywords: assessment practices; clinical competence; improvement; undergraduate; South Africa.

Background

Assessment should form an integral part of curriculum design in higher education.[1] Biggs explains that the outcomes of a programme, training and assessment should complement each other.[2]

The South African Qualifications Authority provides principles for credible assessment, among which are validity, reliability, fairness and practicability.[3] Blueprinting is another important component of assessment, and ensures the reliability and validity of assessments.[4,5] An assessment blueprint is a detailed plan (or table) of what is covered in the assessment.[4] A blueprint should form part of the overall assessment planning and should include the content and cognitive levels that will be covered in the assessment process.[4] The cognitive levels include knowledge, comprehension, application, analysis, synthesis and evaluation.[6] These original levels as described by Blooms et al. are displayed in Figure 1.

The validity of assessments can be addressed using appropriate assessment methods and tools.[7] Reliability is influenced by the quality and number of markers and questions, as well as the quality of assessment rubrics.[8,9] To ensure that an assessment is practically feasible, resources,



Source: Bloom MS, Krathwohl DR. Taxonomy of educational objectives: The classification of educational goals, by a committee of college and university examiners. In: Handbook 1: Cognitive domain. New York, NY: Longman; 1956

FIGURE 1: Bloom's taxonomy.

including assessors, patients, space, finances and equipment, should be considered when planning and performing the assessment.[4]

In outcomes-based curricula, such as medicine, the core competencies should be stated clearly in relation to the requirements of regulatory bodies.[10] Moynihan et al.[11] define core competencies as:

> [*T*]he essential minimal set of a combination of attributes, such as applied knowledge, skills, and attitudes, that enable an individual to perform a set of tasks to an appropriate standard efficiently and effectively.

The Health Professions Council of South Africa (HPCSA) prescribes core competencies that should be incorporated in the training of undergraduate medical students in South Africa.[12] These competencies were derived from the original Canadian Medical Society (CANMEDS) document[13] and adapted for the South African context. The roles of a healthcare practitioner are central to these competencies: communicator, collaborator, health advocate, scholar, professional and leader, and manager.[12]

At the University of the Free State (UFS), a 5-year outcomes-based Bachelors of Medicine and Bachelors of Surgery (MBChB) programme is presented as training for medical doctors. The programme is divided into three phases over 10 semesters. Clinical training takes place in phase III (semesters 6–10). Clinical students rotate through six clinical blocks per year where they receive clinical and theoretical training. During rotations, continuous and end-of-block assessments take place. At the end of the academic year, students do a summative assessment in all disciplines. To progress to the next year or graduate (final year), students need to pass both the theoretical and practical component of each discipline separately. Knowledge, skills and attitudes are trained and assessed during this phase.[14] This article is part of an overarching project to address the quality of undergraduate medical assessment. In other parts of the study, the students' experiences and opinions were gathered, and the reliability of assessments was determined. Finally, lecturers discussed and made recommendations on how to improve current assessment practices to ensure defendable results.

The aim of this article was to report on current assessment practices in the clinical phase of the undergraduate medical programme at the UFS. The objectives were to describe different assessment methods that are used, the planning of the assessments, assessors and moderation practices, as well as how core competencies are assessed. Opinions on pass and fail decisions were also gathered, and recommendations for improving current assessments were obtained.

Methods

A descriptive cross-sectional study design was used. Mainly quantitative data were gathered by means of a questionnaire and were supported by qualitative data that provide clarifying information and recommendations by participants.

Study population and sampling

The study population consisted of teaching and learning coordinators (T&Ls) appointed in various clinical disciplines and module leaders of modules in disciplines that lacked T&Ls. The 13 clinical disciplines in phase III of the MBChB programme were eligible for inclusion.

A pilot study was conducted on two senior lecturers who were not part of the study population, to ensure that questions were clear and followed a logical sequence. Recommendations from a biostatistician were also incorporated before the questionnaire was finalised. One duplicate question was removed and the order of questions was changed to improve flow.

Measurement

A questionnaire was developed, taking the principles of questionnaire development into account.[15] Questions in the questionnaire were based on a framework to benchmark the quality of clinical assessment in a South African undergraduate medical programme.[16]

A self-administered, hard-copy questionnaire was distributed to T&Ls and/or module leaders in clinical disciplines at a phase III working group meeting. The staff members were invited to participate in the survey voluntarily. An information leaflet accompanied the questionnaire. Eligible staff members who were not present at the meeting received an electronic copy of the questionnaire, with an explanatory e-mail. An information leaflet and a hard copy of the questionnaire were also delivered to their offices. Participants returned questionnaires to the researcher in hard copy format or via e-mail. All participants signed informed consent. Data collection took place during September 2019.

The questionnaires obtained data about the different types or formats of assessment used, assessment planning and blueprinting, alignment of assessment with outcomes and

training, the assessment of core competencies required by the HPCSA, moderation practices and recommendations for improving assessment. Clarification data on how the core competencies, as described by the Medical and Dental Board of South Africa (part of the HPCSA), are assessed were grouped per competency. In addition, suggestions and recommendations on how to improve assessment were obtained.

Analysis of data

Data from the questionnaires were transferred to Excel datasheets by the researcher. The process of data transfer was done twice, to ensure integrity and accuracy. The Department of Biostatistics, Faculty of Health Sciences, did the data analysis of quantitative data with Analytics Software & Solutions (SAS) Version 9.4. Descriptive statistics, including frequencies and percentages, were calculated. Qualitative data were grouped by the first author according to themes.

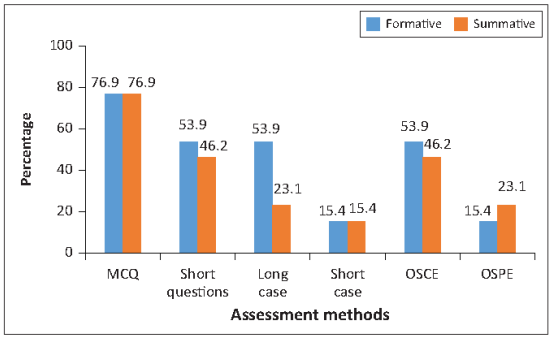
Ethical consideration

This study was approved by the Health Sciences Research Ethics Committee of the UFS (UFS-HSD 2019/0001/2304). Authorities at the UFS permitted the inclusion of UFS staff members in the study, and all participants signed informed consent. Although it was possible to identify individuals and disciplines from the questionnaires, no person or discipline was identified during the reporting of the data. All data were managed confidentially.

Results

All 13 disciplines in the study population returned completed questionnaires: general surgery, internal medicine, paediatrics, obstetrics and gynaecology, psychiatry, family medicine, urology, orthopaedics, otorhinolaryngology, ophthalmology, oncology, nuclear medicine and anaesthesiology.

Results show that different assessment methods were used for formative and summative assessment to assess theoretical knowledge and clinical skills, respectively. Workplace-based assessment (WBA), in the form of direct observation in the training area, was performed by 30.1% of disciplines.



MCQ, multiple-choice questions; OSCE, objective structured clinical examinations; OSPE, objective structured practical examination.

FIGURE 2: Percentages of different assessment methods used for formative and summative assessments.

Multiple-choice questions (MCQs) were used for formative and summative assessments by 76.9% of disciplines. Objective structured clinical examination (OSCE) was used by 53.9% of disciplines for formative assessment and by 46.2% for summative assessment. More long cases were used for formative assessment than for summative assessment (53.9% vs. 23.1%), while more objective structured practical examination (OSPE) was used for summative assessments than for formative assessments (23.1% vs. 15.4%). Figure 2 displays the percentages of different assessment methods used for formative and summative assessments.

Current assessment practices were evaluated based on assessment factors, assessor factors and moderation and feedback. Most disciplines always (46.2%) or usually (38.5%) blueprinted their assessments. Resources were not taken into consideration in the planning of assessment by 15.4% of disciplines. Table 1 shows how often various assessment factors were taken into consideration during the planning of assessment.

The results also show that assessors are not well prepared for the assessments in which they were involved: 38.5% had never received formal training before the assessment, while 30.8% had never been involved in assessment preparation. These results are shown in Table 2.

Results also show that, in most disciplines, the practices of feedback and moderation were not well established.

TABLE 1: The use of various assessment factors for assessment planning (%).

Assessment factor	Always	Usually	Sometimes	Never
Blueprinting	46.2	38.5	0.0	15.4
Blooms taxonomy level	30.8	46.2	0.0	23.0
Alignment with teaching	53.9	46.2	0.0	0.0
Alignment with module outcomes	46.2	53.9	0.0	0.0
Resources (patients, assessors, finances, etc.)	61.5	23.1	0.0	15.4
Standardised assessment tools	38.5	30.8	7.7	23.0
Assessment dates	30.8	30.8	0.0	38.5

TABLE 2: Assessor factors that may influence the quality of assessment.

Assessor factor	Always	Usually	Sometimes	Never
Assessors received formal training	7.7	23.1	30.8	38.5
Assessors received informal training	38.5	30.8	7.7	23.1
Assessors may use subjective marking	23.1	23.1	7.7	46.2
Assessors are involved with the preparation of the assessment	7.7	30.8	30.8	30.8
Assessors agree on the tool before the assessment	15.4	53.9	7.7	23.1

TABLE 3: Feedback and moderation practices.

Feedback and moderation	Always	Usually	Sometimes	Never
Students receive memorandums after assessment	15.4	23.1	7.1	53.9
Students receive feedback after assessment	15.4	38.5	23.1	23.1
Students receive results within 10 days of the assessment	53.9	7.7	15.4	23.1
Moderation takes place before the assessment	30.8	46.2	7.7	15.4
Moderation takes place after the assessment	15.4	38.5	23.1	23.1

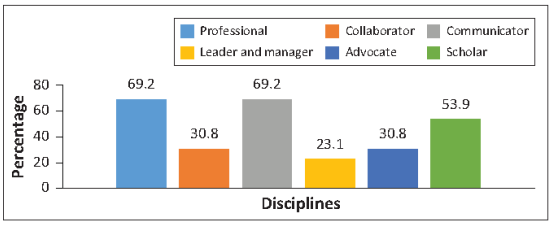


FIGURE 3: Percentages of disciplines that assessed the core competencies.

Only two disciplines (15.4%) always gave feedback after assessment, and only two disciplines made use of post-assessment moderation as a standard practice. Table 3 displays the results of feedback and moderation practices.

Respondent comments regarding feedback and moderation included the following:

> 'Moderation should not be a paper exercise. Moderators must help to improve the assessment.'
>
> 'Our department needs to introduce post assessment moderation.'

Most disciplines' participants indicated that they assess some of the core competencies prescribed by the HPCSA. Figure 3 displays the percentages of disciplines that assessed the six core competencies.

Seven disciplines (53.8%) indicated that the clinical trainers assess professionalism during the rotation, while one department reported making use of patient feedback. One department presents a session on the core competencies. Collaboration was mainly assessed during interprofessional training sessions. Communication was not formally assessed in most disciplines. Two disciplines assessed communication formally during case presentations, and seven disciplines indicated that they assess it during case presentations as part of the overall assessment. One discipline reported assessing communication through referrals to other healthcare workers. Although 23.1% of disciplines indicated that they assess the competency 'Leader and manager', they did not indicate how this was done. Being a 'Health advocate' was assessed mainly through observation of patient–student contact. This included how the student manages resources, provides holistic care and develops alternative management plans. More than half the departments indicated that they assessed the core competency of 'Scholar' – they did this by considering preparation for assessments as sole element.

All disciplines (100%) regarded their assessments as fair, while 92.3% indicated that their assessments were reliable and of an appropriate standard. Two disciplines (15.4%) believed that their assessments were not appropriate for assessing knowledge and skills. Half the disciplines indicated that all students who passed were competent to register as entry-level healthcare practitioners, and 30.8% indicated that students who failed were not competent to register as entry-level doctors.

An open-ended question was posed to the respondents regarding suggestions on how to improve the current assessment. Recommendations centred on the types and process of assessment, integration and planning of assessment and resources. The following suggestions and recommendations were transcribed verbatim.

Types and process of assessment:

> 'Maybe less emphasis on marks in the final cases and more formative assessment. Not all students perform best in high pressure clinical assessment.'
>
> 'There is a need for alternative assessments. Simulated cases and formative assessment should be used.'
>
> 'We need to change assessment procedures.'
>
> 'Students that pass block assessments should not need to do the end of year assessment again. We should be able to declare them competent or not competent after a rotation.'
>
> 'Patients change during assessment or change their story, it is not reliable.'
>
> 'Standard setting and rubrics in clinical assessment can decrease subjectivity.'

Integration and planning of assessment:

> 'What about one integrated OSCE?'
>
> 'The whole department should be involved with training, setting of papers and assessment.'
>
> 'We need to plan assessment from the beginning.'

Resources:

> 'Lack of trained educators, staff and resources need to be addressed.'
>
> 'Summative assessments are very labour intensive.'
>
> 'The use of patients in summative assessment is problematic due to numbers (of assessments per day).'
>
> 'We need expertise and support in IT [*Information Technology*]. We can't spend so much time on this.'

To conclude the questionnaire, respondents could make final comments. All but one were regarding the Nelson Mandela Fidel Castro Medical Programme (NMFCMP) students who share the training platform with the UFS students. Their comments were:

> 'The additional burden of the NMFCMP students have a negative impact on training and assessment. The extra number of students, as well as the extra effort to train them, decrease the student-lecturer ratios for current students.'
>
> 'More students pose problems with patients and logistics during assessment.'
>
> 'The current assessment is too labour intensive for the personnel numbers.'
>
> 'The logistics with so many students is a nightmare. We need alternative assessment like workplace assessment.'
>
> 'Resources should be addressed, like more lecturers and personnel to cope with these (NMFCMP) students.'

One final quote:

> 'There is always scope for improvement, but we do good. If the clinical training is good, the assessment should confirm what you already know.'

Discussion

Because the response rate was 100%, results are representative of assessment by clinical disciplines in the clinical phase of the undergraduate medical programme at the UFS. The results indicate the assessment practices in the different disciplines, while recommendations made reflect personal opinions of assessors responsible for assessments, and not necessarily that of the discipline involved.

As van der Vleuten[17] states, 'Any single assessment method can never be perfect on all criteria and in reality assessment always involves a compromise'. Therefore, different assessment methods should be used to assess students' competence in clinical medical training. More than three-quarters of disciplines in this study used MCQs to assess theoretical knowledge. The advantage of using MCQs in assessment includes feasibility. In addition, when questions are well constructed, validity and reliability improve. Half the disciplines used short-written questions, which have the advantage that logic, reasoning and problem-solving can be assessed. The disadvantages of short-written questions include that the marking is labour intensive and leaves scope for subjective opinions and rater (marker) bias.[18]

Performance tests or assessments are used when learners need to demonstrate their competence and are appropriate for clinical medicine.[19] These types of assessment reflect the level 'does' as described in Miller's pyramid.[20] Objective structured clinical examinations and WBA are examples of performance assessments. Workplace-based assessment is described as the observation of students while they are performing skills and competencies in the workplace ('does'); they receive immediate feedback, to improve, reinforce or certify a skill.[21,22] Less than one-third of disciplines in this study used direct observation or WBA for assessment. As WBA is authentic and tests performance,[21] it is recommended that this method of assessment should play a bigger role in clinical assessment.[5] Workplace-based assessment and training provide an opportunity to certify entrustable professional activities. An example of an entrustable professional activity is a student who demonstrates effective neonatal resuscitation every time it is done in the workplace; the skill can be certified, and there is no need for additional summative assessment.[23] Half the disciplines used OSCEs as part of their assessment. Despite being labour intensive, the validity and reliability of OSCEs make it an excellent assessment method to measure clinical competence.[24] Disciplines should be encouraged to use OSCEs to assess clinical competence. Half the disciplines used long cases during formative assessment. As a formative assessment method, observed long cases are appropriate for assessing holistic care, communication and problem-solving.[24,25]

A quarter of disciplines used unobserved long cases during summative assessment. This practice should be discouraged, as neither communication nor clinical skills can be assessed unless they are observed.[26,27] During summative assessment, students do only two or three long cases, making this a less valid and reliable assessment method because of the small numbers.[28]

Constructive alignment is essential during curriculum planning.[1] This involves the alignment of assessment with course outcomes and training.[2] Part of this planning should also include blueprinting[4] and determining the level of difficulty of the assessment using Bloom's taxonomy.[6] In this study, alignment between assessment, outcomes and training was always or usually taken into consideration during assessment planning. However, in less than half the assessments in the different clinical disciplines, blueprinting and the use of Bloom's taxonomy were considered in assessment planning. In some disciplines, it was not done at all. This is an area where assessment can be improved.

The feasibility of an assessment is dependent on resources.[4] However, resources were not consistently considered when assessment was planned. In general, the planning for assessment seems poor, with about a third of assessments performed without standardised assessments tools.

The reliability of an assessment is influenced by the assessment, the assessor and the student.[9] In this study, assessors were not well prepared or trained for assessments. Less than a third of assessors always or usually received formal assessment training before assessments. Two-thirds of assessors had (always or usually) received informal training before the assessment, and less than half of assessors were usually involved with assessment planning. The above assessor factors may contribute to unreliable assessments. Resources and opportunities are available to address the lack of formal assessment training. Assessors should be encouraged and supported to attend these courses as part of professional development.

More than half the disciplines allowed subjective marking. For many years, objectivity had been regarded as a cornerstone of assessment – the introduction of MCQs made objectivity possible.[29] In clinical practice, the management of patients is not unidimensional, and different approaches may all be correct. Ten Cate and Regehr[29] argue that 'subjective expert judgements by medical professionals are not only unavoidable but actually should be embraced as the core of assessment of medical trainees'.

Effective feedback in the workplace supports learning and competence development, which, in turn, improves patient care.[30] For feedback to be effective, it should be focused, specific and on time.[31] Feedback after assessment is a routine practice in only a few disciplines: only 15.4% of disciplines always gave feedback after assessments. If students do not receive feedback, they may continue to do things wrongly

without knowing it. These poor feedback practices should be flagged and addressed, to improve clinical competence and ultimately patient care.[21,31] Most disciplines indicated that they do not make memorandums available after assessments. This may be because many disciplines use MCQs and prefer to protect their question banks. Just more than half the disciplines always make marks available to students within the prescribed time frame of 2 weeks. When feedback and marks are late, students are already busy with a subsequent rotation and may not see the feedback as a learning opportunity.

From a quality assurance viewpoint, moderation of assessments is of utmost importance. Moderation is prescribed by the UFS assessment policy.[32] Moderation usually or always takes place before assessment in at least 75% of disciplines, but only in half of disciplines after assessment. This is a missed opportunity to improve assessment and assessment practices.

Of the six core competencies prescribed by the HPCSA,[12] only three – 'Professional', 'Communicator' and 'Scholar' – are assessed in more than half the disciplines. The competencies 'Professional' and 'Communicator' are assessed well during clinical training and patient presentations. 'Collaborator' is assessed mainly in disciplines where formal interprofessional training takes place. As students work daily with staff members from other disciplines, as well as in groups, the competency of 'Collaborator' can also be assessed through feedback from the team members. Although the competency 'Scholar' was assessed in most disciplines, it mainly took form of preparation for assessment (learning), and aspects such as the creation, dissemination and translation of knowledge were not assessed. Introducing a personal portfolio for each student, where all assessments of competencies throughout their training are recorded, may be effective in this regard.

Just less than half the T&Ls believed that some students who pass the summative assessment in clinical medicine are not competent to become entry-level healthcare practitioners fit for internship – they believe this despite assessors 'certifying' the students competent during assessment. This may be because of the phenomenon of failure-to-fail, where assessors pass incompetent students to prevent dealing with the consequences of failure.[33] It may also happen because they are not confident in the quality of assessments.

Valuable recommendations were made on how to improve the quality of assessment. Most of the respondents suggested less emphasis on summative assessment with more formative assessments. A move towards WBA and formative assessment with feedback is recommended when clinical competence must be assessed.[30] Better assessment planning, the use of standardised tools and better training of assessors were also proposed. Although these were individual recommendations, it will benefit to determine the need for training and provide support for all assessors.

The NMFCMP students are South African students trained in Cuba under the government to government agreement between South Africa and Cuba. These students train in Cuba and then return to South Africa where they are absorbed in the different undergraduate medical programmes to complete their last 18 months of training in South Africa.[34] These students are included in the normal formative and end-of-block assessments of the universities. The increase in student numbers, with the added NMFCMP students, creates tremendous pressure on the current training platform with limited resources. This may be a risk for poor quality of assessment and need to be addressed.

The feedback from students, as well as a focus group interview with the T&Ls in the different modules, may assist to form a holistic picture of current assessment practices. This may also explore resistance or opportunities for assessor training. With this added information, a formal proposal for the improvement of undergraduate medical assessment can be made. For now, better implementation of moderation practices, specifically post-assessment moderation, will contribute to improved quality assurance.

Conclusion

Current assessment practices for clinical students in the undergraduate medical programme at the UFS covers the spectrum that is necessary to assess all the different competencies. Multiple-choice questions and OSCEs, which are valid and reliable assessment methods, are used frequently. The lack of trained assessors, poor feedback and moderation practices should be addressed. More formative assessments, and less emphasis on summative assessment, should be investigated. Workplace-based and continuous assessment may be good ways to ensure the effective assessment of clinical competence.

Acknowledgements

Competing interests

The authors have declared that no competing interests exist.

Authors' contributions

H.B. was responsible for the conceptualisation of the study, protocol development, data collection and writing of the manuscript. J.B. and L.J.V.d.M. were the promoters who assisted with the conceptualisation and planning of the study, as well as critical evaluation and final approval of the manuscript. G.J. assisted with the concept and methodology, performed data analysis and assisted with the interpretation and write-up.

Funding information

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Data availability statement

Data will be available on request with permission of the authors.

Disclaimer

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of any affiliated agency of the authors.

References

1. Biggs J. Enhancing teaching through constructive alignment. High Educ. 1996;32(3):347–364. https://doi.org/10.1007/BF00138871
2. Biggs J. Aligning teaching for constructing learning. High Educ Acad. 2003;1(4):347–364.
3. South African Qualifications Authority. Criteria and guidelines for assessment of NQR registered unit standards and qualifications [homepage on the Internet]. South African Qualifications Authority; 2001 [cited 2019 Oct 12]. Available from: http://cdn.lgseta.co.za/resources/guidelines/2.4.1%20SAQA%20Criteria%20and%20Guidelines%20for%20Assessment.pdf
4. Norcini J, Friedman Ben-David M. Concepts in assessment. In: A practical guide for medical teachers. 5th ed. London: Elsevier, 2017:285–291.
5. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. Med Teach. 2007;29(9):855–871. https://doi.org/10.1080/01421590701775453
6. Bloom MS, Krathwohl DR. Taxonomy of educational objectives: The classification of educational goals, by a committee of college and university examiners. In: Handbook 1: Cognitive domain. New York, NY: Longman; 1956.
7. Royal K. Four tenets of modern validity theory for medical education assessment and evaluation. Adv Med Educ Pract. 2017;(8):567–570. https://doi.org/10.2147/AMEP.S139492
8. Manterola C, Grande L, Otzen T, Garcia N, Salazar, Quiroz G. Reliability, precision or reproducibility of the measurements. Methods of assessment, utility and applications in clinical practice. Rev Chil de Infectol. 2018;35(6):680–688. https://doi.org/10.4067/S0716-10182018000600680
9. Maughan S, Tisi J, Whitehouse G, Burdett N. A review of literature on marking reliability research [homepage on the Internet]. Slough: Natuaional Foudation for Educational Research; 2013 [cited 2019 Sep 12]. Available from: https://www.nfer.ac.uk/publications/mark01/mark01.pdf
10. Spady WG. Outcome-based education: Critical issues and answers [homepage on the Internet]. Arlington, VA: American Association of School Administrators; 1994 [cited 2019 Nov 11]. Available from: https://eric.ed.gov/?id=ED380910
11. Moynihan S, Paakkari L, Välimaa R, Jourdan D, Mannix-McNamara P. Teacher competencies in health education: Results of a Delphi study. PLoS One [serial online]. 2015 [cited 2019 Nov 11];10(12). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4667995/
12. Health Professions Council of South Africa. Core competencies for undergraduate students in clinical associate, dentistry and medical teaching and learning programmes in South Africa [homepage on the Internet]. Health Professions Council of South Africa; 2014 [cited 2019 Oct 12]. Available from: http://www.hpcsa.co.za/uploads/editor/UserFiles/downloads/medical_dental/MDB%20Core%20Competencies%20-%20ENGLISH%20-%20FINAL%202014.pdf
13. Seely J, Wade J. CanMEDS [homepage on the Internet]. The Royal College of Physicians and Surgeons of Canada; 1996 [cited 2019 Oct 12]. Available from: http://www.royalcollege.ca/rcsite/canmeds/about/contributors/canmeds-1996-contributors-e
14. University of the Free State. Faculty of Health Sciences rule book. School of Medicine. Undergraduate Qualifications 2019 [homepage on the Internet]. University of the Free State; 2019 [cited 2019 Sep 12]. Available from: https://apps.ufs.ac.za/dl/yearbooks/335_yearbook_eng.pdf
15. Pietersen J, Maree K. Surveys and the use of questionnaires. In: Maree K, editor. First steps in research. 2nd ed. Pretoria: Van Schaik, 2019; p. 215–224.
16. Brits H, Bezuidenhout J, Van der Merwe LJ. A framework to benchmark the quality of clinical assessment in a South African undergraduate medical programme. S Afr Fam Pract. 2020;62(1):a5030. https://doi.org/10.4102/safp.v62i1.5030
17. Van der Vleuten CPM. Revisiting 'Assessing professional competence: From methods to programmes.' Med Educ. 2016;50(9):885–888. https://doi.org/10.1111/medu.12632
18. Paniagua M, Swygert K, Downing S. Written tests: Writing high-quality constructed-response and selected-response. In: Yudkowski R, Park Y, Downing S, editors. Assessment in health professions education. 2nd ed. New York, NY: Routledge; 2019; pp. 109–126.
19. Yudkowski R, Park Y, Downing S. Introduction to assessment in health professions education. In: Yudkowski R, Park Y, Downing S, editors. Assessment in health professions education. 2nd ed. New York, NY: Routledge; 2019; p. 3–16.
20. Miller GE. The assessment of clinical skills/competence/performance. Acad Med. 1990;65(9):S63. https://doi.org/10.1097/00001888-199009000-00045
21. Lefroy J. Action research : Towards excellence in teaching, assessment and feedback for clinical consultation skills [homepage on the Internet]. [PhD]. Keele University; 2018 [cited 2019 Oct 12]. Available from: http://eprints.keele.ac.uk/5170/1/LefroyPhD2018.pdf#page=146
22. McBride M, Adler M, McGaghie W. Workplace-based assessment. In: Yudkowski R, Park Y, Downing S, editors. Assessment in health professions education. 2nd ed. New York, NY: Routledge; 2019; pp. 160–172.
23. Chen H, Van den Broek W, Ten Cate O. The case for use of entrustable professional activities in u...: academic medicine. Acad Med. 2015;90(4):431–436. https://doi.org/10.1097/ACM.0000000000000586
24. Patricio M, Julião M, Fareleira F, Carneiro A. Is the OSCE a feasible tool to assess competencies in undergraduate medical education? Evidence from a BEME Systematic Review. Med Teach. 2013;35(6):503–514. https://doi.org/10.3109/0142159X.2013.774330
25. Amin Z, Seng C, Eng K, editors. Practical guide to medical student assessment [homepage on the Internet]. Singapore: World Scientific Publishing; 2006 [cited 2019 Nov 11]. Available from: https://www.worldscientific.com/worldscibooks/10.1142/6109
26. Kamarudin MA, Mohamad N, Siraj MNABHH, Yaman MN. The relationship between modified long case and objective structured clinical examination (OSCE) in final professional examination 2011 Held in UKM Medical Centre. Procedia – Soc Behav Sci. 2012;60:241–248. https://doi.org/10.1016/j.sbspro.2012.09.374
27. Wass V, Jolly B. Does observation add to the validity of the long case? Med Educ. 2001;35(8):729–734. https://doi.org/10.1046/j.1365-2923.2001.01012.x
28. Wass V, Jones R, Der Vleuten CV. Standardized or real patients to test clinical competence? The long case revisited. Med Educ. 2001;35(4):321–325.
29. Ten Cate O, Regehr G. The power of subjectivity in the assessment of medical trainees. Acad Med. 2019;94(3):333–337. https://doi.org/10.1097/ACM.0000000000002495
30. Govaerts M. Workplace-based assessment and assessment for learning: Threats to validity. J Grad Med Educ. 2015;7(2):265–267.
31. Kelly E, Richards JB. Medical education: giving feedback to doctors in training. BMJ. 2019;19:i4523. https://doi.org/10.1136/bmj.l4523
32. University of the Free State. Guidelines for the implementation of external moderation [homepage on the Internet]. University of the Free State; 2009 [cited 2019 Oct 12]. Available from: https://www.ufs.ac.za/docs/default-source/all-documents/guidelines-for-the-implementation-of-external-moderation-404-eng.pdf?sfvrsn=2e2be421_0
33. Hughes LJ, Mitchell M, Johnston ANB. 'Failure to fail' in nursing – A catch phrase or a real issue? A systematic integrative literature review. Nurse Educ Pract. 2016;20:54–63. https://doi.org/10.1016/j.nepr.2016.06.009
34. Motala M, Van Wyk J. Where are they working? A case study of twenty Cuban-trained South African doctors. Afr J Prim Health Care Fam Med. 2019;11(1):a1977. https://doi.org/10.4102/phcfm.v11i1.1977

Departmental questionnaire

Assessment in the clinical phase of the MBChB programme

1. **Department:** .....

2. **In your assessment of medical students in Phase III (clinical phase) please indicate with a cross which of the following continuous (formative) assessment methods you use:**

2.1 Continious assessment:

Logbook	
Case presentations	
Tests (formal and informal)	
Assignments	
Rotation mark	
Direct observation of skills	

What other assessments do you use for continuous assessment? .....

.....

2.2 End of block assessment

MCQ test	
Short question test	
Long clinical case/s	
Short clinical case/s	
OSCE	
OSPE	

What other assessments do you use for end of block assessment? .....

.....

2.3 End of year assessment

MCQ test	
Short question test	
Long clinical case/s	
Short clinical case/s	
OSCE	
OSPE	

3. **How often do you take the following in consideration when planning assessments?**

3.1 For individual assessments e.g. OSCE, clinical cases, oral or paper

	Always	Usually	Sometimes	Never
Blueprinting to ensure even spread				
Blooms taxonomy to determine cognitive level				
Assessment aligned with teaching				
Assessment aligned with outcomes				
Resources available (e.g. patients, assessors, finances)				
When and how feedback will be given				
Standardised assessment tools/rubrics				

3.2 For the whole assessment (Clinical and Theory combined)

	Always	Usually	Sometimes	Never
Blueprinting to ensure even spread				
Blooms taxonomy to determine cognitive level				
Assessment aligned with teaching				
Assessment aligned with outcomes				
Resources available (e.g. patients, assessors, finances)				
When and how feedback will be given				
Standardised assessment tools/rubrics				

4. **Regarding assessors/examiners:**	Always	Usually	Sometimes	Never
Do you ensure that examiners received formal training before assessments?				
Do you ensure that examiners received informal training or orientation before assessments?				
Do you allow for subjective assessments by examiners?				
Are your examiners involved in the preparation of assessment tools/rubrics before assessments?				
Do your examiners agree with the assessment tools/rubrics before they are used?				

5. **Regarding administation around assessments:**	Always	Usually	Sometimes	Never
Do you compile memoranda for tests/assessments?				
Do you make memoranda for tests available to students?				
Do you send tests for moderation before assessments?				
Does the moderator verify the results after marking?				
Do you make results available within 10 days of assessment?				
Does the first assessment take place within 1 month of the start of the rotation?				

6. **HPCSA required Competencies:**

Please indicate if and how you test the following competencies?

A healthcare practitioner to be a:	Tested	How
Professional		
Collaborator		
Communicator		
Leader and manager		
Health advocate		
Scholar		

7. **In your opinion, indicate if the current assessment practices in your department are:** (Motivate)

Appropriate	Y	N	.....
Reliable	Y	N	.....
Fair	Y	N	.....
Of good standard	Y	N	.....

8.	**Do you think that all students who pass the final assessment in your department are competent to be an intern?**	Y	N
9.	**Do you think that all students who fail should receive extended training time?**	Y	N

10. **Do you have any suggestions on how to improve the current assessment practices?**

.....

.....

.....

.....

Thank you

3

INFORMATION LEAFLET

Study title: BEST PRACTICES FOR QUALITY ASSESSMENT IN THE CLINICAL PHASE OF UNDERGRADUATE MEDICAL TRAINING.

Dear Colleague

I, Hanneke Brits, am doing research for my PhD study. Research is just the process to learn the answer to a question. In this study, I want to learn more about the current assessment practices in the clinical phase of the MBChB programme to assist me to compile best practice recommendations to ensure quality assessment for our students.

This study will consist of different parts. In part 1, I will study specific policies and practices and compile a theoretical guideline for quality assessment in the clinical phase of the MBChB programme. In part 2, I would like to assess current assessment perceptions and practices in the clinical phase of the MBChB programme. In the first part, I ask your opinion on the assessment in an anonymous questionnaire. Questionnaires will be filled by both students and lecturers in the clinical phase of the MBChB programme.

In part 3, I will discuss the differences between the current practices and the theoretical requirements for quality assessment with the Teaching and Learning coordinators and module leaders of the different departments to compile recommendations on how to ensure and maintain quality assessment in the MBChB programme. This discussion will be video and audio taped, but no person or department will be identified in any report.

There will be no **risks** involved when participating in the study and you will not receive any compensation for your participation.

A **benefit** of being in the study is that you may contribute to sustainable quality assessment in the clinical phase of the MBChB programme.

Participation is voluntary, and refusal to participate will involve no penalty or loss of benefits to which you are otherwise entitled; you may discontinue participation at any time without penalty or loss of benefits to which you are otherwise entitled.

Confidentiality: Efforts will be made to keep personal information confidential. No personal name will appear on any of the data collection instruments. Absolute confidentiality cannot be guaranteed. Organizations that may inspect and/or copy this research records for quality assurance and data analysis include groups such as the Health Sciences Research Ethics Committee.

If results are published, this may lead to cohort identification.

I am **inviting** you to participate in this research project by filling in the anonymous questionnaire and participate in the focus group interview.

For further information/reporting of study-related adverse events please contact Prof Hanneke Brits at britsh@ufs.ac.za or the HSREC Chair and Secretariat **Ethics@ufs.ac.za** or (051) 4017794/5 for reporting of complaints/problems.

2

CHAPTER 5:

ARTICLE 4 – EVALUATION OF ASSESSMENTS MARKS OF AN UNDERGRADUATE MEDICAL TRAINING PROGRAMME: WHERE ARE WE AND HOW CAN WE IMPROVE?

The article was prepared according to the submission guidelines of the *African Journal of Health Professions Education.*

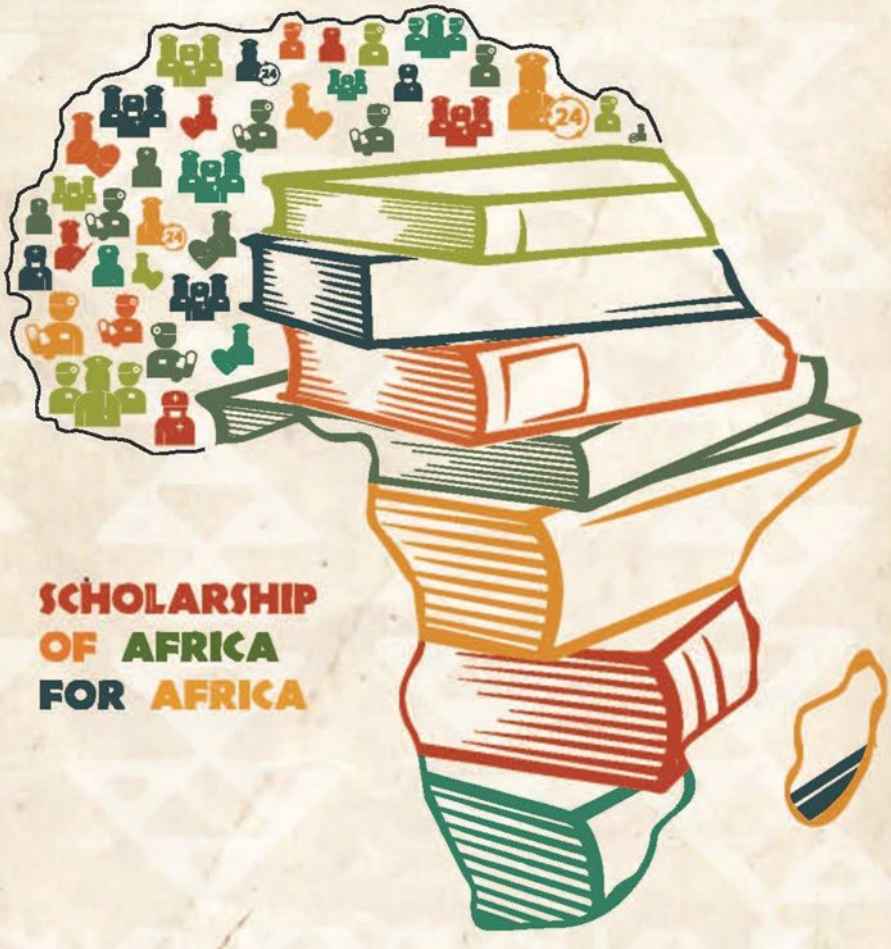
Journal:	African Journal of Health Professions Education
Publisher:	Health & Medical Publishing Group
Country of publication:	South Africa
ISSN:	2078-5127 (ONLINE)
DOI:	

Citation:

Brits, H., Joubert, G., Bezuidenhout, J. & Van der Merwe, L.J. Evaluation of assessment marks in the clinical years of an undergraduate medical training programme: Where are we and how can we improve?

Submitted:	12 May 2020
Accepted:	Under review

African Journal of Health Professions Education



ISSN 2078 - 5127

Overview

ISSN

2078-5127 (ONLINE)

Focus and scope

The AJHPE is a journal for health professions educators. It carries research articles, short scientific reports, letters, editorials, education practice, personal opinion and other topics related to the education of health care professionals. It also features African education-related news, obituaries and general correspondence. A double blind review process is followed, which means that both the reviewer and author identities are concealed from each other throughout the review process. A majority of manuscripts will be sent to one or two reviewers, under the management of an editor assigned to the submission. Reviewers as experts in the field provide comments to authors and editors on the importance, originality and scientific merit of the manuscript and suggest changes which may improve the quality and validity of the manuscript.

Editor

HB Research

Evaluation of assessment marks in the clinical years of an undergraduate medical training programme: Where are we and how can we improve?

H Brits, MBChB, MFamMed, MHPE, FCFP, G Joubert, BA, MSc., J Bezuidenhout, BA, HDE. BA Honn (Psyc.), PGDipl. (HPE), MEd (Psyc. Ed.), DTechEd., L J Van der Merwe, MBChB, MMedSc, DA(SA), PhD (HPE).

Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa

Corresponding author: H Brits (britsh@ufs.ac.za)

Background: In high-stakes assessments, the accuracy and consistency of the decision to pass or fail a student is as important as the reliability of the assessment. The aim of this study was to evaluate the quality of assessment results in the clinical phase of an undergraduate medical programme and to identify areas for improvement.
Methods: A cohort analytical study design was used. The final, end-of-block marks and the end-of-year assessment marks of both fourth-year and final-year medical students over three years were compared for decision reliability, test-retest reliability, stability and reproducibility.
Results: 1 380 marks in 26 assessments were evaluated. The G-index of agreement for decision reliability ranged from 0.86 to 0.98. In 88.9% of assessments, the test-retest correlation coefficient was <0.7. Mean marks for end-of-block and end-of-year assessments were similar; however, the standard deviation of differences between end-of-block and end-of-year assessment marks were high. Multiple-choice questions (MCQ) and objective structured clinical examinations (OSCE) yielded good reliability results.
Conclusion: The reliability of pass/fail outcome decisions was good. The test reliability, as well as stability and reproducibility of individual student marks, could not be accurately replicated. The use of MCQs and OSCEs are practical examples of where the number of assessments can be increased to improve reliability. In order to increase the number of assessments and to reduce the stress of high-stake assessment more workplace-based assessment with observed clinical cases can be recommended.

INTRODUCTION

The pass/fail decision in summative assessment for medicine and other professional qualifications holds many consequences for the various stakeholders.[1-3] Failure and having to repeat modules have financial and emotional implications for students, while they may lose trust in the training institution.[4] Student failure may affect throughput rates, as well as the reputation of the faculty or university.[4] However, passing an incompetent student may affect both patients and the healthcare system, e.g., through loss of life and avoidable expenses; it could also lead to misconduct claims against individuals or institutions.[5,6] Miller[7] emphasises this important responsibility relating to assessment:

> *If we are to be faithful to the charge placed upon us by society to certify the adequacy of clinical performance... then we can no longer evade the responsibility for finding a method that will allow us to do so.*

If we are to be able to defend the outcome of high-stakes exams, where the outcome has major consequences,[8,9] the assessment must meet the basic requirements of validity, reliability and fairness.[10-11] From a theoretical perspective, it is possible to improve the quality of assessment by addressing criteria such as validity, reliability and fairness.[12]

An assessment is considered valid when it measures what it is supposed to measure.[13-14] In the case of clinical medicine, competence must be measured. Validity in clinical assessment is usually evaluated using Miller's assessment framework.[15] According to this model, a valid assessment for competence must be on the "show how" and "does" level. However, when validity is increased by assessing in real-life situations, the reliability of the assessment may decline, due to subjective judgements and the lack of standardisation.[16] Before the validity of an assessment can be evaluated, its reliability must be established.[14]

The reliability of a clinical assessment is defined as the degree to which a test measures the same concept in different assessments and obtains stable or reproducible results.[17,18] Reproducibility, a synonym for reliability, is described as the closeness of or variation in results of successive measurements of the same assessment carried out under the same or nearly the same conditions.[19] With any assessment some form of "measurement error" will occur. This error should be as low as possible to ensure accurate assessment. The calculation of this error determines the reliability of an assessment.[18] Reliability can be evaluated with various measures, depending on the data that is available and what you want to establish.[20]

From a theoretical viewpoint, an assessment can be considered fair if everybody is subjected to the same assessment, under the same conditions, and all are marked

by the same assessors using the same mark sheets.[21] In practice, an assessment is fair when the interpretation of the results is transparent and just, and when nobody is disadvantaged in the process.[22]

One of the aims of assessment evaluation should always be to improve the quality of assessment for all stakeholders.[2] A fine balance should exist between traditional and innovative assessment methods by selecting judiciously sound assessment methods above tradition or convenience.[9] The decision to change or improve assessment practices or to move towards more innovative assessments should be based on facts rather than on preferences.[9] As a first step to improve the quality of assessment in the clinical years of undergraduate medical training the quality of current assessment must be established.

Pass/fail decisions are made on predetermined criteria. In high-stakes assessments, the accuracy and consistency of the decision to pass or fail a student is as important as the reliability of the test or assessment.[3,7,23,24] Decision reliability is a term used to measure the consistency with which pass/fail decisions are made.[3]

The best way to evaluate the reliability of a clinical assessment is to assess the same participants under similar circumstances on more than one occasion,[25] which is almost impossible in real life situations. The reliability of an assessment can be improved by using standardised questions and mark sheets, and multiple and trained markers, and by increasing the number of questions.[20] A high correlation between the different test scores (r >0.7) is indicative of test-retest reliability.[26]

The undergraduate medical programme at the University of the Free State (UFS) is a five-year, outcomes-based programme that runs over 10 semesters. The clinical phase is presented from semester 6 to 10. In the clinical phase of the programme, students are assessed in different disciplines. Some disciplines are grouped together to form a module. For example, in the fourth year, the surgery module consists of general surgery, orthopaedics, ophthalmology and otorhinolaryngology. Modules are presented in blocks. Students rotate between different blocks to cover all modules presented in the specific year. At the end of each rotation (block), students are assessed by the end-of-block assessment. In the fourth year, students must pass all disciplines to progress to the fifth year. If students meet minimum requirements in the fifth year, but fail certain disciplines, they are required to repeat only the failed disciplines. Admission to the final end-of-year assessment in the fourth and final year requires that students meet end-of-block academic as well as attendance requirements. Students in the fourth and final years must pass all disciplines in all the modules, including each of the clinical and theoretical components individually (if applicable), to pass the final end-of-year assessment.[27] Regarding clinical cases, students must also pass more than 50% of the cases, irrespective of the overall clinical mark obtained. If a student fails the end-of-year assessment (fourth and final years), but meets minimum requirements for reassessment, the student is allowed to do a reassessment within one week of the end-of-year assessment.[27] The pass mark for assessments is predetermined at 50% as per university regulations. No formal standard setting process exits. Assessments are blueprinted and assessments rubrics or memoranda are moderated before assessments. See Figure 1 for a flow diagram of the assessment process.

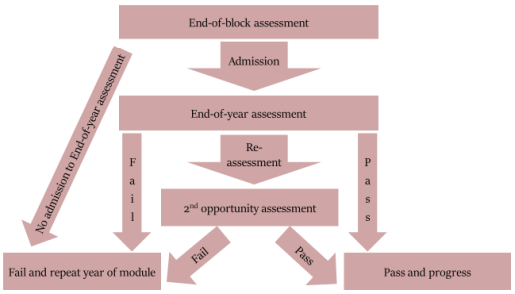


Figure 1: Flow diagram of assessment process and outcome in the fourth and final years of the undergraduate medical programme, UFS.

The end-of-block assessment and the end-of year assessment cover the same content and are generally conducted by the same assessors (academic staff in clinical departments). Both these assessments consist of theoretical as well as clinical assessments. Different disciplines structure their assessments differently, which make comparison between disciplines not feasible. No regulation or specific reason was found for conducting an end-of-year assessments after the end-of-block assessments and it is possibly more traditional than evidence based.

The aim of this study was to evaluate the assessment results of high-stakes assessments in the clinical phase of the undergraduate medical programme. This will assist to make recommendations for improving the quality of current assessments in the undergraduate medical programme, with validity, reliability and fairness in mind. The objectives were as follows:

1. To determine the decision reliability of the current summative assessments, and whether pass/fail decisions can be defended;
2. To determine the test-retest correlation between different assessments;

3. To compare the reliability results of different assessment methods.

METHODS

A cohort analytical study design was used. The study population consisted of all the fourth-year and fifth (final) year undergraduate medical students at the UFS who participated in the last end-of-block and end-of-year assessments of 2016, 2017 and 2018. The last end-of-block marks (obtained during the last rotation of the year) and the end-of-year assessment marks obtained during the final assessment at the end of the academic year, were used for data analysis. Data were collected retrospectively. Between the last end-of-block assessment and the end-of-year assessment, no formal training and very little learning takes place, which make these assessments comparable, but not identical.

The authors used an aggregated approach to look at the reliability of assessments, as an individual approach was impossible due to the variability in the way each discipline designs multiple choice questions (MCQs), clinical cases and objective structures clinical evaluations (OSCEs) and or objective structured practical evaluations (OSPEs). The reliability of the theoretical and clinical assessments were determined separately. Theoretical assessments consisted of papers with MCQs only and papers with a combination of MCQs and mostly short written questions. Clinical assessments included clinical cases, OSCEs and OSPEs. In clinical cases the student assesses a patient unobserved and then reports on findings while the assessors clarify findings and ask predetermined questions. The term OSCE is used for assessments in the form of clinical stations with patients or simulated patients. Students were directly observed at these clinical stations. The term OSPE was used for assessment involving unmanned stations, where students had to interpret diagnostic investigations, e.g., x-rays or laboratory results. Different disciplines use different combinations of assessments, however disciplines use the same combinations during end-of-block and end-of-year assessments.

Table 1 categorises the disciplines as either surgical or medical, indicates the study year(s) in which a discipline is presented, and lists the different assessment methods used for each discipline. General surgery, orthopaedics, urology, otorhinolaryngology, ophthalmology, anaesthetics and obstetrics and gynaecology are classified as surgical disciplines (n=7). Internal medicine, paediatrics, family medicine, oncology and psychiatry are classified as medical disciplines (n=5).

Table 1: Classification of disciplines, study years of presentation and types of assessment per discipline

Discipline	Classification	Study year	Assessment types
Discipline A	Surgical	4 and 5	Theory, clinical
Discipline B	Surgical	4 and 5	Theory, clinical
Discipline C	Surgical	4 and 5	Theory, clinical
Discipline D	Surgical	4	Theory
Discipline E	Surgical	4	Theory
Discipline F	Surgical	5	Theory, clinical
Discipline G	Surgical	5	Theory
Discipline H	Medical	4 and 5	Theory, clinical
Discipline I	Medical	4 and 5	Theory, clinical
Discipline J	Medical	4 and 5	Combined
Discipline K	Medical	4	Theory
Discipline L	Medical	5	Combined

Data collection

Student marks, corresponding with respective student numbers, were obtained from the official marks database used by the Faculty of Health Sciences. This is an extensive database with numerous datasets, available in Excel spreadsheets for each student per discipline and per assessment. It is a secure database with password protection – only authorised access is permitted. All marks, including that of reassessments, were used to compare final pass/fail outcome decisions.

Data management and analysis

The Department of Biostatistics performed data analysis using SAS Version 9.4. Calculations were done per discipline for fourth-year and fifth-year students separately.

The decision-reliability between the final end-of-block and end-of-year assessment was calculated using 2x2 tables. Due to the skewed data, kappa values could not be calculated[28] and a value of ≥ 0.7 on Holley and Gilford's G-index of agreement[29] (as an alternative for categorical judgement) was considered as reliable. The Holley and Gilford's G-index of agreement allows for correlation in the presence of skewed data. As a final step to evaluate the pass/fail outcome decisions, the reassessment outcome decision was compared with the final end-of-block and end-of-year assessment outcome decisions.

To determine test-retest reliability between final end-of-block and end-of-year assessment marks, Pearson correlation coefficients were calculated. A correlation coefficient of ≥ 0.7 was considered as reliable.[26]

The mean and standard deviation of differences between end-of-block and end-of-year assessment marks were calculated. This mean is used as an indication of assessment stability. The percentage of students whose

marks for the end-of-block and end-of-year assessments differed by less than 10% for the two assessments was calculated to assess reproducibility. The assessment was considered reliable if the reproducibility was at least 80%.

For clinical cases, the individual student marks obtained in consecutive assessments performed on the same day were also compared. The mean of the different cases was compared to determine test consistency, and the variance in marks (standard deviation) obtained by individual students was calculated to determine reproducibility.

Ethical considerations, quality and rigour of data management

Ethical approval to conduct the study was obtained from the Health Sciences Research Ethics Committee of the UFS (UFS-HSD 2019/0001/2304) and permission to use student data was granted by the relevant university authorities. All data were managed confidentially and only student numbers were used. No student or discipline is identified in the published results.

RESULTS

A total of 1 380 marks in a total of 26 administered assessments were evaluated. In Table 2 the numbers of students included in the study per discipline are indicated for the different years. Some disciplines are presented in only one of the study years (see Table 1). The study used the marks of 12 disciplines within the medical programme.

Table 2: Number of students per discipline for different study years

Discipline	4th year		
	2016	2017	2018
Discipline A	30	37	26
Discipline B	30	37	26
Discipline C	30	37	26
Discipline D	30	37	26
Discipline E	28	37	27
Discipline F	21	22	17
Discipline G	21	22	17
Discipline H	21	19	19
Discipline I	22	22	16
	5th year		
	2016	2017	2018
Discipline A	23	29	35
Discipline B	23	29	35
Discipline C	14	20	18
Discipline F	23	29	35
Discipline G	31	28	37
Discipline H	19	19	19
Discipline I	31	28	37
Discipline J	19	18	21
Discipline L	18	18	21

Decision reliability of pass/fail decisions

In two of the 12 4th year assessments, and seven of the 14 5th year assessments, the pass/fail decisions in the final end-of-block concurred with the end-of-year assessments, all students passed. In the remaining disciplines, there were between 92.5% and 98.9% agreement of the same pass/fail decision outcome between the end-of-block and end-of-year assessments. The G-index of agreement values ranged from 0.86 to 0.98.

Three fourth-year students obtained marks below 50% in the final end-of-block assessment. They subsequently failed the end-of-year assessment too, as well as the reassessment, and therefore had to repeat the year. No fifth-year students obtained marks below 50% in the final end-of-block assessment, or failed the year. Three fourth-year students and two fifth-year students passed the final end-of-block assessments, and then failed a subcomponent of a discipline/module in the end-of-year assessment. All these students qualified for reassessment, according to the rules, and all passed the reassessment and, therefore, passed the year.

Test-retest correlation of end-of-block and end-of-year results

In the fourth and fifth years respectively, 12 and 15 assessments were compared for test-retest correlation. Three assessments in the fourth year had correlation coefficients of ≥ 0.70. None of the assessments in the fifth year had correlation coefficients of ≥ 0.70. These results are displayed in Table 3.

Stability of assessment marks per discipline

Table 4 summarises the differences between the final end-of-block and the end-of-year assessment marks per discipline and per study year. The mean differences between marks obtained in the final end-of-block and end-of-year assessments varied between -11.4% (Discipline K, fourth-year group) and 7.5% (Discipline F, fifth-year group), with Discipline K emerging as a clear outlier.

Reproducibility and assessment methods

The percentage of students whose final end-of-block and end-of-year assessment marks were within a 10% range varied between 33.3% (Discipline K fourth year) and 98.9% (Discipline I fifth year). The individual marks of students varied considerably, as indicated by the high standard deviation, particularly for the fourth-year group.

In Table 5 these percentages are given for the different assessment methods.

Differences between marks for consecutive clinical cases

In three disciplines, students were assessed on two or three clinical cases on the same day. The mean marks obtained per discipline were within 4.5% of each other.

between 0 and 45% for different cases in the same discipline. In Table 6, the mean, standard deviation, minimum and maximum of differences in student marks obtained for consecutive cases are indicated per discipline.

Table 3: Correlation between final end-of-block and end-of-year assessment marks, per discipline, study year and type of assessment

Discipline	4th year			5th year		
	Theory	Clinical	Comb	Theory	Clinical	Comb
Discipline A	0.39* (p<0.01) #	0.47 (p<0.01)		0.61 (p<0.01)	0.34 (p<0.01)	
Discipline B	0.32 (p<0.01)	0.48 (p<0.01)		0.23 (p<0.01)	0.24 (p=0.03)	
Discipline C			0.60 (p<0.01)	0.34 (p=0.01)	0.67 (p<0.01)	
Discipline D	0.80 (p<0.01)					
Discipline E	0.50 (p<0.01)					
Discipline F				0.57 (p<0.01)	0.40 (p<0.01)	
Discipline G				0.25 (p=0.01)		
Discipline H	0.61 (p<0.01)			0.62 (p<0.01)	0.35 (p<0.01)	
Discipline I	0.78 (p<0.01)	0.93 (p<0.01)		0.64 (p<0.01)	0.32 (p<0.01)	
Discipline J			0.23 (p=0.7)			0.66 (p<0.01)
Discipline K	0.43 (p<0.01)					
Discipline L						0.46 (p<0.01)

* correlation coefficient # p-value

Table 4: Differences between the end of block and end-of-year assessment marks, per discipline and study year.

Discipline	4th year		5th year	
	Mean*	Standard Deviation	Mean	Standard Deviation
Discipline A Theoretical	3.29	11.04	-0.87	7.25
Discipline A Clinical	-3.44	9.21	-4.51	7.63
Discipline B Theoretical	1.90	10.67	-0.06	9.82
Discipline B Clinical	-1.22	10.89	-0.64	7.63
Discipline C Theoretical	2.05	6.81	2.06	8.62
Discipline C Clinical			0.79	6.12
Discipline D Theoretical	-1.91	4.94		
Discipline E Theoretical	-1.51	6.27		
Discipline F Theoretical			-2.29	9.10
Discipline F Clinical			7.45	8.59
Discipline G Theoretical			-4.16	10.63
Discipline H Combined	-0.78	5.69	1.11	4.68
Discipline I Theoretical	4.43	4.75	4.08	6.01
Discipline I Clinical	-1.36	2.10	-7.27	7.36
Discipline J Combined	2.65	12.38	0.97	5.55
Discipline K Theoretical	-11.42	12.84		
Discipline L Combined			-1.44	5.30

* A positive mean indicates that the end-of-year marks were higher than the end-of-block marks, while a negative mean indicates that the end-of-year marks were lower than the end-of-block marks.

Table 5: Percentage of students whose final end-of-block assessment and end-of-year marks were within a 10% range, by discipline, year group and assessment method

	Theory		Clinical		
	MCQ	Combined paper	Clinical case	OSCE	OSPE
4th Year					
Discipline A	67.7#				74.2
Discipline B	65.6*				64.5
Discipline C	88.3				
Discipline D		95.7			
Discipline E		91.4			
Discipline H	91.5				
Discipline I	90.2			98.9	
Discipline J					56.7
Discipline K		33.3			
5th year					
Discipline A	88.5		71.3		
Discipline B	72.4*				81.6
Discipline C		75.0		90.4	
Discipline F	69.0*				62.1
Discipline G		64.6			
Discipline H	98.3		60.3		
Discipline I	87.5		63.5		
Discipline J				94.8	
Discipline L				96.5	

#This assessment originally consisted of 30 questions, but the number of questions increased in 2017 *These assessments consisted of 30 or fewer questions per assessment.

Table 6: Differences in marks obtained for consecutive cases per discipline

Discipline	Cases	Mean difference	Standard deviation	Minimum	Maximum
Discipline A	Case 1+2	1.99	12.21	-32	45
	Case 1+3	1.85	12.78	-35	37
	Case 2+3	-0.15	12.98	-37	30
Discipline H	Case 1+2	-2.32	14.99	-45	38
	Case 1+3	-4.36	14.55	-44	35
	Case 2+3	-2.04	13.20	-36	43
Discipline I	Case 1+2	1.67	11.44	-25	30

DISCUSSION

The results presented here may be considered representative of the selected study population as all the student marks were available, in a usable format, in the database.

The aim when evaluating the quality of an assessment should be to identify areas that can be improved in the assessment.[30] Data for this study were obtained with this aim in mind rather than to pronounce judgement on the reliability of current assessment methods and practices.

Calculating the reliability of pass/fail outcome decisions using a kappa-coefficient is described in literature.[9,31] In this study, very few students failed, and the small numbers made the kappa statistic inappropriate for this measurement.[32] A G-index of agreement was, therefore, calculated.[28] In almost half the disciplines investigated, the agreement between the outcomes obtained in assessments was 100%. For the remaining disciplines, the G-index of agreement was above 0.85. The decision reliability on pass/fail outcome decisions for clinical assessments in the undergraduate medical programme at the UFS can, therefore, be considered excellent. The comprehensive end-of-block assessments, the strict admission requirements to the end-of-year assessment, and the reassessment opportunity may be reasons for this finding. Each individual student result, as well as discipline specific results, are discussed at the examination admission and final examination meeting to ensure defensible outcomes. With the current measures in place and the addition of standard setting to ensure accurate pass/fail decisions, during end-of-block assessments, the necessity of an end-of-year assessment may be reconsidered.

The test-retest correlations were low and did not reach a value of ≥0.7 for any of the fifth-year students' assessments. This indicates poor reliability for individual assessments. The reliability of an assessment can be affected by the students, the test and the markers.[17,20] Student factors that could contribute to the low test-retest correlations include that students who had passed the recent final end-of-block assessment might be confident about passing the end-of-year assessment, and then opt to study for disciplines/modules in which they had passed the end-of-block assessment some time ago. The added stress of high-stakes assessment, together with uncertainty about future work and placement, could also influence students' performance. Performance stress during high-stakes assessments is well described.[33,34] The effect of additional stress is unpredictable – it can have a positive or negative effect on academic performance.[35] More and regular low-stakes assessments may address the student factors described above. Test factors that could have played a role in this study include that, even though the same content was assessed in both the final end-of-block and end-of-block assessments, the questions differed and no formal standard setting was performed. Furthermore, not all competencies could be tested in all assessments, with very few assessments performed summative assessments. Competency in one case has a poor reproducibility for another case.[1] Finally, the markers stayed the same during both assessments, with the exception of a few additional external assessors. By increasing the number of assessments during rotations the reliability of overall assessment can also be improved.

The mean marks obtained in the end-of-block compared to end-of-year assessments did not differ much. The exception was the theoretical assessment in one discipline in the fourth year, where the end-of-year mark obtained was 11.4% lower than the final end-of-block module mark. The reason for this difference is not clear, though moderation reports of these assessments could provide some insight. The small variation in the mean marks (end-of-block vs. end-of-year) per discipline may be an indication that the assessments were of the same standard. However, the standard deviation was high for all assessments, indicating large differences in the marks obtained by individual students in the two assessments. These differences occurred in theoretical as well as clinical assessments. Poor validity of the assessments or the student factors discussed above may be reasons these differences.

Assessment methods varied across disciplines and, therefore, direct comparisons could not be made between different assessment methods. For theoretical assessments, MCQ papers with more than 30 questions produced student marks within a 10% range, indicating reproducibility. Reproducibility could not be proved for assessments with fewer than 30 questions. The reproducibility of assessments can be improved by increasing the number of questions.[20] Clinical OSCEs yielded good reproducibility results, while OSPEs and clinical cases did not. Patrício and co-workers[36] analysed the results of 366 articles on OSCEs performed in undergraduate medical education, and concluded that OSCEs produced reliable results and were feasible for assessing competence. An OSCE in itself is not reliable, but can produce reliable results if adequate sampling, good quality questions and mark sheets, time allocation per station and trained assessors are used.[19,20,37] OSPEs lack clinical interaction and demonstration of competence, making OSPEs almost equal to written questions.[38]

Clinical cases or long cases are renowned for their poor validity and reliability.[39] Evaluation of the marks obtained for consecutive clinical cases revealed a high standard deviation, despite a stable mean mark. A difference of up to 45% was observed in marks obtained for different clinical cases performed by the same student.

A possible reason may be patient selection and reuse of patients for the assessment. It is difficult to find enough suitable, similar and stable patients to use in clinical cases, making long cases less practical and reliable for summative assessment.[40] Assessors also need to make subjective judgements of competence which may influence reliability. Nevertheless, clinical cases have a definite role to play in low-stakes and formative assessment of which the aim is learning.[1] An advantage of using long cases is that a student can be assessed holistically on an actual case.[40] This advantage is lost when the student's examination of the patient is unobserved, and is followed by the student reporting his/her findings.[41] It has been calculated that 10 clinical cases are necessary to achieve acceptable reliability with clinical cases.[42] These numbers are only possible when workplace-based assessments are used.[9] From the above it is recommended that clinical cases only be used for formative assessment.

Reliability is only one aspect of quality assessment to ensure clinical competence. To achieve quality assessment of clinical competence, students should be assessed in real life, or in near-real-life situations. Assessing clinical competence is a complex procedure, with many dimensions requiring different assessment methods.[1,30,43] The highest level of competence, according to Miller's framework[7] for assessment, is "does". To ensure the competence of future medical professionals, we should assess them frequently, and in the workplace, and move away from overemphasis on high-stakes assessments.[44] Miller[7] states that,

> *No single assessment method can provide all the data required for judgment of anything so complex as the delivery of professional services by a successful physician.*

However, real-life situations are not stable and reproducible. This poses challenges ensuring the reliability of assessments.[45] It is important to take the quality of the assessment process as a whole into account, and to avoid merely focusing on validity, reliability or fairness as individual components to improve the assessment.

Limitations to the study

The quality of pass/fail decisions for the individual assessments (end-of-block and end-of-year) were not formally established before these assessments were compared with each other. However, the outcome of each assessment per student is discussed during the examination admission meeting and the examination results meeting to ensure accurate decisions.

The validity and fairness of the assessments were not assessed in this paper. This paper is only a step in the process to assess the quality of assessment.

The end-of-block and end-of-year assessments that were compared are not identical, but comparable. It is almost impossible to get identical assessments in clinical medicine as it is performed in real life situations.

Results of students were grouped together per discipline and not displayed per individual student per discipline. The aim of this paper was however not to look at individual students or assessments, but at a collective.

CONCLUSION

The reliability of pass/fail outcome decisions in clinical assessments in the undergraduate medical programme involved in this study was found to be good. The necessity of end-of-year assessment after a comprehensive end-of-block assessments may be questioned. The test reliability, as well as stability and reproducibility of individual student marks, were less acceptable. The use of MCQs and OSCEs are practical examples of where the number of assessments can be increased to improve reliability. In order to increase the number of assessments and to reduce the stress of high-stake assessment more workplace-based assessment with observed clinical cases can be recommended.

REFERENCES

1. Amin Z, Seng CY, Eng KH (eds). Practical guide to medical student assessment. 2006. Singapore: World Scientific Publishing.

2. Hays RB, Hamlin G, Crane L. Twelve tips for increasing the defensibility of assessment decisions. Med Teach. 2015;37(5):433-436.

3. Möltner A, Tımbıl S, Jünger J. The reliability of the pass/fail decision for assessments comprised of multiple components. Soc Behav Sci. 2015;32(4):Doc42. doi:10.3205/zma000984

4. Najimi A, Sharifirad G, Amini MM, Meftagh SD. Academic failure and students' viewpoint: The influence of individual, internal and external organizational factors. JEHP. 2013;2:22. http://doi.org/10.4103/2277-9531.112698

5. Bates DW, Slight SP. Medication errors: What is their impact? Mayo Clin Proc. 2014;89(8):1027–1029. doi:10.1016/j.mayocp.2014.06.014

6. Swaminath G, Raguram R. Medical errors I: The problem. Indian J Psychiatry 2010;52(2):110–112. http://doi.org/10.4103/0019-5545.64580

7. Miller GE. The assessment of clinical skills/ competence/ performance. Acad Med. 1990;65:S63– 67. doi:10.1097/00001888-199009000-00045

8. SAQA (South African Qualification Authority) 2015. National policy and criteria for designing and implementing assessment for NQF qualifications and part-qualifications and professional designations in South Africa. SAQA. http://www.saqa.org.za/docs/pol/2015/National%20Policy%20for%20Assessment.pdf (accessed 12 September 2019).

9. Yudkowski R, Park YS, Downing SM. Introduction to assessment in health professions education. In R Yudkowski, YS Park, SM Downing (eds). Assessment in health professions education. 2nd ed. 2019. New York: Routledge.

10. Räisänen M, Tuononen T, Postareff L, Hailikari T, Virtanen V. Students' and teachers' experiences of the validity and reliability of assessment in a bioscience course. High Educ Stud. 2016;6(4):181–189. http://dx.doi.org/10.5539/hes.v6n4p181

11. Van der Vleuten CPM. The assessment of professional competence: Developments, research and practical implications. Adv Health Sci Educ. 1996;1:41–67. https://doi.org/10.1007/BF00596229

12. Peck C. Principles of sound assessment practice in health professions education. ECPP. 2017;5(5):150-157.

13. Van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: from methods to programmes. Med Educ. 2005;39(3):309–317. doi:10.1111/j.1365-2929.2005.02094.x

14. Katzenellenbogen J, Joubert G. Data collection and measurement. In R Ehrlich, G Joubert (eds). Epidemiology - A research manual for South Africa. 3rd ed. 2014. Cape Town: Oxford University Press.

15. Pangaro L, Ten Cate O. Frameworks for learner assessment in medicine: AMEE Guide No. 78. Med Teach. 2013;35(6):e1197–e1210. doi:10.3109/0142159X.2013.788789.

16. Williams R, Klamen D, Mcgaghie W. Cognitive, social and environmental sources of bias in clinical performance ratings. Teach Learn Med. 2003;15:270–292. doi:10.1207/S15328015TLM1504_11.

17. Manterola C, Grande L, Otzen T, García N, Salazar P, Quiroz G. Reliability, precision or reproducibility of the measurements. Methods of assessment, utility and applications in clinical practice. Rev Chilena Infectol. 2018;35(6):680-688. https://dx.doi.org/10.4067/S0716-10182018000600680

18. Pietersen J, Maree K. Standardisation of a questionnaire. In K Maree (ed.). First steps in research. 7th Impression. 2016. Pretoria: Van Schaik Publishers.

19. Bartlett JW, Frost C. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. UOG. 2008;31:466–475. doi:10.1002/uog.5256

20. Tisi J, Whitehouse G, Maughan S, Burdett N. A review of literature on marking reliability research. Report for Ofqual. Slough: NFER. 2013. https://www.nfer.ac.uk/publications/mark01/mark01.pdf. (accessed 12 September 2019).

21. Kane M. Validity and fairness. Language Testing 2010;27(2):177–182. doi:10.1177/0265532209349467

22. Gipps C. Fairness in assessment. In C Wyatt-Smith, JJ Cumming (eds). Educational assessment in the 21st century. 2009. Dordrecht: Springer.

23. Gugiu C, Gugiu M. Determining the minimum reliability standard based on a decision criterion. J Exp Educ. 2018;86(3):458-472. doi:10.1080/00220973.2017.1315712

24. Stoker HW, Impara JC. 7 Basic psychometric issues in licensure testing. In JC Impara (ed.). Licensure testing: purposes, procedures, and practices, 1995. 12. pp. 167–186. Lincoln, NE: Buros. http://digitalcommons.unl.edu/buroslicensure/12. (accessed 12 September 2019).

25. Heale R, Twycross A. Validity and reliability in quantitative studies. Evid-Based Nurs. 2015;18:66-67. http://dx.doi.org/10.1136/eb-2015-102129

26. Sauro J. 2015. How to measure the reliability of your methods and metrics. Measuring U website. https://measuringu.com/measure-reliability/. (accessed 12 September 2019).

27. UFS (University of the Free State). Faculty of Health Sciences rule book. School of Medicine. Undergraduate qualifications 2019. https://apps.ufs.ac.za/dl/yearbooks/335_yearbook_eng.pdf. (accessed 12 September 2019).

28. Xu S, Lorber MF. Interrater agreement statistics with skewed data: Evaluation of alternatives to Cohen's kappa. J Consult Clin Psychol. 2014;82(6):1219.

29. Holley JW, Guilford JP. A note on the G index of agreement. EPM. 1964;24:749–753. doi:10.1177/001316446402400402

30. Opposs D, He O. The reliability programme. Final report. 2011. Ofqual. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/578899/2011-03-16-the-reliability-programme-final-report.pdf. (accessed 12 September 2019).

31. McHugh ML. Interrater reliability: the kappa statistic. Biochem Medica. 2012;22(3):276–82.

32. Sim J, Wright CC. The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. Phys Ther. 2005;85(3):257–268. https://doi.org/10.1093/ptj/85.3.257

33. Attali Y. Effort in low-stakes assessments: What does it take to perform as well as in a high-stakes setting? Educ Psychol Meas. 2016;76(6):1045–1058.

34. Beilock SL, Carr TH. On the fragility of skilled performance: What governs choking under pressure? J Exp Psychol Gen. 2001;130:701–725.

35. Sandi C. Stress and cognition. Wiley Interdiscip Rev Cogn Sci. 2013;4(3):245–261. https://doi.org/10.1002/wcs.1222

36. Patrício MF, Julião M, Fareleira F, Carneiro AV. Is the OSCE a feasible tool to assess competencies in undergraduate medical education? Med Teach. 2013;35(6):503-514. doi:10.3109/0142159X.2013.774330

37. Brits H, Bezuidenhout J, Van der Merwe LJ. A framework to benchmark the quality of clinical assessment in a South African undergraduate medical programme. S Afr Fam Pract. 2020;62(1), a5030. https:// doi.org/10.4102/safp. v62i1.5030

38. Khan KZ, Gaunt K, Ramachandran S, Pushkar P. The objective structured clinical examination (OSCE): AMEE Guide No. 81. Part II: organisation & administration. Med Teach. 2013;35:e1447-63. doi:10.3109/0142159X.2013.818635

39. Ponnamperuma GG, Karunathilake IM, McAleer S, Davis MH. The long case and its modifications: a literature review. Med Educ. 2009;43:936–941. doi:10.1111/j.1365-2923.2009.03448.x

40. Kamarudin MA, Mohamad N, Awang MN, Siraj BHH, Yaman MN. The relationship between modified long case and objective structured clinical examination (OSCE) in final professional examination 2011 held in UKM Medical Centre. Procedia Soc Behav Sci. 2012;60:241-248. doi:10.1016/j.sbspro.2012.09.374

41. Wass V, Jolly B. Does observation add to the validity of the long case? Med Educ. 2001;3:729–734. doi:10.1046/j.1365-2923.2001.01012.x

42. Wass V, Jones R, Van der Vleuten C. Standardized or real patients to test clinical competence? The long case revisited. Med Educ. 2001;35(4):321–325. https://doi.org/10.1046/j.1365-2923.2001.00928.x

43. Norman G. Postgraduate assessment - reliability and validity. Transactions of the Colleges of Medicine of South Africa. 2003;47:71-75.

44. Liu C. An introduction to workplace-based assessments. Gastroenterol Hepatol Bed Bench. 2012;5(1):24–28.

45. Clauser BE, Margolis MJ, Swanson DB. Issues of validity and reliability for assessments in medical education. In ES Holmboe, SJ Durning, RE Hawkins (eds). Practical guide to the evaluation of clinical competence. 2nd ed. 2018. Philadelphia: Elsevier.

CHAPTER 6:

ARTICLE 5 – QUALITY ASSESSMENT IN UNDERGRADUATE MEDICAL TRAINING: HOW TO BRIDGE THE GAP BETWEEN WHAT WE DO AND WHAT WE SHOULD DO

The article was prepared according to the submission guidelines of the *Pan African Medical Journal*.

Journal:	Pan African Medical Journal
Publisher:	The Manuscript Hut
Country of publication:	Kenya
ISSN:	1936-8688 (ONLINE)
DOI:	10.11604/pamj.2020.36.79.23658

Citation:

Brits, H., Bezuidenhout, J. & Van der Merwe, L.J. 2020. Quality assessment in undergraduate medical training: how to bridge the gap between what we do and what we should do. Pan African Medical Journal. 2020;36:79. DOI: 10.11604/pamj.2020.36.79.23658

Submitted:	21 May 2020
Accepted:	27 May 2020
Published:	9 June 2020


The Pan African Medical Journal
May - August, 2020
Volume 36
ISSN: 1937 – 8688
An Open Access Journal
PAMJ.ORG
Unleashing African Research
PanAfrican Medical Journal

Overview

ISSN

1937-8688 (ONLINE)

Focus and scope

We believe that scientific work done in Africa should be rapidly and freely made available to all researchers worldwide. Aim: To create, stimulate and perpetuate a culture of information sharing and publishing amongst researchers and other health actors of the African health scene in ways that will contribute to availability of health information, better understanding of Africa specificities and overall, to improve health outcomes of the people on the continent. Scope: We publish original scientific research on clinical, public health, social, political, economic and all other factors affecting the health of populations in Africa.

Managing Editors

Editor´s pick

Research

Quality assessment in undergraduate medical training: how to bridge the gap between what we do and what we should do

Hanneke Brits et al.

09 Jun 2020

PanAfrican Medical Journal

Open Access

Research

Quality assessment in undergraduate medical training: how to bridge the gap between what we do and what we should do

Hanneke Brits[1,&], Johan Bezuidenhout[2], Lynette Jean Van der Merwe[3]

[1]Department of Family Medicine, School of Clinical Medicine, University of the Free State, Bloemfontein, South Africa, [2]Health Sciences Education, Faculty of Health Sciences, University of the Free State, Bloemfontein, South Africa, [3]Undergraduate Programme Management, School of Clinical Medicine, University of the Free State, Bloemfontein, South Africa

[&]Corresponding author: Hanneke Brits, Department of Family Medicine, School of Clinical Medicine, University of the Free State, Bloemfontein, South Africa

Keywords: Quality assessment, focus group interview, clinical competence

Domain: Health education

Received: 21 May 2020 - Accepted: 27 May 2020 - Published: 09 Jun 2020

Abstract

Introduction: the outcome of the undergraduate medical training programme in South Africa is to produce competent medical doctors who can integrate knowledge, skills and attitudes relevant to the South African context. Training facilities have a responsibility to ensure that they perform this assessment of competence effectively and defend the results of high-stakes assessments. This study aimed to obtain qualitative data to suggest practical recommendations on best assessment practices to address the gaps between theoretical principles that inform assessment and current assessment practices. **Methods:** a focus group interview was used to gather this data. The teaching and learning coordinators for five of the six modules that are offered in the clinical phase of the undergraduate medical programme participated in the focus group interview. The focus group interview proceeded as planned and took 95 minutes to complete. The responses were transcribed and recorded on a matrix. **Results:** the lack of formal feedback to students was identified as an area of concern; feedback plays an important role to promote student learning and improve patient care. The role of teaching and learning coordinators as drivers of quality assessment were recognized and supported. All participants agreed on the outcome of the programme and the central role of the outcome in all assessments. **Conclusion:** the training of assessors and the implementation of workplace-based assessment and assessment portfolios were recommended and can also address feasibility challenges. Participants recommended decreasing summative assessments and only performing these for borderline students.

Research | Volume 36, Article 79, 09 Jun 2020 | 10.11604/pamj.2020.36.79.23658

This article is available online at: http://www.panafrican-med-journal.com/content/article/36/79/full/

Introduction

Quality assessment requires that the type and content of the assessment is aligned with the outcome of the training programme.[1] The outcome of the undergraduate medical training programme in South Africa is to produce competent medical doctors who can integrate knowledge, skills and attitudes relevant to the South African context.[2] Assessment of clinical competence is a complex process, due to a number of factors, which include the constant emergence of new best-practice medical evidence,[3] the theory-practice gap between what is taught and what is observed in clinical practice[4-6] what is feasible,[7] and the challenges of assessment in real-life situations that may compromise the reliability of the assessment.[8]

Competence assessment must satisfy various stakeholders, which include patients and the general public, training providers, regulatory bodies and students. Training facilities have a responsibility to ensure that they perform this assessment task effectively and can defend the results of high-stakes assessments.[9]

A paper describing a framework to benchmark the quality of clinical assessment in a South African undergraduate medical programme, provides context-specific theoretical principles for undergraduate medical assessment.[10] Assessment reports and quantitative studies (In press) on current assessment practices used for undergraduate medical students at the University of the Free State (UFS) showed that these principles are not always adhered to, which may compromise the defensibility of high-stakes assessments.

This study aimed to obtain qualitative data to suggest practical recommendations on best assessment practices to address the gaps between theoretical principles, that inform assessment, and current assessment practices. These recommendations will be combined with other research results to prepare a proposal to inform quality assessment at the UFS.

Methods

Research design

A focus group interview (FGI) was used to triangulate theory (i.e. theoretical principles that inform assessment) with current assessment practices, to compile recommendations that should assist with quality assessment in undergraduate medical training. A FGI can be used in a mixed-methods design to triangulate qualitative and quantitative data from different sources,[11] as was done in this study. Various definitions exist for an FGI, and some researchers even use the terms FGI and focus group discussion (FGD) interchangeably.[12] The difference between an FGI and an FGD is that the main objective of an FGI is to obtain answers to specific questions while, in an FGD, the interaction between the group members and the group dynamics are as important as the information gathered.[12,13]

Merton and Kendall (in Cohen et al.)[14] first described the concept of an FGI in 1946 and concluded that:

- During an FGI, there is a greater degree of interviewer control;
- The people participating in the interview should share experiences;
- The interview questions are based on previous data analysis; and
- Subjective experiences of people who have been exposed to the same experience are gathered.

The strength of a focus group is that it stimulates new or forgotten ideas and that members can build on the input of others. Some of its limitations are that it can be difficult to get members together, the group may not be representative, and some group members may dominate others.[14-15]

Participants

In an FGI, between five and 12 members interact, debate and argue their opinions on a specific issue. The participants of the focus group should represent the target population. Members that participate should do so voluntarily, should be knowledgeable on the subject and able to communicate in a group.[11] The clinical phase at the UFS comprises six modules. The six teaching and learning (T&L) coordinators of these

modules were invited to participate in the FGI. Five of these T&L coordinators participated in the FGI.

Facilitator

The facilitator asks specific questions with the view to obtain answers to specific questions.[13] It is important for the facilitator to monitor the group dynamics and ensure participation by all members. The facilitator must be in control of the situation and should avoid too much or too little personal participation.[12] A facilitator with experience in higher education and in conducting FGIs was used to facilitate the process.

Questions

An FGI is not merely a general discussion, but is focused on a specific topic. Usually, the discussion starts broadly and, then, spirals inwards to address the research question/s.[16] The questions asked during this FGI derived from an assessment framework for undergraduate medical programmes,[10] as well as the results of current assessment practices (In press) and publications with recommendations for undergraduate medical assessment.[1,2,9,17] The guidelines for developing "good focus group questions", which include that the questions must be short, clear, open-ended and directional, as described by Krueger and Casey,[18] were followed. Questions were categorised and grouped. All the questions were available in the facilitator and participant guides which the facilitator and participants received before the FGI.

Logistics

An FGI should last between 60 and 90 minutes.[19] To capture all the information, the facilitator needs to take notes of the discussions and non-verbal cues. It can be helpful to record or videotape the discussion, and to use a co-facilitator to take notes and write down observations too.[12] The researcher arranged a neutral venue, confirmed the availability of the facilitator and participants and provided refreshments. The facilitator received all the necessary documents well in advance of the FGI. The researcher met with the facilitator in person about the process to be followed and to clarify uncertainties and agreed on the process. All participants received a participant guide one week before the FGI and a reminder to attend one day before the FGI was conducted on 29 January 2020.

Data collection

The aim of an FGI is not consensus, but rather the gathering of rich ideas.[11] The facilitator asked one question at a time and encouraged active participation by all participants. Discussions continued until all participants were satisfied with the answer to a particular question. If no answer or more than one answer or suggestion were offered, the facilitator encouraged participation until no new ideas were produced. More than one answer or disagreement between opinions were allowed.

Pretesting of focus group and explorative interview

No test run of the FGI was done, as it is important to obtain the collaborative feedback of the whole group. The validity of the questions asked in the FGI was discussed in an explorative interview with the promotors, and was based on previous experience of the researcher.

Analysis of data and reporting

An audio recording of the FGI was transcribed by the researcher immediately after the FGI concluded. The researcher used a video recording to verify the accuracy of the transcription. A matrix, as suggested by Onwuegbuzie et al.,[20] was used to transfer the answers of the specific questions. Data were reported under specific categories and questions. The audio recording was used again to verify the information on the matrix.

Ethical considerations

Ethical approval for the study was obtained from the Health Sciences Research Ethics Committee, UFS (UFS-HSD 2019/0001/2304). UFS authorities approved the inclusion of personnel. Informed consent was obtained from participants for participation and for making the audio and video recordings. Participants were not identified and a participant number was allocated to each, which is also used for data reporting.

Quality and rigour of the data management

To ensure the credibility of the data collection, all the research questions were clarified with the promotors. The facilitator ensured active participation by all participants, and clarified concepts to improve the quality of the data. Local, national and international assessment guidelines were included to make the recommendations transferable to other institutions.

The focus group participants and interview process were clearly described for the purpose of assessing the dependability of the results. Confirmability was ensured by audio and video recording of the process and verifying results after completion of the result template.

Results

The T&L coordinators for five of the six modules that are offered in the clinical phase of the undergraduate medical programme attended the FGI. The process proceeded as planned and took 95 minutes to complete. The audio recording was of good quality, with all conversations clearly audible and respondents identifiable. The participants provided answers to all the questions in the FGI, and all participants contributed and gave original suggestions and participated in the discussions. No participant dominated or withdrew during discussions.

The results of the focus group interview are displayed in three tables according to the adjusted template suggested by Onwuegbuzie et al.[20]

In table 1 the results for the outcome of the programme, competence, validity and reliability are displayed. Table 2 addresses the results for fairness, feasibility, educational effect and assessment methods and Table 3 quality assurance, training and general comments.

Discussion

The FGI met the requirements for a good FGI regarding participants, the facilitator, the questions, logistics, explorative interview and data collection and analysis. The results are also representative of the study population, with five of the possible six participants included.

The first question was around the outcome of die undergraduate medical programme. All the participants agreed with the outcome as is, namely, to produce a competent medical doctor who can integrate knowledge, skills and attitudes relevant to the South African context. This clear outcome should be kept in mind during all assessments. This outcome is in line with the regulations stipulated in the Health Professions Act of South Africa, the South African Quality Assurance Authority and the assessment policy of the UFS.[2,17,21]

The next questions focused on competence and the way it is assessed. Clinical competence must be assessed on the "Does" level, according to Miller's pyramid.[22] It was mentioned that the actual demonstration of this competence only occurs during internship, which is still part of training (students must complete internship and community service before registering as independent medical practitioners with the HPCSA). A suggestion to implement pass/fail stations and not only an average of 50% or above to pass, was well accepted. A discussion on the difficulty to ensure competence with a pass mark of 50% (the pass mark according to the UFS assessment policy) provided more questions than answers. It must be recognised that a mark of 50% indicate that the student is competent and not "half competent". All assessors should be aware of how they allocate marks and the implication thereof. Further discussion in this regard was recommended to clarify the meaning of 50% in the context of competence.

During questions regarding validity, good practices were shared and recommendations made. It was agreed that T&L coordinators should take responsibility for assessments, to ensure the validity of assessments. Blueprinting of all assessments should be done. Blueprinting will improve content validity, and using appropriate assessment methods will improve construct validity.[10] There is no need to add additional assessment methods, as most assessment methods described for undergraduate clinical assessment[23] are currently used at the UFS. It was recognized that a shortages in the workforce favour the use of less labour-intensive assessment methods, e.g., multiple-choice questions rather than longer written questions that can assess higher cognitive levels. The lack of trained assessors also limits the use of workplace-based assessment (WBA) and assessment/competency portfolios to assess competence. To address the workforce issue, all clinicians should be trained as assessors, and registrars can be included in the assessment process. By including registrars, they are trained on the important skill of assessment, and it may help to spread the workload. Regarding the assessment of professionalism and "soft skills," the suggestion to implement a "professionalism

portfolio" and implement the graduate attributes policy of the university were supported and should be investigated.

The participants gave valuable input on aspects to improve the quality of assessment, including recommendations on reliability, fairness, educational effect and feasibility. Competency assessment cannot be 100% reliable, but the suggestions to use WBA and assessment/ competency portfolios were recommended to increase the number of assessments. WBA and assessment portfolios are excellent ways to assess competence, but reliability may be compromised.[24] Although portfolios and WBA are labour intensive, these methods are more authentic and the number and type of assessments can increase, thereby contributing to reliability.[25]

The lack of formal feedback to students was identified as an area of concern – feedback plays an important role to promote student learning and improve patient care.[26] Feedback is also a requirement stipulated in the assessment policy of the UFS.[21] The scheduling of formal feedback sessions after assessments may assist with the implementation of formal feedback, a practice that is currently lacking in the undergraduate medical programme.

Participants in the FGI recommended decreasing summative assessments and only performing these for borderline students. This practice will also address some of the problems with the feasibility of summative assessments. Less emphasis on summative assessment is well supported in the literature e.g. assessment results should not depend on a single summative assessment, as competency in one case is a poor predictor of competency in another.[27] Performance stress during high-stakes assessments may also contribute to less reliable outcomes,[28] and a single poor performance should not affect the outcome of years of training.[23]

The lack of post-assessment moderation was identified as a risk for quality assessment. Although procedures and checklists for moderation are available, the implementation is not standard practice in all departments. Quality assurance and moderation are important components of ensuring and maintaining the quality of assessment.[21] An e-mail to remind departments to do moderation, and spot checks, may reinforce the implementation of this important practice.

During the FGI clinical training was also discussed in relation to assessment. Biggs[29] describes the term constructive alignment as comprising outcomes, teaching and training activities and assessment that are planned to complement and support each other. Students indicated in their feedback before the FGI that they want more on-site practical training in wards and clinics (In press). The increase in student numbers and decrease in teacher numbers also decreases supervised, hands-on practical training for students. A suggestion for countering the lack of clinical exposure is to stipulate clearly and monitor available clinical training time. Another factor that affects clinical training negatively is overburdened clinicians, who may not necessarily be good role models and tend to give students time off, so that the clinicians can get clinical work done, rather than spend time on training. This practice may be due to burnout, as evidenced by a study in this academic setting that showed that only 3.4% of the doctors included in the study showed no signs of burnout.[30] The participants mentioned the importance of developing core competencies in undergraduate students, such as professionalism, leadership and scholarship,[31] how to cope in difficult situations, and practicing self-care which should be included in clinical training. The training platform may be an opportunity for students to see how to behave professionally, but also how not to behave. It was discussed that students may not be aware that, although they are trained in tertiary facilities, they are not expected to perform as specialists, but that they should rather use the opportunity to identify clinical signs and develop an approach to a specific problem. Better communication on the outcome of specific training rotations may assist both students and clinicians and was recommended.

The FGI concluded with a discussion on the effect of the introduction of T&L coordinators on student assessment and training. The excellent work of the T&L coordinators was recognised and appreciated. All agreed that the T&L coordinators should continue to play a leading role in student assessment and training.

Limitations and strengths

Only the T&L coordinators of the major disciplines participated in the FGI, and the FGI may have failed to capture contributions by excluding minor disciplines. However, these

smaller disciplines were indirectly represented by the major disciplines.

Strengths of the FGI were that the FGI was conducted according to the planning, and within the guidelines for a FGI, as described in the methods, and that data management met the criteria for credibility.

Conclusion

The clear, agreed-upon outcome, namely, to produce a competent medical doctor who can integrate knowledge, skills and attitudes relevant to the South African context, should be kept in mind during all assessments. The difficulty of how to measure and allocate marks to competence was recognised. The lack of formal feedback to students and blueprinting should be addressed. The important place of WBA and assessment portfolios, with less emphasis on summative assessment were important recommendations from the FGI. A proposal to improve the quality of assessment in the clinical phase of the undergraduate medical programme will be compiled from this and other research information. This proposal will be submitted to the Executive Committee of the School of Clinical Medicine for implementation. Finally, an FGI can be recommended as an appropriate way to get rich data for practical solutions.

What is already known about assessment

- Assessment should be aligned with the outcome of the training programme
- Assessment of clinical competence is a complex process

What this study adds

- Workplace-based assessment should form part of competency assessment
- The difficulty of how to measure and allocate marks to competence was recognised
- Competency and professional portfolios should be implemented

Competing interests

The authors declare no competing interest

Authors' contributions

HB – Conceptualisation of study, protocol development, data collection and writing of paper, JB and LJvdM – Promotors who assisted with conceptualisation and planning of the study, as well as critical evaluation and final approval of the manuscript.

Acknowledgements

Prof Mathys Labuschagne for facilitating the focus group interview
Mrs Hettie Human for language editing

Tables

Table 1: results of the focus group interview displayed for outcome of programme, competence, validity and reliability adjusted according to the template by Onwuegbuzie et al.

Table 2: results of the focus group interview displayed for fairness, feasibility, educational effect and assessment methods adjusted according to the template by Onwuegbuzie et al.

Table 3: results of the focus group interview displayed for quality assurance, training and general comments adjusted according to the template by Onwuegbuzie et al.

References

1. John Norcini, Brownell Anderson, Valdes Bollela, Vanessa Burch, Manuel João Costa, Robbert Duvivier *et al.* Criteria for good assessment: consensus statement and recommendations from the Ottawa 2010 Conference. Med Teach. 2011;33(3):206-14. PubMed | Google Scholar

2. Published under Government Notice R139 in Government Gazette 31886 of 19 February 2009. Regulations Relating to the Registration of Students, Undergraduate Curricula and Professional Examinations in Medicine.

3. Cooke SJ, Johansson S, Andersson K, Livoreil B, Post G, Richards R, Stewart R, Pullin AS. Better evidence, better decisions, better environment: emergent themes from the first environmental evidence conference. Environmental Evidence. 2017;6(1):15. Google Scholar

4. Ajani K, Moez S. Gap between knowledge and practice in nursing. Procedia - Social and Behavioral Sciences. 2011;15:3927-31. Google Scholar

5. Hussein MH, Osuji J. Bridging the theory-practice dichotomy in nursing: The role of nurse educators. J Nurs Educ Pract. 2017;7(3):20-2. Google Scholar

6. Salah AA, Aljerjawy M, Salama A. Gap between theory and practice in the nursing education: The role of clinical setting. Emergency. 2018;24:17-18. Google Scholar

7. Zlatkin-Troitschanskaia O, Pant HA. Measurement advances and challenges in competency assessment in higher education. Journal of Educational Measurement. 2016; 53(3):253-264. Google Scholar

8. Clauser BE, Margolis MJ, Swanson DB. Issues of validity and reliability for assessments in medical education. In ES Holmboe, SJ Durning, RE Hawkins (Eds). Practical guide to the evaluation of clinical competence. Philadelphia, Elsevier. 2018;2nd ed.

9. Richard B Hays, Gary Hamlin, Linda Crane. Twelve tips for increasing the defensibility of assessment decisions. Med Teach. 2015;237(5): 433-436. PubMed | Google Scholar

10. Hanneke Brits, Johan Bezuidenhout, Lynette J Van der Merwe. A framework to benchmark the quality of clinical assessment in a South African undergraduate medical programme. S Afr Fam Pract (2004). 2020 Feb 4;62(1):e1-e9. PubMed | Google Scholar

11. Carey MA, Asbury JE. Focus group research. New York, Taylor and Francis Group, Routledge. 2016. Google Scholar

12. Nyumba T, Wilson K, Derrick CJ, Mukherjee N. The use of focus group discussion methodology: Insights from two decades of application in conservation. Methods in Ecology and Evolution. 2018;9(1):20-32. Google Scholar

13. Boddy C. A rose by another name may smell as sweet but "group discussion" is not another name for a "focus group" nor should it be. Qualitative Market Research: An International Journal. 2005;8(3):248-255. Google Scholar

14. Cohen L, Manion K, Morrison K. Research methods in education. New York, Taylor and Francis Group, Routledge. 2002;6th ed:317-382. Google Scholar

15. Michael D Fetters, Timothy C Guetterman, Debra Power, Donald E Nease Jr. Split-session focus group interviews in the naturalistic setting of family medicine offices. The Annals of Family Medicine. 2016;14(1):70-75. PubMed | Google Scholar

16. Nieuwenhuis J. Qualitative research design and data gathering techniques. In Maree, K. (Ed). First steps in research. 7th Impression. 2016 Pretoria. Van Schaik Publishers: 70-97.

17. South African Qualifications Authority (SAQA), Sabinet Online. **National policy and criteria for designing and implementing assessment for NQF qualifications and part-qualifications and professional designations in South Africa.** Accessed 20th January 2020.

18. Krueger RA, Casey MA. Focus Groups. A practical guide for applied research. Singapore, Sage Publications. 2015;. 5th ed:39-76. **PubMed** | **Google Scholar**

19. Skinner D. Qualitative research methodology: An introduction. In Ehrlich, R & Joubert, G (Eds). Epidemiology - A research manual for South Africa, Cape Town, Oxford University Press. 2014;3rd ed:349-359.

20. Onwuegbuzie AJ, Dickinson WB, Leech NL, Zoran AG. A qualitative framework for collecting and analyzing data in focus group research. International Journal of Qualitative Methods. 2009;8(3):1-21. **PubMed** | **Google Scholar**

21. University of the Free State (UFS). **Assessment policy on the UFS coursework learning programme.** Accessed 20th January 2020.

22. GE Miller. The assessment of clinical skills/competence/performance. Academic Medicine. 1990;65 (9):S63-S67. **PubMed** | **Google Scholar**

23. Yudkowski R, Park YS, Downing SM. Introduction to assessment in health professions education. In Yudkowski R, Park YS, Downing SM (Eds). Assessment in health professions education. 2nd ed. 2019. New York: Routledge.

24. Cees P M van der Vleuten. Revisiting Assessing professional competence: From methods to programmes. Med Educ. 2016 Sep;50(9):885-8. **PubMed** | **Google Scholar**

25. Schumacher DJ, Tekian A, Yudkowsky R. Assessment portfolios. In Yudkowski R, Park YS, Downing SM (Eds). Assessment in health professions education. 2nd ed. 2019. New York: Routledge.

26. Marjan Govaerts. Workplace-based assessment and assessment for learning: Threats to validity. Journal of Graduate Medical Education. 2015;7(2):265-267. **PubMed** | **Google Scholar**

27. Amin Z, Seng CY, Eng KH (Eds). Practical guide to medical student assessment. Singapore: World Scientific Publishing. 2006. **Google Scholar**

28. Attali Y. Effort in low-stakes assessments: What does it take to perform as well as in a high-stakes setting? Educ Psychol Meas. 2016;76(6): 1045-1058. **Google Scholar**

29. Biggs JB. Enhancing teaching through constructive alignment. Higher Education. 1996;32:347-364. **Google Scholar**

30. Sirsawy U, Steinberg WJ, Raubenheimer JE. Levels of burnout among registrars and medical officers working at Bloemfontein public healthcare facilities in 2013. South African Family Practice 2016; 58(6): 213-218. **Google Scholar**

31. HPCSA (Health Professions Council of South Africa). Core competencies for undergraduate students in clinical associate, dentistry and medical teaching and learning programmes in South Africa. Health Professions Council of South Africa. 2014.

Table 1: results of the focus group interview displayed for outcome of programme, competence, validity and reliability adjusted according to the template by Onwuegbuzie et al.

QUESTION	ANSWERS	RESPONDENT				
		1	2	3	4	5
1. OUTCOME OF THE PROGRAMME						
1.1 Do you agree with the outcome of the MBChB (Bachelors of Medicine and Bachelors of Surgery) programme?	1.1.1 Yes	A	A	A	OS	A
	Discussion: A clear outcome is necessary to measure the outcome of any assessment. The outcome is only visible during Internship and maybe it should be tested again then. The outcome of the programme should be kept in mind during all assessments					
1.2 Do you have other suggestions for the outcome?	1.2.1 The International standard should also be included in the definition to make it more global	U	NR	A	OS	A
	Discussion: Although it is broad, it encompasses the important concepts of competence, integration and the relevant context					
2. COMPETENCE						
2.1 Do we assess competence in the final summative assessment in the MBChB programme?	2.1.1 Yes, but I think the bar is set too low and you pass on an average mark and may not be competent in all expected skills.	U	A	A	A	OS
	Discussion: It is difficult to achieve competence before you start to work, and Internship is also part of training and becoming competent. Students are generally competent, but I don't think that we assess competence well enough. 50% is not necessarily a mark that indicates competence. We should ensure that if a student gets 50% that the student is competent and not "half competent", as 50% is the pass mark. This should be discussed at other forums.					
	2.1.2 I think there should be pass/fail stations or you should pass a minimum number of assessments rather than on average.	A	A	OS	A	A
	Discussion: This is a good idea. We should trust our assessment results as we do many different assessments and moderate papers.					
3. VALIDITY						
3.1 To have a valid assessment enough of the content should be assessed. This can be done by blueprinting of all assessments. How can we improve blueprinting of all assessments?	3.1.1 We should start sooner by planning and blueprinting all tests and assessments and not only exams.	NR	A	A	A	OS
	Discussion: I want to emphasize that blueprinting is making life much easier. I luckily inherited the system, but then had to get some training as well. *Information leaflets on Blackboard may help especially with turnover of personnel*					
	3.1.2. All T&L coordinators should do the Health Professions Education assessment course	A	A	A	OS	NR
	3.1.3 T&L coordinators should lead the process in departments to implement blueprinting of all assessments	A	A	OS	A	A
	3.1.4 Informal blueprinting happens, but it should be formalized to have evidence	A	OS	A	A	A
3.2 Students and lecturers think that assessment methods should be improved. What methods can you suggest to improve assessment?	3.2.1. We use all methods and I don't think we need to add methods, however it is difficult with limited workforce.	A	A	OS	A	A
	Discussion: It is easier to use MCQ's than longer questions, although higher cognitive levels may be difficult to assess. The workforce is the problem					
	3.2.2. Continuous and portfolio assessment may be a way to go and we should work towards it.	U	A	A	OS	A
	Discussion: Continuous assessment is problematic, due to assessors not giving the marks or giving 65%. The number of assessments may be beneficial.					
	3.2.3 Longer questions may test concepts better.	OS	A	U	NR	A
	3.2.4.The methods are good, the assessors not always and they may benefit from rubrics and training	A	A	OS	A	A
	Discussion: Assessors need to be trained better, e.g. registrars can be trained in assessment, by allowing them to assess together with a consultant and then discuss the marks. It will benefit both parties and address the work force.					
3.3 According to the Health Professions Council of South Africa (HPCSA) "soft skills and professionalism" should be assessed. How do you suggest that we assess "soft skills" and professionalism throughout the curriculum?	3.3.1 Soft skills are assessed in clinical case presentations, but a specific mark is not allocated to it. We may allocate a specific mark to it.	A	A	OS	NR	A
	3.3.2 In communications stations it can be assessed as well	NR	A	A	OS	NR
	Discussion: Assessments are not normal circumstances and students know how to behave professionally in assessments, however professionalism should be practice and assessed throughout. I like the idea of peer and patient assessment. Peer review may not work, it was tried before. Unprofessional behaviour should be recorded and have consequences e.g. deduction of marks. A "Professional portfolio" for continuous assessment may work. *Facilitator: The university is involved in a programme to promote graduate attributes and we may look how they do it and what is in place from their side.*					
	3.3.3 We should try and latch to the university programme	A	OS	A	A	A
4. RELIABILITY						
4.1 Can/Should all assessments be 100% reliable?	4.1.1 No, it is impossible	A	A	A	A	OS
	4.1.2. But we should try to keep it as reliable as possible, taking the real life situation into account	A	A	A	OS	A
4.2 Which specific measures can be implemented to improve or get reasonable reliability in clinical assessment?	4.2.1 We should use more clinical cases in the workplace (WBA), which is less labour intensive than an exam	U	A	U	OS	A
	Discussion: It is difficult with the limited resources. Assessment rubrics may make it easier to implement, as well as simulated scenarios					

Table 2: results of the focus group interview displayed for fairness, feasibility, educational effect and assessment methods adjusted according to the template by Onwuegbuzie et al.

5. FAIRNESS						
5.1 How can we improve the alignment between outcomes, training and assessment?	5.1.1 Lecturers should be asked to update "outcomes" yearly in line with clinical practice and assessment experience, before the new groups start.	NR	A	OS	A	A
	5.1.2 Student feedback of the module should also be considered	A	A	NR	A	OS
	Discussion: The outcomes should be a framework, rather than specific. This is tricky, because we need specifics to blueprint. The students struggle with the transition between pre-clinical with specific outcomes and clinical training with broader outcomes. They need to mature in this regard. These are senior students and we should not spoon feed them. They must be able to integrate and think rather than concentrating on detail.					
	5.1.3 T&L coordinators must facilitate the process to ensure alignment and fairness	A	A	A	A	OS
6. FEASIBILITY						
6.1 Which resources do you take into account when planning individual and overall assessments?	6.1.1 The basics are assessors, timing and patients. The numbers are calculated according to the number of students.	A	A	OS	A	A
	Discussion: This is more difficult with the addition of training sites, increased student numbers and the Nelson Mandela Fidel Castro Medical Programme (NMFCMP) students.					
6.2 How important is each one of those?	6.2.1 The assessors, timing and patients are most important.	A	A	A	OS	A
	6.2.2 Recently finances must also be considered.	NR	A	OS	NR	A
	Discussion: Patients must come in for exams and also wants compensation, transport money and food. Travel and/or accommodation of External examiners must also be budgeted for. This adds up to a substantial amount. Less summative assessment may help with resources.					
7. EDUCATIONAL EFFECT						
7.1 Feedback is one of the most important aspects of learning. What strategies can be used to ensure effective feedback?	7.1.1 Logistically it is difficult because students start in a new rotation. It may help if a specific session is scheduled on the time tables, say 2 weeks into the new rotation.	A	A	A	A	OS
	7.1.2 Electronic feedback to the group via e-mail or on Black board	OS	A	A	OS	A
	Discussion: Information can include the class average and highlighting of problem areas.					
	7.1.3.Appointments with individual students who struggled with the assessment	A	OS	A	A	A
	7.1.4 Open door policy to come and discuss the assessment with the T&L coordinator, as is currently the practice	A	A	A	A	OS
	7.1.5 Immediate feedback after clinical cases to highlight strengths and areas that need improvement.	OS	A	A	A	A
7.2 How should feedback be given on multiple choice questions?	7.2.1 This is difficult, because we don't want to compromise our databank. However general feedback is given on problem areas after the assessment.	OS	A	A	A	A
	Discussion: Students want the answers, rather than the knowledge and therefor general feedback is given.					
	7.2.2 The students may re-write the test under exam conditions and then the answers are discussed	A	A	A	OS	NR
	7.2.3 Poor performers may come and have a look at their paper in order to identify the root of the problem.	U	OS	U	A	A
8. TYPES OF ASSESSMENT						
8.1 Students and some lecturers suggested only end-of-block assessments and if students pass they need not do a summative assessment at the end of the year again. How do you feel about this suggestion?	8.1.1 It is a good idea and practiced at other universities.	NR	A	A	A	OS
	Discussion: It motivates students to work hard during rotations. A pass mark of 60% was agreed upon. This will ensure that borderline candidates can be assessed again during summative assessment.					
8.2 What is your opinion on a single integrated assessment?	8.2.1 I fully support it, like the Family Medicine OSCE and then all can contribute to the assessment	NR	OS	A	A	A
8.3 With guidelines for good assessment practices in mind, how can we improve our current assessment regarding Workplace based assessment (WBA)?	8.3.1 This is the ideal way forward and we must try to implement it, despite workforce problems.	U	A	A	OS	A
	Discussion: This is the best place to assess real life competence. The students, patients and assessors are there. Peer assessment may be problematic. The use of a "competency portfolio" was suggested and supported.					
8.4 How can we improve formative assessment?	8.4.1We should try to assess and record more student patient encounters	A	A	A	A	OS
	Discussion: This will increase the number and the reliability of assessments. You know which students are competent when you work with them. You can also assess professionalism better. The competency portfolio was mentioned again.					
8.5. What is your opinion on summative assessment?	Although more assessments are good learning opportunities, I think we must try and reduce summative assessment to only borderline candidates.	A	OS	A	A	A
	Discussion: Students may be disadvantaged during their first rotation, because they gain experience and competence throughout the year. However all will have first rotations and it is therefore fair.					

Table 3: results of the focus group interview displayed for quality assurance, training and general comments adjusted according to the template by Onwuegbuzie et al.						
9. QUALITY ASSURANCE						
9.1 What strategies should be implemented to ensure compliance with the UFS pre- and post-assessment moderation practices?	9.1.1 Although it is more work a moderation checklist should be implemented for all assessments	A	A	OS	A	A
	Discussion: It is done informally, without any evidence when needed, therefore it should be formalized. Quality assurance helps to improve assessments and maintain standards.					
10. TRAINING						
10.1 Students suggested more exposure to patients in wards and clinics and less in the classroom. What are your suggestions to improve clinical training?	10.1 1. With more students (and less lecturers) the direct student exposure decreases. Time at training sites should be stipulated and controlled.	OS	A	A	A	A
	Discussion: Most students want to go home as soon as possible, stating that they want to study, which is contradicting what they suggested.					
10.2 Students also suggested more exposure to good clinical role models. What is your response to this?	10.2.1 All clinicians are not necessarily good role models, but students can also learn from the "not so good" on what not to do.	OS	A	A	A	A
	Discussion: Due to workload many people are suffering from burnout. This must also be discussed with students and the importance of self-care must be re-emphasized. The clinical psychologist can assist and attention should be paid to resilience training. Reflective practice and professionalism must also be addressed. Students get very good support at the UFS to cope with stress.					
10.3 Students want to be assessed on what they see during training, but not at specialist hospitals. What is your response to this?	10.3.1 Students get exposure to all levels of care to expose them to different conditions and clinical signs. Students must know what is expected where, to benefit from the extended training platform.	A	A	A	OS	A
	Discussion: In tertiary hospitals the clinical signs are more obvious. At primary care, they get exposure to the burden of disease in their environment. Students are not expected to make specialist diagnosis, but rather display their approach to a specific symptom or sign. Supervision and space is problematic at primary care clinics. Good communication is important to know what is expected.					
11. GENERAL						
11.1 Any last comments regarding the quality of undergraduate clinical assessment and training?	11.1.1 We, the T&L coordinators do a great job under difficult circumstances	A	A	A	OS	A
	11.1.2 The T&L coordinators made a huge difference to the quality of assessment and training	OS	A	A	A	A

CHAPTER 7:
CONTRIBUTION OF THE STUDY

The contributions of this study are the framework to benchmark clinical assessment (cf. Chapter 2) and a proposal prepared for the Executive Committee: School of Clinical Medicine, Faculty of Health Sciences at the University of the Free State on measures to improve clinical assessment. The proposal was written according to the guidelines of Becker, Schütt and Amini (2014) in *Proposal Writing for International Research Projects: A Guide for Teachers*.

Proposal: Best practice recommendations to enhance quality assessment in the clinical phase of the undergraduate medical programme at the University of the Free State.

The implementation of these recommendations will help to ensure accurate judgement on student performance and to bridge the gap that exists between current assessment practices and prescribed policies and guidelines on quality assessment. Results from this study were published in peer-reviewed journals for others to apply in clinical training facilities and programmes.

Best practice recommendations to enhance quality assessment in the clinical phase of the undergraduate medical programme at the University of the Free State

Faculty of Health Sciences

Prepared for: The Executive Committee:
School of Clinical Medicine
Faculty of Health Sciences
University of the Free State

By: Hanneke Brits

May 2020

"If we are to be faithful to the charge placed upon us by society to certify the adequacy of clinical performance... then we can no longer evade the responsibility for finding a method that will allow us to do so."

- Miller 1990

TABLE OF CONTENTS

1. Executive summary iv

2. Introduction 1

3. Problem statement 2

4. Recommendations 4

5. Action plan 11

6. Conclusion 13

7. Final remark 13

8. References 14

EXECUTIVE SUMMARY

It is said that competence is measured by what you do and not what you say you do. The University of the Free State (UFS) prides itself on producing competent medical doctors who can integrate knowledge, skills and attitudes that are relevant for the South African context. It is the responsibility of the School of Clinical Medicine to assess effectively whether graduates have, actually, acquired the required knowledge, skills and attitudes, and to defend the results of assessments that certify competence. Benchmarking current assessment practices for undergraduate medical students in the clinical phase against an assessment framework for undergraduate medical training revealed a range of quality in assessment practices.

The aim of this proposal is to make scientifically sound, yet practically feasible recommendations to address the gaps identified in current assessment practices in the clinical phase of the undergraduate medical programme at the UFS, and to ensure that existing good practices are maintained.

Seven activities for intervention are recommended, namely,

- Ensure alignment between learning outcomes, training (teaching and learning activities) and assessment;
- Improve the content and construct validity of all assessments;
- Improve the reliability of assessments;
- Plan and conduct all assessments with feasibility (e.g., resources) in mind;
- Ensure that all assessments have an educational effect (e.g., through feedback);
- Adapting assessment methodology, so that it moves towards workplace-based assessment (WBA) and integrated assessment, and has less emphasis on summative assessment; and
- Maintaining and improving the quality of assessment through existing structures.

For each activity, specific recommendations will be made, each with an action plan for implementation. Most of these recommendations may be implemented at department level; it will require some training and have minimal financial implications. The implementation of workplace-based assessment (WBA) to certify competence, and a single final integrated assessment, will require the approval of the Executive Committee of the School of Clinical Medicine and adherence to other relevant academic approval processes at the UFS. Implementation of the proposed interventions should improve the validity and reliability of the assessment of undergraduate medical students in the clinical phase, and ensure the defensibility of results.

INTRODUCTION

To produce competent medical doctors who can integrate knowledge, skills and attitudes relevant to the South African context, as prescribed in the Health Professions Act (HPA), the School of Clinical Medicine should train for and assess this competence.[1] The assessment of competence is complex, due to various factors, including the continuous emergence of new best practice medical evidence,[2] the theory-practice gap, between what is taught and what is observed in clinical practice[3-5] and what is feasible,[6] and the challenges related to assessment in real-life situations that may compromise the reliability of the assessment.[7] Assessment of competence must also satisfy various stakeholders, among which patients and the general population, training facilities, regulatory bodies and students. Training facilities, therefore, have the responsibility to perform this assessment task effectively and to defend the results of high-stakes assessments.[8]

This proposal is motivated by the perception that not all assessments in the clinical phase of the undergraduate medical programme at the University of the Free State (UFS) are of high quality. This perception was supported by a study conducted on assessment practices as experienced by undergraduate medical students in the clinical phase of the MBChB programme at this UFS in 2019. This study found that only 43.5% of fourth-year and 44.1% of fifth-year medical students perceived the current assessment as fair.[9] The students indicated, furthermore, their dissatisfaction with some assessment practices and the lack of feedback, and also made valuable recommendations regarding the assessment of professionalism and integrated assessment.[9] Youens and Hall[10] state that, 'instead of treating students as voices crying in the wilderness, we would be far better served if we asked the voices' owners what they think and listened actively to the answers'. It is, therefore, important to value the opinions and perceptions of students regarding the quality of assessment.

In turn, some lecturers in the clinical phase of the MBChB programme at the UFS were of the opinion that certain students were fortunate or "lucky" to have passed the final assessment, and that they may not be competent to qualify as doctors.[10] To ensure a comprehensive overview of current assessment practices, a quantitative study was performed on student marks, and the reliability of marks obtained in a variety of assessments was determined.[11]

To identify the theory-practice gaps between current assessment practices and what is prescribed in the literature, a publication, "A framework to benchmark the quality of clinical assessment in a South African undergraduate medical programme"[12] was used to benchmark current assessment practices in the clinical phase of the undergraduate medical programme at the UFS. When benchmarked against this assessment framework, assessment practices were categorised as either excellent (> 80%), good (60-79%), borderline (45-59%) or poor (< 45%).

Gaps in the current assessment practices and suggestions aimed at bridging these gaps were then discussed in a focus group discussion with teaching and learning (T&L) coordinators of clinical departments of the School of Clinical Medicine.[13] The recommendations from the focus group discussion are included in this proposal, to improve the quality of competency assessment in the clinical phase of the undergraduate medical programme of the UFS.

PROBLEM STATEMENT

Benchmarking of competency assessment practices in the clinical phase of the undergraduate medical programme at the UFS revealed excellent, good, borderline and poor assessment practices. The practices identified as borderline and poor may pose problems for the defensibility of assessment, and may place accreditation of the undergraduate medical programme at risk. Figure 1 displays the benchmarking results.

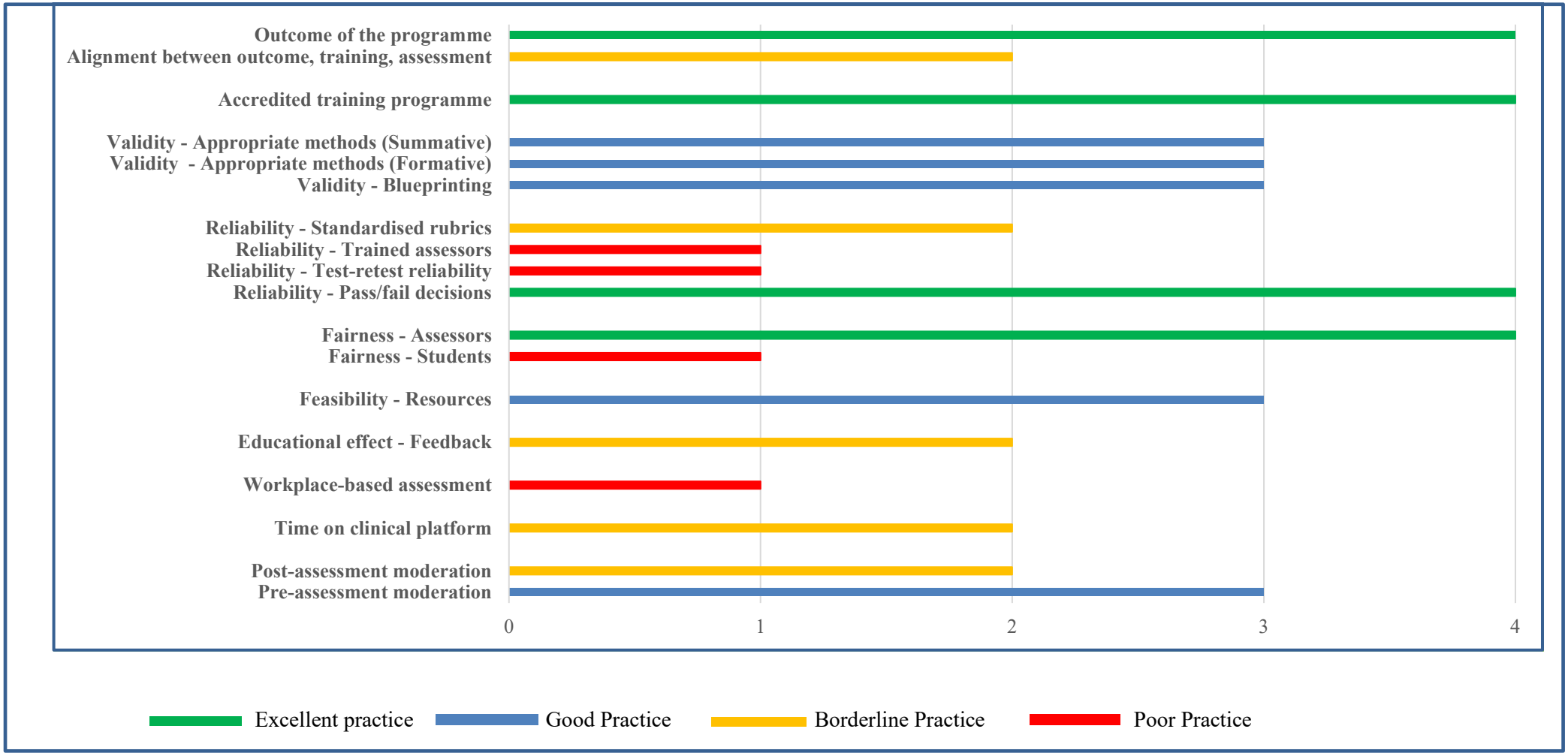


Figure 1: Benchmarking results of current assessment practices compared with an assessment framework

RECOMMENDATIONS

To bridge the gaps between excellent/good and borderline/poor practices, and to improve quality of competency assessment in the clinical phase of the undergraduate medical programme at the UFS, the following recommendations are made:

- Ensure alignment between learning outcomes, training (T&L activities) and assessment;
- Improve the content and construct validity of all assessments;
- Improve the reliability of assessments;
- Plan and conduct all assessments with feasibility in mind;
- Ensure that all assessments have an educational effect;
- Adapt assessment methodology, with a move towards WBA and integrated assessment, with less emphasis on summative assessment; and
- Maintain and improve the quality of assessment through existing structures.

<u>Ensure alignment between learning outcomes, training (T&L activities) and assessment</u>

The outcome of the undergraduate medical training programme at the UFS is producing competent medical doctors who can integrate knowledge, skills and attitudes relevant to the South African context.[14] This outcome is in line with recommendations in the Health Professions Act of South Africa[1] and the South African Qualifications Authority (SAQA).[15] The integration of knowledge, skills and attitudes ultimately takes place on the clinical training platform and in real-life situations, and not in classrooms.

Recommendation 1: Departments to adjust rosters and rotations to ensure maximum clinical exposure during training. Doing so will ensure that students are exposed to common conditions (burden of disease) and are trained in situations where knowledge, skills and attitudes are integrated. In this way, constructive alignment between outcomes, training and assessment will also be addressed.

Recommendation 2: Introduce continuous assessments in clinical areas while students perform duties and receive training. This recommendation can be implemented by using WBA. The WBA should focus on the core competencies prescribed by the Health Professions Council of South Africa, namely, Healthcare Professional as: "Professional, Collaborator, Communicator, Health Advocate, Leader and Manager and Scholar".[16] The introduction of a competency portfolio for recording achieved clinical competencies during WBA will provide evidence of entrustable professional competences that do not need to be assessed again.[17]

Recommendation 3: Introduce a "professionalism portfolio" that is to be used throughout undergraduate training. Attitudes and professionalism should be assessed under usual working and training conditions, and not only during formal assessments. It should be the student's responsibility to get at least two conduct assessments per rotation. All unprofessional conduct should also be captured in the portfolio and reported to the management of the undergraduate programme.

<u>Improve the content and construct validity of all assessments</u>

Content validity of assessments can be improved by blueprinting of individual as well as overall assessments.[18]

Recommendation 4: All assessments must be blueprinted. An assessment matrix or blueprint must be implemented for the overall module assessment, as well as for individual assessments. This blueprint should be included in rotation planning.

The construct validity of assessment can be improved by applying appropriate assessment methods.[18] Assessment planning should ensure sufficient and systematic sampling of questions and assessment methods.

Recommendation 5: Ensure that all questions are unambiguous and clear. All questions should be clear, to avoid any confusion on what is expected in the assessment. Ensuring moderation of all assessments will assist in this regard.

Recommendation 6: Ensure appropriate assessment methods to assess specific competencies. The performance of a skill must be demonstrated, as it cannot be tested in a theoretical question. To ensure that knowledge, skills and attitudes are assessed, a variety of assessment methods should be used.

Improve the reliability of assessments

The reliability of a clinical assessment is defined as the degree to which a test measures the same concept in different assessments and obtains stable or reproducible results.[19,20] Assessment results of the undergraduate medical programme at the UFS show that pass/fail decisions are almost 100% accurate, but test-retest reliability, as well as stability and reproducibility of individual student marks in the clinical phase, are not good. Only 3/26 (11.5%) of the assessments had a correlation coefficient of ≥ 0.7, which is an indication of good reliability.[11] The reliability of an assessment is influenced by the assessment, the assessor and the student.[21]

Recommendation 7: Increase the number of assessments and use a variety of assessment methods. In the case of reliability, more is better, and a greater number of assessments with different assessors will contribute to better reliability.[22] The reliability of assessments will improve by combining multiple-choice questions that are reliable, WBA that is authentic, and observed clinical cases that can assess integration.

Recommendation 8: Ensure that standardised assessment rubrics are used for all clinical assessments. There exist many standardised assessment and feedback tools that ensure effective assessment and feedback during formative assessment, e.g., mini clinical evaluation

exercises (mini CEX), clinical encounter cards (CEC), direct observation of procedural skills (DOPS).[23]

Although no perfect pass mark exists, the pass mark at the UFS is set at 50%.[24] Criterion-referenced measurement is regarded as the best method to assess competence.[25] Clinical competence cannot, however, be measured by a single mark (or percentage), as competence comprises multiple facets and the integration of different skills.[25] Assessing clinical competence is even more complex in clinical practice, when a single mistake may prove fatal, e.g., using an incorrect drug dosage.

To be certified as competent, a student needs to pass an assessment with at least 50%. Students, therefore, need to meet 100% of the minimum criteria for competence to score 50% when pass/fail decisions are at stake.

Recommendation 9: Assessment rubrics should be set to ensure a score of 50% if minimum competence criteria are met. Assessors should grasp the concept that minimum criteria are necessary for competence and they should be trained on using the assessment tools.

<u>Plan and conduct all assessments with feasibility (resources) in mind</u>

The feasibility of assessments is determined by the number of trained assessors, the availability of stable patients with good clinical signs, the timing and duration of the assessment and the financial implications of assessments.[26] Increased student numbers and the recommendation to include more assessments to increase reliability places strain on assessment feasibility.

Recommendation 10: Train registrars and clinical staff members to do regular WBA. Training will empower clinicians and ensure that, in future, there are enough competent assessors. WBA will ensure that there are enough patients for assessments, without the extra financial expenditure required to arrange additional final summative assessment opportunities that use clinical cases involving patients.

Ensure that all assessments have an educational effect (e.g., through feedback)

Assessment should promote learning by encouraging students to study for assessments and providing constructive feedback.[27] According to the assessment policy of the UFS, students must receive feedback after assessments.[28] However, Brits et al. found that only 15.4% of disciplines routinely gave feedback after assessments.[10]

Recommendation 11: Schedule formal, compulsory feedback sessions on student rosters two weeks after end-of-block assessments. Timely feedback will allow students and assessors the opportunity to identify areas that must receive more attention and clear up areas of confusion regarding learning content and assessment methods. The timing of feedback is important to ensure the best educational effect, so that good practices are reinforced and poor performance can be addressed and rectified.[29]

Recommendation 12: Give group feedback via electronic platforms e.g., Blackboard. This method of feedback should address problem areas during assessment, as well as class averages. Students can get an idea of their competence compared to the rest of the class, which may motivate further learning.

Recommendation 13: Give immediate feedback after patient cases and WBA on what went well and what aspects need improvement. In this way, learning will be reinforced. Feedback training should be provided to ensure that assessors are competent in this area and students get the maximum educational effect of feedback.

Recommendation 14: Implement compulsory meetings for poorly performing students with T&L coordinators. These meeting will enable the provision of targeted support. Attendance of these meetings will provide evidence of the student and the educational institution's efforts regarding obtaining and providing support respectively.

Assessment methodology should be adapted, with a move towards WBA and integrated assessment, and less emphasis on summative assessment

WBA is described as the observation of students while they are performing skills and competencies in the workplace. It includes immediate feedback, to improve, reinforce or certify a skill.[30,31] Less than a third (30.1%) of disciplines in the undergraduate medical programme use WBA.[10] SAQA states that "there shall be some component of the final assessment that assesses a student's ability to adopt an integrated approach to clinical knowledge and reasoning", and that students should have more than one opportunity to display their knowledge.[32]

Recommendation 15: All disciplines should implement some form of WBA, using standardised assessment tools. See Recommendation 8 for a description of these tools. The educational effect of assessment will improve with constructive feedback during low-stakes assessments, like WBA, with the added benefit of improving reliability, validity and feasibility of assessment.

Recommendation 16: Implement a final integrated medical assessment at the end of the final year to replace module-specific assessments. All disciplines should be involved in this assessment. Theoretical knowledge can be tested with multiple-choice questions, and clinical skills with written and clinical objective structured clinical evaluations (OSCEs). The Family Medicine OSCE in the final year can serve as an example of such an integrated assessment.

Assessment drives learning,[33] and reward is a good motivator for learning.[34] By combining assessment with reward, students' motivation to learn may increase. Promotion is practised at various universities; this involves students who perform well during formative assessments being exempted from summative assessments at the end of the year.[24,35]

Recommendation 17: Implement a promotion system for students who perform well in assessments during rotations (formative/summative). Students who pass block assessments and obtain >60% for both the theoretical and clinical components could be exempted from summative assessments in these modules at the end of the year. Only borderline students (who obtained <60% in block assessments during the year) would then need to do summative assessments in these modules.

Maintaining and improving the quality of assessment through existing structures

To improve the quality of undergraduate medical training, the School of Clinical Medicine has appointed T&L coordinators for all major clinical disciplines. These T&L coordinators are UFS staff members who ensure the smooth running of the undergraduate medical programmes in their departments and modules, thereby supporting overburdened clinical staff on the joint staff establishment with administrative teaching responsibilities. A number of administrative staff members have also been appointed to support the clinical departments with student training in the undergraduate medical programme.

Recommendation 18: Ensure that all T&L coordinator and administrative posts are filled, and promote the professional development of T&L coordinators. Proactive measures must be in place to ensure timeous appointments in vacant posts.

Moderation of all assessments is part of the quality assurance process prescribed by the UFS.[28] Although these policies exist, 75% of disciplines in the clinical phase of the undergraduate medical programme at the UFS do pre-assessment moderation, but only half do post-assessment moderation.[10] Formal moderation reports form part of the required documentation during programme accreditation.

Recommendation 19: Ensure the implementation of the existing moderation policies and procedures for all assessments. The T&L coordinators, in consultation with the academic heads of departments, should take responsibility for moderation. T&L coordinators should

collaborate on best practices for implementing moderation in their respective disciplines. All the necessary documents to assist with this process are available to simplify the process.

ACTION PLAN

Implementing these recommendations will require a number of processes on different managerial levels. The recommendations that require approval from higher authority should be assessed for feasibility and due processes should be followed to ensure implementation. The effective implementation of these recommendations will require the availability of resources and realistic timelines. In Figure 2, the level of authority, resources and timelines for implementing recommendations are displayed.

Most of these recommendations are operational and can be implemented immediately by departments. Training of T&L coordinators and assessors should be a priority. The Division Health Sciences Education at the Faculty of Health Sciences, UFS, has the necessary skills and expertise to assist in this regard.

Action plan - Quality of assessment recommendations						
	Pre-requisites	**Administrative responsibility**	**Resources**	**Implementation level**	**Responsible driver of process**	**Full implementation date**
1: Departments to adjust rosters and rotations to ensure maximum clinical exposure during training.	None	Departmental administration	Training of T&L	Departmental/ Operational	T&L coordinator	Jan-21
2: Introduce a "Competency portfolio" for continuous assessments in clinical areas while students perform duties and receive training	UFS Assessment policy in place	Departmental administration	Training of T&L + clinical personnel	Departmental/ Operational	HOD + T&L coordinator	Jan-21
3: Introduce a "Professionalism portfolio" to be used throughout undergraduate training.	Approval by EXCO	SCM administration	Training of T&L + clinical personnel	Programme / Regulations	EXCO SCM	Jan-21
4: All assessments must be blueprinted	UFS Assessment policy in place	None	Training of T&L	Departmental/ Operational	T&L coordinator	Immediately
5: Ensure that all questions are unambigious and clear	UFS Assessment policy in place	None	Training of T&L	Departmental/ Operational	T&L coordinator	Immediately
6: Ensure appropriate assessment methods to assess specific competencies	UFS Assessment policy in place	None	Training of T&L + clinical personnel	Departmental/ Operational	T&L coordinator	Jan-21
7: Increase the number of assessments and use a variety of assessment methods.	UFS Assessment policy in place	Departmental administration	Training of T&L + clinical personnel	Departmental/ Operational	T&L coordinator	Jan-21
8: Ensure that standardized assessment rubrics are used for all clinical assessments.	UFS Assessment policy in place	Departmental administration	Training of T&L + clinical personnel	Departmental/ Operational	T&L coordinator	Immediately
9: Assessment rubrics should be set to ensure a score of 50% if minimum competence criteria are met	UFS Assessment policy in place	None	Training of T&L + clinical personnel	Departmental/ Operational	HOD + T&L coordinator	Jan-21
10: Train registrars and clinical staff members to take responsibility for regular workplace-based assessments (WBA).	UFS Assessment policy in place	None	Training of T&L + clinical personnel	Departmental/ Operational	HOD + T&L coordinator	Jan-21
11: Schedule formal compulsory feedback sessions on student rosters two weeks after end of block assessments.	None	Departmental administration	Training of T&L	Departmental/ Operational	T&L coordinator	Jan-21
12: Give group feedback via electronic platforms e.g. Blackboard.	None	Departmental administration	Training of T&L	Departmental/ Operational	T&L coordinator	Immediately
13: Implement immediate feedback sessions after patient cases and WBA's	UFS Assessment policy in place	None	Training of T&L	Departmental/ Operational	HOD + T&L coordinator	Immediately
14: Implement compulsory meetings for poor performing students to meet with T&L coordinators.	None	Departmental administration	Training of T&L	Departmental/ Operational	T&L coordinator	Immediately
15: All disciplines should implement some form of WBA using standardized assessment tools e.g. Competency portfolio	UFS Assessment policy in place	Departmental administration	Training of T&L + clinical personnel	Departmental/ Operational	HOD + T&L coordinator	Jan-21
16: Implement a final integrated medical assessment (IMA) at the end of final year instead of module specific assessments.	Approval by EXCO	Departmental administration	Training of T&L + clinical personnel	Programme / Regulations	EXCO SCM	Nov-21
17: Implement a promotion system for well performing students	Approval by EXCO	Departmental administration	Training of T&L + clinical personnel	Programme / Regulations	EXCO SCM	Nov-21
18: Ensure that all T&L posts are filled and T&L coordinators are empowered to take responsibiity for assessments	Approval by EXCO	Departmental administration	None	Programme / Regulations	EXCO SCM	Immediately
19: Ensure the implementation of the existing moderation policies and procedures for all assessments.	UFS Assessment policy in place	Departmental administration	Training of T&L	Departmental/ Operational	HOD + T&L coordinator	Immediately
*** T&L = Teaching and Learning coordinators SCM - School of Clinical Medicine EXCO - Executive Committee HOD - Head of Department**						

CONCLUSION

The benchmarking exercise to evaluate assessment practices in the clinical phase of the MBChB programme at the UFS gave a good overview of the current status quo. Most of the practices that compromise the quality of assessment can be addressed on an operational level and will not be costly to address. This includes training of assessors, increasing the duration of clinical training and WBA, providing students with effective feedback and blueprinting and moderation of all assessments. Assessor training will improve the quality of assessments, and will contribute to the professional development of assessors. T&L coordinators will play an important role in ensuring the implementation of the recommendations to improve the quality of assessment in the clinical phase of the undergraduate medical programme. Although a move away from summative assessment in modules, towards continuous WBA, will require regulatory changes, the ultimate effect of improved validity and reliability will benefit all stakeholders.

FINAL REMARK

The Health Sciences Education division possesses the expertise to perform the relevant training for the T&L coordinators and clinicians as proposed in this document. Coordination of the suggested changes is possible through the Phase III working group.

REFERENCES

1. Health Professions Act 56 of 1974. Regulations relating to the registration of students, undergraduate curricula and professional examinations in medicine. Government Notice R139 in Government Gazette 31886 of 19 February 2009. https://www.hpcsa.co.za/Uploads/MDB/Rules%20%26%20Regulations/regulations_gnr_139_2009.pdf. Accessed 20 January 2020.

2. Cooke SJ, Johansson S, Andersson K, Livoreil B, Post G, Richards R, Stewart R, Pullin AS. Better evidence, better decisions, better environment: emergent themes from the first environmental evidence conference. Environmental Evidence. 2017;6(1):15. https://doi.org/10.1186/s13750-017-0092-0

3. Hussein MH, Osuji J. Bridging the theory-practice dichotomy in nursing: The role of nurse educators. J Nurs Educ Pract. 2017;7(3):20-5. https://doi.org/10.5430/jnep.v7n3p20

4. Ajani K, Moez S. Gap between knowledge and practice in nursing. Procedia - Social and Behavioral Sciences. 2011;15:3927-31. doi:10.1080/14703297.2013.860879

5. Salah AA, Aljerjawy M, Salama A. Gap between theory and practice in the nursing education: The role of clinical setting. Emergency. 2018;24:17-8. http://dx.doi.org/10.19080/JOJNHC.2018.07.555707

6. Zlatkin-Troitschanskaia O, Pant HA. Measurement advances and challenges in competency assessment in higher education. Journal of Educational Measurement. 2016;53(3):253-64.

7. Clauser BE, Margolis MJ, Swanson DB. Issues of validity and reliability for assessments in medical education. In ES Holmboe, SJ Durning, RE Hawkins (eds). Practical guide to the evaluation of clinical competence. 2nd ed. 2018. Philadelphia: Elsevier.

8. Hays RB, Hamlin G, Crane L. Twelve tips for increasing the defensibility of assessment decisions. Med Teach. 2015;37(5):433-436. https://doi.org/10.3109/0142159X.2014.943711

9. Brits H, Bezuidenhout J, Van der Merwe LJ, Joubert G. Students' voices: Assessment in undergraduate clinical medicine. Pan African Medical Journal. 2020:36;130. doi:10.11604/pamj.2020.36.130.221

10. Brits H, Bezuidenhout J, Van der Merwe LJ, Joubert G. Assessment practices in undergraduate clinical medicine training. What do we do and how we can improve? Afr J Prm Health Care Fam Med. 2020;12(1), a2341. https://doi.org/10.4102/ phcfm.v12i1.2341

11. Brits H, Joubert G, Bezuidenhout J, Van der Merwe LJ. Evaluation of assessment marks in the clinical years of an undergraduate medical training programme: Where are we and how can we improve? Submitted to African Journal of Health Professions Education.

12. Brits H, Bezuidenhout J, Van der Merwe LJ. A framework to benchmark the quality of clinical assessment in a South African undergraduate medical programme. S Afr Fam Pract. 2020;62(1), a5030. https://doi.org/10.4102/safp. v62i1.5030

13. Brits H, Bezuidenhout J, Van der Merwe LJ. Quality assessment in undergraduate medical training: how to bridge the gap between what we do and what we should do. Pan African Medical Journal. 2020;36:79. doi:10.11604/pamj.2020.36.79.23658

14. University of the Free State. School of Clinical Medicine vision and mission statement [homepage]. [cited 2018 Apr 13]. Available from: https:// www.ufs.ac.za/health/departments-and-divisions/school-of-medicine-home.

15. South African Qualifications Authority. Criteria and guidelines for assessment of NQR registered unit standards and qualifications [Internet]. SAQA; 2001 [cited 2019 Oct 12]. Available from: http://cdn.lgseta.co.za/resources/guidelines/2.4.1%20SAQA%20 Criteria%20and%20Guidelines%20for%20Assessment.pdf

16. Health Professions Council of South Africa. Core competencies for undergraduate students in clinical associate, dentistry and medical teaching and learning programmes in South Africa [homepage on the Internet]. 2014 [cited 2018 Mar 12]. Available from: http://www.hpcsa.co.za/uploads/editor/UserFiles/downloads/medical_dental/MDB% 20Core%20Competencies%20-%20ENGLISH%20-%20FINAL%202014.pdf.

17. Chen HC, Van der Broek WES, Ten Cate O. The case for use of entrustable professional activities in undergraduate medical education. Academic Medicine. 2015;90(4):431-436. doi: 10.1097/ACM.0000000000000586

18. Tavakol M, Dennick R. The foundations of measurement and assessment in medical education. Medical Teach. 2017 Oct 3;39(10):1010-5. https://doi.org/10.1080/0142159X.2017.1359521

19. Manterola C, Grande L, Otzen T, García N, Salazar P, Quiroz G. Reliability, precision or reproducibility of the measurements. Methods of assessment, utility and applications in clinical practice. Rev Chilena Infectol. 2018;35(6):680-688. https://dx.doi.org/10.4067/S0716-10182018000600680

20. Pietersen J, Maree K. Standardisation of a questionnaire. In K Maree (ed.). First steps in research. 7th Impression. 2016. Pretoria: Van Schaik Publishers.

21. Maughan S, Tisi J, Whitehouse G, Burdett N. A review of literature on marking reliability research [Internet]. Slough: National Foundation for Educational Research; 2013 Jun [cited 2019 Sep 12]. Available from: https://www.nfer.ac.uk/publications/mark01/mark01.pdf

22. Tisi J, Whitehouse G, Maughan S, Burdett N. A review of literature on marking reliability research. Report for Ofqual. Slough: NFER. 2013. https://www.nfer.ac.uk/publications/mark01/mark01.pdf. (accessed 12 September 2019).

23. Norcini J, Burch V. Workplace-based assessment as an educational tool: AMEE Guide No. 31. Med Teach. 2007;29(9):855–71. https://doi.org/10.1080/01421590701775453

24. UFS 2020. General rules for undergraduate qualifications, postgraduate diplomas, Bachelor Honours degrees, Master's degrees, doctoral degrees, higher doctorates, honorary degrees and the convocation. https://www.ufs.ac.za/docs/default-source/policy-institutional-documents/2020-general-rules-for-publication-28-november-2019---with-web-links-to-legislation-and-govt-gazettes.pdf?sfvrsn=dafe9d21_2

25. Yudkowski R, Park Y, Downing S. Introduction to assessment in health professions education. Written tests: writing high-quality constructed-response and selected-response. In R Yudkowski, Y Park, S Downing (eds). Assessment in health professions education. 2nd ed. New York: Routledge; 2019.

26. Norcini J, Friedman Ben-David M. Concepts in assessment. In: A practical guide for medical teachers. 4th ed. London: Elsevier; 2013. pp. 285–91.

27. Govaerts M. Workplace-based assessment and assessment for learning: Threats to validity. Journal of Graduate Medical Education 2015;7(2):265–7. https://doi.org/10.4300/JGME-D-15-00101.1

28. University of the Free State. Assessment policy on the UFS coursework learning programme [homepage on the Internet]. 2016 [cited 2019 May 11]. Available from: https://www.ufs.ac.za/docs/default-source/all-documents/assessment-policy-on-the-ufs-coursework-learningprogrammes.pdf?sfvrsn=3716c321_0.

29. McSweeney F. Students' views on assessment. Other resources. 2014;12. https://arrow.dit.ie/aaschssloth/1217

30. Lefroy J. Action research: towards excellence in teaching, assessment and feedback for clinical consultation skills [Internet] [PhD]. Keele University; 2018 [cited 2019 Oct 12]. Available from: http://eprints.keele.ac.uk/5170/1/LefroyPhD2018.pdf#page=146

31. McBride M, Adler M, McGaghie W. Workplace-based assessment. In R Yudkowski, Y Park, S Downing (eds), Assessment in health professions education. 2nd ed. New York: Routledge; 2019.

32. South African Qualifications Authority. The National Qualifications Framework curriculum development [homepage on the Internet]. 2000 [cited 2019 May 15]. Available from: http://www.saqa.org.za/docs/pol/2000/curriculum_dev.pdf.

33. Wormald BW, Schoeman S, Somasunderam A, Penn M. Assessment drives learning: an unavoidable truth? Anat Sci Educ. 2009 Oct;2(5):199-204.) DOI:10.1002/ase.102

34. Marvin CB, Shohamy D. Curiosity and reward: Valence predicts choice and information prediction errors enhance learning. J Exp Psycho Gen. 2016 Mar;145(3):266. http://dx.doi.org/10.1037/xge0000140

35. University of Pretoria. 2020. Faculty of Health Sciences Year book. https://www.up.ac.za/yearbooks/2020/MED-faculty/rules/Faculty%20Rules

CHAPTER 8:
CONCLUSIONS, IMPLICATIONS AND RECOMMENDATIONS

8.1 INTRODUCTION

Chapter 1 introduced the reader to the importance of quality assessment in an undergraduate medical programme (cf. 1.2.5). This information was supplemented by an outline of the undergraduate medical programme and assessment practices applied at the UFS (cf. 1.2.1). The chapters that followed were in the form of articles that addressed an assessment framework (cf. Chapter 2), current assessment practices from different views (cf. Chapters 3–5) and triangulation of the results during a focus group interview (cf. Chapter 6). Chapter 7 was a proposal to the Executive Committee of the School of Clinical Medicine on best practice recommendations to enhance quality assessment in the clinical phase of the MBChB programme at the UFS (cf. Chapter 7).

The research concludes with Chapter 8. This chapter aims to summarise the research and reflect on the research process and results. This chapter includes a summary of the study and study results, a personal reflection and some recommendations on the way forward.

8.2 OVERVIEW OF THE STUDY

The theoretical grounding of best practices in assessment occured throughout this thesis. The first chapter was an overview of the components of (cf.1.2.3) and requirements for quality assessments (cf. 1.2.5) The research design, with pragmatism as the point of departure (cf. 1.8), was used to investigate the theory behind current assessment practices and to make recommendations to enhance best practices in competency assessment.

In Article 1 (cf. Chapter 2) a rapid review was performed on all policies and regulations pertaining to assessment practices in undergraduate medical training. Twenty-five policy documents and 24 other documents were used to compile a theoretical framework for quality assessment in the undergraduate medical programme.

With the study methodology and outcomes of each article (cf. Chapters 3,4,5 and 6) in

mind, literature searches were performed as background to the specific aspect of quality assessment addressed in each paper.

Universities and training facilities should be able to defend their assessment results when they certify that they produce competent medical practitioners who can integrate knowledge, skills and attitudes relevant to the South African context. This assessment and certification of clinical competence of entry-level medical practitioners are complex and multi-dimensional. The defined problem that was addressed by this study is that assessment in the clinical phase of the MBChB programme at the UFS has not been reviewed critically or benchmarked against local and international standards. This thesis intended to benchmark clinical assessment practices against an assessment framework and then propose an action plan on how to bridge the gap between theory and practice when assessing clinical competence.

Pragmatism was the paradigm that was used to address the practical problem of uncertainty regarding the quality of assessment. From a theoretical perspective, an abductive approach was used to achieve inference. A mixed method methodology was implemented by using both quantitative and qualitative methods to collect data. During triangulation, alignment of and gaps between theory and practice were identified and solutions recommended. Finally, a proposal with an action plan was drafted to enhance the quality of clinical assessment in the MBChB programme.

An in-depth study of assessment in the clinical phase of undergraduate medical training revealed a paucity of assessment standards. Without standards, it is difficult to benchmark current assessment practices and to make recommendations for improvement.

The overall goal of the study was to contribute to high quality assessment practices in the clinical phase of the MBChB programme by recommending assessment practices of which the outcome can be confidently defended and justified. Three research questions were formulated and answered in this study.

8.2.1 Research question 1: What are the current policies and guidelines to benchmark quality assessment in the clinical phase of the MBChB programme?

Before this study, no content and context-specific assessment framework existed to benchmark clinical assessment in MBChB programmes. In Part 1 of this study, a rapid literature review of local, national and international official regulations and policies, supported by best evidence practices, were used to compile an assessment framework to benchmark current assessment practices for undergraduate clinical medicine. In this framework, the three components of quality assessment, namely, accreditation, assessment and quality assurance were addressed (cf. 1.2.3). Guidelines for assessment content and standards, assessment types, assessment methods and principles of assessment were included in the framework. The importance of constructive alignment between outcomes, training and assessment were emphasised in the framework. The framework also addressed the balance between formative, summative and continuous assessment and the important place of workplace-based assessment (WBA) to ensure clinical competence. Guidelines were provided on how to improve the validity, reliability, fairness, feasibility, educational effect and acceptability of clinical assessments.

It is important to realise that clinical assessment is complex and that no assessment is perfect. An assessment framework can help to improve assessment, but it cannot guarantee quality assessment. To answer the first research question, an article with the title, *A framework to benchmark the quality of clinical assessment in a South African undergraduate medical programme,* was published (cf. Chapter 2).

8.2.2 Research question 2: What are the current assessment practices employed in the MBChB programme at the UFS and how valid and reliable are these assessments?

In Part 2 of this study, assessment practices in the clinical phase of the MBChB programme were reviewed. Information to answer this question was gathered from three sources, namely, students, lecturers and student marks, to ensure that different views were included.

Firstly, a questionnaire with open and closed-ended questions was completed by clinical students in the MBChB programme, to get the students' perspectives on assessment. Less than half the students were of the opinion that the assessments were fair, with lack of blueprinting and an incorrect level of assessment major contributors to this opinion. Two thirds believed that the assessment was aligned with outcomes, however, training was not aligned with the assessment. Almost all the students reported a lack of feedback after assessment. If they passed an assessment, fewer than half of fourth-year students, compared to three quarters of fifth-year students, were confident about their knowledge and skills. Valuable suggestions from the students included ways of assessing professionalism, timing of assessments and feedback and training of assessors. An article with the title, *Students' voices: Assessment in undergraduate clinical medicine,* addressed the students' perspective on assessment (cf. Chapter 3).

Secondly, the T&L coordinators and module leaders of all the clinical departments involved in undergraduate medical training completed questionnaires on the assessment methods used in their departments and made recommendations for ways to improve current assessment practices. The departments used various assessment methods to assess the competencies required of entry-level health care professionals. WBA was performed by a third of disciplines, while MCQs and OSCEs were the main methods used during assessment. The study found that not all assessors were well prepared for assessment, with more than a third never having received any formal training in assessment. Few disciplines made use of post-assessment moderation as a standard practice or gave feedback after assessments. Current assessment practices for clinical students in the undergraduate medical programme at the UFS cover the spectrum that is necessary to assess all the different competencies. The implementation of more formative assessments, with less emphasis on summative assessment, was recommended. It was also suggested that WBA and continuous assessment may be better ways to ensure the effective assessment of clinical competence. An article, *Assessment practices in undergraduate clinical medicine training. What do we do and how we can improve?* addressed aspects of validity and reliability in the assessment practices used at the UFS (cf. Chapter 4).

Thirdly, current assessment practices were evaluated for reliability. A total of 1 380 marks in 26 assessments were evaluated. This evaluation revealed that the reliability of pass/fail outcome decisions were good. It may, therefore, not be necessary to do a summative assessment if a student passed the end-of-block assessment, as these two assessments

produced the same pass/fail results. The test reliability, as well as stability and reproducibility of individual student marks, were not good. MCQ tests with fewer than 30 questions, short questions, and unobserved long cases were not reliable and it is recommended that they should not be used in summative assessments. To improve the reliability of assessments, more low-stakes assessments should be performed. MCQs, OSCEs and WBAs would be the most effective way of ensuring competence with reliability in mind. An article, *Evaluation of assessment marks in the clinical years of an undergraduate medical training programme: Where are we and how can we improve?* addressed the reliability of the assessment practices in the clinical phase of the MBChB programme (cf. Chapter 5).

8.2.3 Research question 3: What best practices can be implemented to enhance quality assessment in the clinical phase of the MBChB programme?

In Part 3 of the study, the T&L coordinators used the results of Part 2 of the study to benchmark current assessment practices against the assessment framework compiled in Part 1. This was done during a focus group interview with an experienced facilitator. The benchmarking of current assessment practices in the clinical phase of the undergraduate medical programme at the UFS revealed excellent, good, borderline and poor assessment practices. These poor practices included a lack of alignment between outcomes, assessment and training, as well as a lack of standardised assessment rubrics, untrained assessors, lack of feedback and absence of post-assessment moderation. Solutions were proposed on how to improve these practices, while good practices were also shared and celebrated. Discussion of best assessment practices, including the roles of validity, reliability, feasibility, fairness and the educational effect of assessment, contributed to recommendations on how to address the gap between theory and current assessment practices. The article reporting on the focus group interview, *Quality assessment in undergraduate medical training: How to bridge the gap between what we do and what we should do,* was published in a peer-reviewed journal (cf. Chapter 6).

Finally, the researcher compiled a proposal with action plan on how to enhance quality assessment in the clinical phase of the MBChB programme at the UFS. To bridge the gaps between excellent/good and borderline/poor practices, and to improve quality assessment, the following areas for improvement were identified:

- Ensure alignment between learning outcomes, training (teaching and learning activities) and assessment;
- Improve the content and construct validity of all assessments;
- Improve the reliability of assessments;
- Plan and conduct all assessments with feasibility (e.g., resources) in mind;
- Ensure that all assessments have an educational effect (e.g., through feedback);
- Assessment methodology should be adapted with a move towards WBA and integrated assessment, with less emphasis on summative assessment; and
- Maintain and improve the quality of assessment through existing structures.

An action plan with 19 specific recommendations under these areas was proposed. Most of the practices that compromise the quality of assessment can be addressed on an operational level, and will not be costly. This includes training assessors, WBA, effective feedback to students and blueprinting and moderating all assessments. Assessor training will improve the quality of assessments, and will also contribute to the professional development of assessors. T&L coordinators will play an important role in ensuring the implementation of the recommendations to improve the quality of assessment in the clinical phase of the undergraduate medical programme. The move away from an emphasis on summative assessment, towards continuous assessment, is already in the assessment policy of the UFS (UFS 2016b), and will only require rule changes. Continuous WBA will have the ultimate effect of improving validity and reliability, which will benefit all stakeholders. A proposal, *Best practice recommendations to enhance quality assessment in the clinical phase of the MBChB programme,* was submitted to the Executive Committee of the School of Clinical Medicine for consideration and possible implementation (cf. Chapter 7). A brief overview of the study is displayed in Figure 8.1.

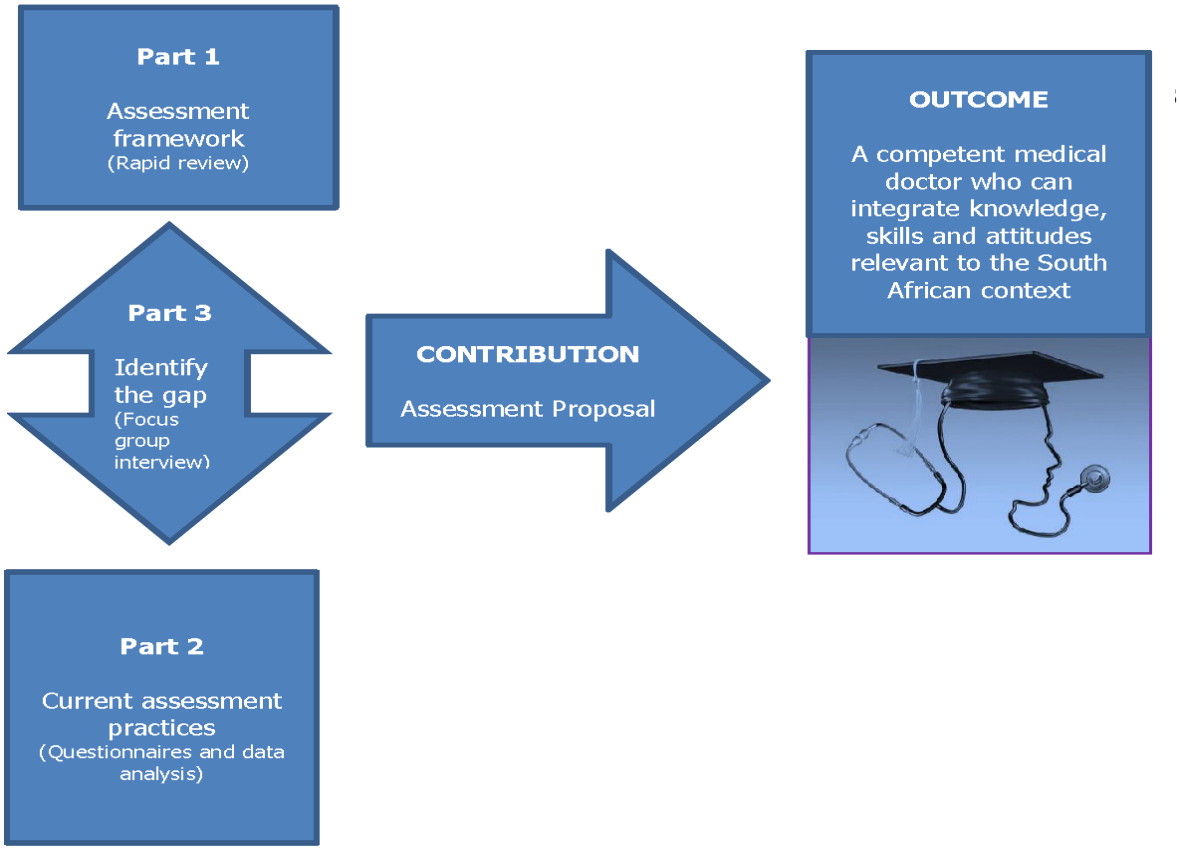


Figure 8.2: A brief overview of the study (Compiled by researcher Brits 2020)

8.3 IMPLICATIONS OF THE STUDY

8.3.1 Goal and contribution

This thesis was approached to contribute to quality assessment practices in the clinical phase of the MBChB programme. To achieve this goal, the researcher addressed the research questions regarding assessment policies, practices and recommendations in line with the title of this thesis: Best practices for quality assessment in the clinical phase of undergraduate medical training. This first contribution of this thesis was an assessment framework for benchmarking quality clinical assessment in an undergraduate medical programme, which was published in a peer-reviewed journal (cf. Chapter 2). Referring to local, national and international official regulations and policies, supported by best evidence practices to compile this assessment framework makes the framework transferable to other universities, who can use it to benchmark their MBChB programmes (cf. Chapter 2). A

second contribution by this study is a proposal and action plan on how to implement best practices for quality assessment in the MBChB programme at the UFS. (cf. Chapter 7). The recommendations on best practices for quality assessment were derived from theory and practice, which make it implementable in different clinical settings.

8.3.2 Final conclusion

To certify clinical competence (which is the outcome of undergraduate medical training programmes) numerous and different assessment methods are necessary (cf. 1.2.4, Chapters 2, 5, 6 and 7). Although the pass/fail decision of current assessments are good and good assessment practices exist, areas for improvement were also identified (cf. Chapters 4 and 5). The final conclusion from this thesis is that the assessment of clinical competence should be done through multiple assessments, with emphasis on WBAs. A move away from high-stakes summative assessment, to continuous assessment in the workplace that meets the requirements of validity, reliability, fairness, feasibility, educational effect and acceptability, is recommended (cf. Chapters 2, 4, 5, 6 and 7).

8.3.3 Study limitations and critique

Although the researcher and supervisors put in effort to ensure the trustworthiness of the study, the following limitations are recognised:

- Some literature was excluded, due to the search criteria employed in this study. It is possible that information was overlooked, due to a paucity of literature on undergraduate clinical assessment, which limited the use of forward and backward searching for more literature. Some sources of information may have been missed if the keywords did not fit the search criteria.
- The researcher may have failed to gather all possible data during the data collection processes. Despite the response rate of >80% for students and 100% for the clinical disciplines in the questionnaires, the responses may not be 100% accurate, as participants were asked to give their opinions, perceptions and recommendations. Only T&L coordinators were included in the focus group interview, and only five of the six participated.
- The requirements of different peer-reviewed journals, for instance, place limitations on number of words, number of tables and figures and references, which forced the researcher to focus the results. Some information was, therefore, not published or

reported on in this thesis. However, the researcher was able to include all information deemed important to meet the objectives of this study.

- The researcher was not completely familiar with all research methods employed, e.g., a rapid review for data gathering and the focus group interview. Outside information was obtained and the guidelines were followed meticulously to ensure sound research methodologies.
- The assessment practices of only one university were evaluated. Therefore, the recommendations to improve assessment practices may not apply to other universities.

8.3.4 Recommendations

Quality assurance is a process that is applied to ensure high quality. Although it is important to improve quality, it is equally important to maintain quality. The researcher recognises the following areas that may need further investigation to ensure sustainable quality assessment in the clinical phase of MBChB programmes:

- The first recommendation is to implement a quality improvement cycle for assessment practices. Through a systematic process it may be possible to evaluate the impact of the recommendations made in Chapter 7 of this study and improve the quality of clinical assessment further.
- With the outcome of the undergraduate medical programme well described and agreed upon, and assessment guidelines in place, the second recommendation of this study is to investigate alignment of teaching and training with outcomes and assessment.
- Although not arising from this study, but in the light of the fourth industrial revolution and the worldwide Covid-19 outbreak, the use of technology and innovative measures to enhance quality clinical assessment and training should be investigated.

8.3.5 Trustworthiness of the research

The researcher invested a great deal of effort to ensure the trustworthiness of this thesis. Credibility was improved through method and source triangulation (cf. 1.10.1), while the inclusion of local, national and international official regulations and policies to compile the assessment framework make the framework transferable to other universities to benchmark their MBChB programmes (cf. 1.10.2, Chapter 2). The recommendations on best practices

for quality assessment derived from theory and practice, which make the recommendations implementable in different clinical settings (cf. Chapter 7). Accurate description of all the processes followed by the different methods makes the study dependable or repeatable (cf. 1.10.3). To ensure confirmability, the data interpretation was discussed and clarified with the promotors and a biostatistician (cf. 1.10.4). The validity and reliability of the data collection were improved by using piloted questionnaires, double-checking results, and by triangulation (cf. 1.10.5 and 1.10.6).

8.3.6 Personal experience

Minor administrative problems were experienced with the approval of this study, which were solved by adjusting time frames and exercising patience. The data management and interpretation process provided the researcher with the opportunity to improve her skills under the expert guidance of a biostatistician.

The researcher opted to submit this thesis via the article route, as she had some experience of clinical article writing. On a positive note, the completion and submission for peer review of each article felt like an accomplishment. Each accepted paper was celebrated. These small steps kept the momentum of the project going. Individual articles had to meet the requirements of different journals, which allowed the researcher to gain experience of the different submission requirements and platforms. It was challenging to ensure continuity in the thesis, while at the same time, ensuring that each article was also a stand-alone publication. The researcher definitely recommends that students consider taking the article route for PhD studies.

There is a saying that, during PhD studies, you learn more and more about less and less, and eventually know everything about nothing. I wish to differ, as I hope that I now know a little more about quality clinical assessment and different research methodologies and methods.

8.4 FINAL REMARK

I was able to compile assessment criteria, critically review current assessment practices from different viewpoints, and compile best practice recommendation to enhance the quality of assessment in the clinical phase of the MBChB programme.

REFERENCES

Bartlett, J.W. & Frost C. 2008. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. *Ultrasound Obstetrics & Gynecology,* 31:466–475.

Bates, D.W. & Slight, S.P. 2014. Medication errors: What is their impact? *Mayo Clinic Proceedings,* 89(8):1027–1029.

Becker, M., Schütt, B. & Amini, S. 2014. *Proposal writing for international research projects: A guide for teachers*. Berlin: axeptDESIGN. https://portal.uni-koeln.de/sites/international/aaa/92/92pdf/92pdf_PROGRANT_Proposal_Writing_SCREEN.pdf

Biggs, J.B. 1996. Enhancing teaching through constructive alignment. *Higher Education*. 32:347-364. DOI: 10.1007/BF00138871

Boud, D. 1998. *Assessment and learning – unlearning bad habits of assessment.* Presentation at the conference, Effective Assessment at University, University of Queensland. http://www.tedi.uq.edu.au/conferences/A_conf/papers/Boud.html

Bowen, G. 2009. Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2):27–40. DOI: 10.3316/QRJ0902027

Brits, H., Bezuidenhout, J., van der Merwe, L.J. 2020. A framework to benchmark the quality of clinical assessment in a South African undergraduate medical programme. *South African Family Practice*, 62(1):a5030. DOI: 10.4102/safp.v62i1.5030

Brown, L., Douglas, V., Garrity, J. & Shepherd, C.K. 2012. What influences mentors to pass or fail students. *Nursing Management*, 19(5).

Chen, H.C., Van der Broek, W.E.S. & Ten Cate, O. 2015. The case for use of entrustable professional activities in undergraduate medical education. *Academic Medicine*, 90(4):431–436.

Creswell, J.W. 2014. *Research design qualitative, quantitative and mixed methods approaches* (4th edition). California, Thousand Oaks, CA: Sage Publications.

Creswell, J.W. & Plano Clark, V.L. 2011. *Designing and conducting mixed methods research* (2nd edition). Thousand Oaks, CA: Sage Publications.

Creswell, J.W. & Poth, C.N. 2018. *Qualitative inquiry and research design. Choosing among five approaches* (4th edition). Thousand Oaks, CA: Sage Publications.

Crotty, M., 1998. *Foundations of social research: Meaning and perspective in the research process*. London: Sage Publications. pp. 1–17.

Cruess, R.L., Cruess, S.R. & Steinert, Y. 2016. Amending Miller's pyramid to include professional identity formation. *Academic Medicine,* 91:180–185.

Danermark, B, Ekström, J., Jakobson, L. & Karlsson, J.C. 2001. *Explaining society. An introduction to critical realism in the social sciences*. London: Routledge.

Darling-Hammond, L., Herman, J., Pellegrino, J., Abedi, J., Aber, J.L., Baker, E., Bennett, R., Gordon, E., Haertel, E., Hakuta, K., Ho, A., Linn, R.L., Pearson, P.D., Popham, J., Resnick, L., Schoenfeld, A.H., Shavelson, R., Shepard, L.A., Shulman, L. & Steele, C.M. 2013. *Criteria for high-quality assessment*. California, Stanford Center for Opportunity Policy in Education.

Denscombe, M. 2014. *The good research guide: For small-scale social research projects* (5th edition). Berkshire: Open University Press.

Epstein, R.M. 2007. Assessment in medical education. *New England Journal of Medicine,* 326(4):387–396.

Fouché, C.B. & Delport, C.S.L. 2011. Introduction to the research process. In A.S. de Vos, H. Strydom, C.B. Fouché & C.S.L. Delport, (Eds). *Research at grass roots. For the social sciences and human service professions* (4th edition). Pretoria: Van Schaik Publishers. pp. 61–67.

Gainsbury, S. 2010. Mentors passing students despite doubts over their ability. *Nursing Times*, 106(16):2–3.

Guba, E.G. 1981. Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Communication and Technology Journal*, 29(2):75–91. https://link.springer.com/article/10.1007/BF02766777

Han, C. & Riazi, M. 2018. The accuracy of student self-assessments of English-Chinese bidirectional interpretation: a longitudinal quantitative study. *Assessment & Evaluation in Higher Education.* 3;43(3):386-98. DOI: 10.1080/02602938.2017.1353062

Heale, R. & Twycross, A. 2015. Validity and reliability of quantitative studies. *Evidence Based Nursing*, 18(3):66–67.

HPCSA (Health Professions Council of South Africa). 2014. *Core competencies for undergraduate students in clinical associate, dentistry and medical teaching and learning programmes in South Africa*. 2014. http://www.hpcsa.co.za/uploads/editor/UserFiles/downloads/medical dental/MDB%20Core%20Competencies%20-%20ENGLISH%20-%20FINAL%202014.pdf [Accessed 12 March 2018].

HPCSA (Health Professions Council of South Africa). 2019. *General Requirements of Approved Educational Institutions.* https://www.hpcsa.co.za/Uploads/MDB/Rules%20%26%20Regulations/regulations_gnr_139_2009.pdf [Accessed 17 April 2020]

Hughes, L.J., Mitchell, M. & Johnston, A.N. 2016. 'Failure to fail' in nursing – A catch phrase or a real issue? A systematic integrative literature review. *Nurse Education in Practice*, 20:54–63.

Ivankova, N.V., Creswell, J.W. & Plano Clark, V.L. 2011. Foundations and approaches to mixed methods research. In K. Maree (Ed.). *First steps in research*. 7th Impression. Pretoria. Van Schaik Publishers. pp 256–280.

Jiang, Y. 2001. Abductive reasoning as pragmatic inference: towards a formal theory for pragmatics. In H. Pan (Ed.). *Studies in Chinese linguistics.* Vol II. Hong Kong: SOAS.

Johnson, R.B. & Onwuegbuzie, A.J. 2004. Mixed methods research: a research paradigm whose time has come. *Educational Researcher,* 33(7):14–26.

Katzenellenbogen, J. & Joubert, G. 2014. Data collection and measurement. In R. Ehrlich & G. Joubert (Eds). *Epidemiology - A research manual for South Africa* (3rd edition). Cape Town: Oxford University Press. pp. 111–16.
Kelley, T.L. 1927. *Interpretation of educational measurements*. Oxford: World Book Co.

Kis, V. 2005. *Quality assurance in tertiary education: Current practices in OECD countries and a literature review on potential effects*. Organisation for Economic Co-operation and Development. https://www.oecd.org/education/skills-beyond-school/38006910.pdf [Accessed 11 September 2018].

Kovács, G & Spens, K.M. 2005. Abductive reasoning in logistics research. *International Journal of Physical Distribution & Logistics Management,* 35(2):132–144. DOI: 10.1108/ 09600030510590318

Lather, P. 1990. Reinscribing otherwise: The play of values in the practices of the human sciences. In E. Guba (Ed.). *The paradigm dialog: Options for social science inquiry*. London: Sage Publications. pp. 315–332.

Leinster, S. 2013. The undergraduate curriculum and clinical teaching in the early years. In J.A. Dent & R.M. Harden (Eds). *A practical guide for medical teachers* (3rd edition). London: Churchill Livingstone, Elsevier. pp. 16–22.

Li, D. 2004. Trustworthiness of think-aloud protocols in the study of translation processes. *International Journal of Applied Linguistics*, 14(3):301–313. DOI: 10.1111/j.1473-4192.2004.00067.x

Manghani, K. 2011. Quality assurance: Importance of systems and standard operating procedures. *Perspectives in Clinical Research*, 2(1):34–37. DOI: 10.4103/2229-3485.76288

Maxwell, G.S. 2010. Moderation of student work by teachers. In B. McGaw, P. Peterson & E. Baker (Ed.). *International encyclopedia of education*. 3. Elsevier. pp. 457–463.

Merriam-Webster. 2019. Dictionary. https://www.merriam-webster.com/dictionary/ [Accessed 3 September 2019].

Miller, G.E. 1990. The assessment of clinical skills/competence/performance. *Academic Medicine,* 65(Suppl 9):S63–S67.

Moynihan, S., Paakkari. L., Välimaa, R., Jourdan, D. & Mannix-McNamara, P. 2015. Teacher Competencies in Health Education: Results of a Delphi Study. *PLoS ONE,* 10(12): e0143703. https://doi.org/10.1371/journal.pone.0143703

Morgan, D.L. 2014. Pragmatism as a paradigm for social research. *Qualitative Inquiry,* 20(8):1045–1053.

Najimi, A., Sharifirad, G., Amini, M.M. & Meftagh, S.D. 2013. Academic failure and students' viewpoint: The influence of individual, internal and external organizational factors. *Journal of Education and Health Promotion*, 2:22. DOI: 10.4103/2277-9531.112698

Norcini, J., Anderson, B., Bollela, V., Burch, V., Costa, M.J., Duvivier, R., Galbraith, R., Hays, R., Kent, A., Perrott, V. & Roberts, T. 2011. Criteria for good assessment: Consensus statement and recommendations from the Ottawa 2010 Conference*. Medical Teacher*, 33(3):206–214. DOI: 10.3109/0142159X.2011.551559

Norcini, J. & Friedman Ben-David, M. 2013. Concepts in assessment. In J.A. Dent & R.M. Harden (Edis). *A practical guide for medical teachers* (4thedition)*.* London: Elsevier. pp. 285–291.

Onwuegbuzie, A.J. & Leech, N.L. 2007. Validity and qualitative research: An oxymoron? *Quality and Quantity*, 41:233–249. DOI: 10.1007/s11135-006-9000-3

Onwuegbuzie, A.J., Dickinson, W.B., Leech N.L. & Zoran, A.G. 2009. A qualitative framework for collecting and analyzing data in focus group research. *International Journal of Qualitative Methods*, 8(3):1–21.

Panadero, E., Brown, G.T. & Strijbos, J.W. 2016. The future of student self-assessment: A review of known unknowns and potential directions. Educational Psychology Review. 1;28(4):803-30. https://doi.org/10.1007/s10648-015-9350-2

Pangaro, L. & Ten Cate, O. 2013. Frameworks for learner assessment in Medicine: AMEE Guide No. 78. *Medical Teacher*, 35(6):e1197-e1210. DOI: 10.3109/0142159X.2013.788789.

Patil, S.Y., Gosavi, M., Bannur, H.B. & Ratnakar, A. 2015. Blueprinting in assessment: A tool to increase the validity of undergraduate written examinations in pathology. *International Journal of Applied and Basic Medical Research,* 5(S1):S76–S79.

Schwandt, T.A., Lincoln, Y.S. & Guba, E.G. 2007. Judging interpretations: But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. *New Directions for Evaluation*, 114:11–25. DOI: 10.1002/ev.223

Seely, J.F and Co-workers. 1996. *CanMEDS.* The Royal College of Physicians and Surgeons of Canada. http://www.royalcollege.ca/rcsite/canmeds/about/ contributors/canmeds-1996-contributors-e [Accessed 5 January 2018].

Shapton, M. 2006. Failing to fail students in the caring professions: Is the assessment process failing the professions? *Journal of Practice Teaching & Learning,* 7(2):39–54.

Spady, W.G. 1994. *Outcome-based education. Critical issues and answers*. Arlington: American Association of School Administrators.

Swaminath, G. & Raguram, R. 2010. Medical errors – I : The problem. *Indian Journal of Psychiatry*, 52(2):110–112. DOI: 10.4103/0019-5545.64580

Tashakkori, A., & Teddlie, C. 2010. *Handbook of mixed methods in social & behavioral research*. Thousand Oaks, CA: Sage Publications.

Ten Cate, O. 2013. Nuts and bolts of entrustable professional activities*. Journal of Graduate Medical Education*, 5(1):157–158. DOI: 10.4300/JGME-D-12-00380.1

Ten Cate, O., Chen, H.C., Hoff, R.G., Peters, H., Bok, H. & Van der Schaaf, M. 2015. Curriculum development for the workplace using entrustable professional activities (EPAs). AMEE Guide No. 99. *Medical Teacher,* 37(11):983–1002. DOI: 10.3109/ 0142159X.2015.1060308

Thatcher, R. 2010. Validity and reliability of quantitative electroencephalography. *Journal of Neurotherapy,* 14:122–152.

Twycross, A. & Shields, L. 2004. Validity and reliability - What's it all about? Part 2 Reliability in quantitative studies. *Paediatric Nursing,* 16(10):36.

Tobin, G.A. & Begley, C.M. 2004. Methodological rigour within a qualitative framework. *Journal of Advanced Nursing*, 48:388–396. DOI: 10.1111/j.1365-2648.2004.03207.x

Torres, M.B. & Cochran, A. 2016. Accuracy and content of medical student midclerkship self-evaluations, *The American Journal of Surgery,* 211(6):1153–1157. DOI: 10.1016/ j.amjsurg.2015.11.030.

UNESCO. 2015. *Principles of good practice in learning assessment.* The ACER Centre for Global Education Monitoring. http://uis.unesco.org/sites/default/files/documents/ principles-good-practice-learning-assessments-2017-en.pdf [Accessed 7 March 2018].

UFS (University of the Free State) (2016a). *Policy on Master's and Doctoral studies.* https://www.ufs.ac.za/docs/default-source/all-documents/masters-and-doctoral-studies.pdf?sfvrsn=ffd2c621_0. [Accessed 12 September 2019].

University of the Free State (UFS) (2016b). *Assessment policy on the UFS coursework learning programme.* https://www.ufs.ac.za/docs/default-source/all-documents/assessment-policy-on-the-ufs-courseworklearningprogrammes.pdf?sfvrsn=3716c321_0 [Accessed 1 May 2018].

University of the Free State (UFS) (2019). *Faculty of Health Sciences rule book. School of Medicine. Undergraduate qualifications 2019*. http://apps.ufs.ac.za/dl/yearbooks/315_yearbook_ eng.pdf [Accessed 1 May 2019].

Van der Bank, C.M. & Van der Bank, M. 2014. Quality assurance in higher education: A case study of the Vaal University of Technology. *Journal of Education and Social Research,* 4(1):395–406.

Van der Vleuten, C.P.M. 1996. The assessment of professional competence: Developments, research and practical implications. *Advances in Health Sciences Education*, 1:41–67.
Van der Vleuten, C.P.M. & Schuwirth, L.W.T. 2005. Assessing professional competence: from methods to programmes. *Medical Education*, 39(3):309–317.

Van der Vleuten, C., Sluijsmans, D. & Joosten-ten Brinke, D. 2017. Competence assessment as learner support in education. In M. Mulder (Ed.). *Competence-based vocational and professional education*. Springer International Publishing. pp. 607–630.

Wass, V. & Van der Vleuten, C. 2004. The long case. *Medical Education*, 38:1176–1180.

Watson, H.E. & Harris, B. 1999. *Supporting students in practice placements in Scotland*. Glasgow Caledonian University, Department of Nursing and Community Health.

Whiting, P., Wolff, R., Mallett, S., Simera, I & Savović, J. 2017. A proposed framework for developing quality assessment tools. *Syst Rev*. **6:**204. https://doi.org/10.1186/s13643-017-0604-6

APPENDICES

Appendix A: Ethics approval letter
Appendix B: Gatekeepers' approval
Appendix C: Language editor declaration
Appendix D: Turnitin Report

Appendix A: Ethics approval letter

UNIVERSITY OF THE FREE STATE
UNIVERSITEIT VAN DIE VRYSTAAT
YUNIVESITHI YA FREISTATA

UFS·UV
HEALTH SCIENCES
GESONDHEIDSWETENSKAPPE

Health Sciences Research Ethics Committee

12-Apr-2019

Dear **Prof Hanneke Brits**

Ethics Clearance: **BEST PRACTICES FOR QUALITY ASSESSMENT IN THE CLINICAL PHASE OF UNDERGRADUATE MEDICAL TRAINING**

Principal Investigator: **Prof Hanneke Brits**

Department: **Family Medicine Department (Bloemfontein Campus)**

APPLICATION APPROVED

Please ensure that you read the whole document

With reference to your application for ethical clearance with the Faculty of Health Sciences, I am pleased to inform you on behalf of the Health Sciences Research Ethics Committee that you have been granted ethical clearance for your project.

Your ethical clearance number, to be used in all correspondence is: **UFS-HSD2019/0001/2304**

The ethical clearance number is valid for research conducted for one year from issuance. Should you require more time to complete this research, please apply for an extension.

We request that any changes that may take place during the course of your research project be submitted to the HSREC for approval to ensure we are kept up to date with your progress and any ethical implications that may arise. This includes any serious adverse events and/or termination of the study.

A progress report should be submitted within one year of approval, and annually for long term studies. A final report should be submitted at the completion of the study.

The HSREC functions in compliance with, but not limited to, the following documents and guidelines: The SA National Health Act. No. 61 of 2003; Ethics in Health Research: Principles, Structures and Processes (2015); SA GCP(2006); Declaration of Helsinki; The Belmont Report, The US Office of Human Research Protections 45 CFR 461 (for non-exempt research with human participants conducted or supported by the US Department of Health and Human Services- (HHS), 21 CFR 50, 21 CFR 56; CIOMS; ICH-GCP-E6 Sections 1-4; The International Conference on Harmonization and Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH Tripartite), Guidelines of the SA Medicines Control Council as well as Laws and Regulations with regard to the Control of Medicines, Constitution of the HSREC of the Faculty of Health Sciences.

For any questions or concerns, please feel free to contact HSREC Administration: 051-4017794/5 or email EthicsFHS@ufs.ac.za.

Thank you for submitting this proposal for ethical clearance and we wish you every success with your research.

Yours Sincerely

Dr. SM Le Grange
Chair : Health Sciences Research Ethics Committee

Health Sciences Research Ethics Committee
Office of the Dean: Health Sciences
T: +27 (0)51 401 7795/7794 | E: ethicsfhs@ufs.ac.za
IRB 00006240; REC 230408-011; IORG0005187; FWA00012784
Block D, Dean's Division, Room D104 | P.O. Box/Posbus 339 (Internal Post Box G40) | Bloemfontein 9300 | South Africa

Appendix B: Gatekeepers' approval

UNIVERSITY OF THE FREE STATE
UNIVERSITEIT VAN DIE VRYSTAAT
YUNIVESITHI YA FREISTATA
UFS UV

Office of the Vice-Rector: Research and Internationalisation
Kantoor van die Viserektor: Navorsing en Internasionalisering

18-Apr-2019

Dear Prof Hanneke Brits

UFS AUTHORITIES APPROVAL

Research Project Title:
BEST PRACTICES FOR QUALITY ASSESSMENT IN THE CLINICAL PHASE OF UNDERGRADUATE MEDICAL TRAINING

This letter serves as confirmation that your request to collect data from students and/or staff members at the University of the Free State for your research project has been approved.

Kind Regards

PROF RC WITTHUHN
VICE-RECTOR: RESEARCH & INTERNATIONALISATION
CHAIR: SENATE RESEARCH ETHICS COMMITTEE

205 Nelson Mandela Drive/Rylaan
Park West/Parkwes
Bloemfontein 9301
South Africa/Suid-Afrika

P.O. Box / Posbus 339
Bloemfontein 9300
South Africa / Suid-Afrika
T:+27(0)51 401 2116
F:+27(0)51 401 3752
WitthuhnRC@ufs.ac.za
www.ufs.ac.za

Appendix C: Language editor declaration

Declaration

6 August 2020

Hester Sophia Human
PO Box 4
Otjiwarongo
Namibia

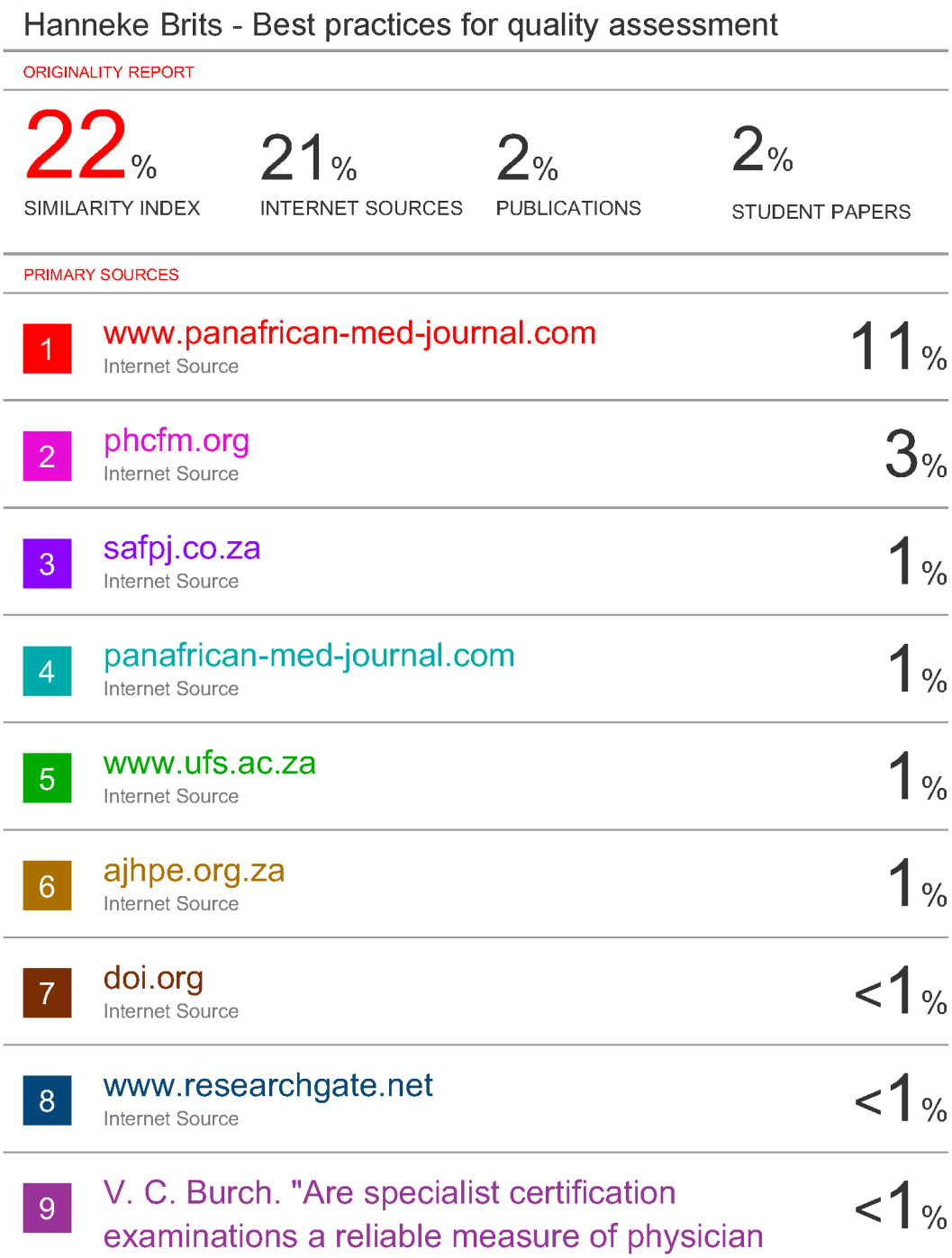
Student: Hanneke Brits

Thesis: Best practices for quality assessment in the clinical phase of undergraduate medical training

I confirm that I edited this this thesis and the articles involved. I checked the references and made recommendations for changes to the text.

Hettie Human
WRITER | EDITOR | TRANSLATOR | INTERPRETER
+264 813 359 120 | hettie.human@gmail.com

Appendix D: Turnitin report



Hanneke Brits - Best practices for quality assessment

ORIGINALITY REPORT

22%	21%	2%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.panafrican-med-journal.com Internet Source	11%
2	phcfm.org Internet Source	3%
3	safpj.co.za Internet Source	1%
4	panafrican-med-journal.com Internet Source	1%
5	www.ufs.ac.za Internet Source	1%
6	ajhpe.org.za Internet Source	1%
7	doi.org Internet Source	<1%
8	www.researchgate.net Internet Source	<1%
9	V. C. Burch. "Are specialist certification examinations a reliable measure of physician	<1%

	competence?", Advances in Health Sciences Education, 10/2008 Publication	
10	www.hpcsa.co.za Internet Source	<1%
11	www.samj.org.za Internet Source	<1%
12	van der Bank, CM, and Marjoné van der Bank. "Quality Assurance in Higher Education: A Case Study of the Vaal University of Technology", Journal of Educational and Social Research, 2014. Publication	<1%
13	sk.sagepub.com Internet Source	<1%
14	www.royalcollege.ca Internet Source	<1%
15	scholar.ufs.ac.za Internet Source	<1%
16	www.bmj.com Internet Source	<1%
17	Grace Anne Thompson. "A grounded theory of music therapists' approach to goal processes within their clinical practice", The Arts in Psychotherapy, 2020 Publication	<1%

18	John Norcini, Vanessa Burch. "Workplace-based assessment as an educational tool: AMEE Guide No. 31", Medical Teacher, 2009 Publication	<1%
19	www.ajhpe.org.za Internet Source	<1%
20	R. B. Hays, R. P. Strasser, T. K. Sen Gupta. "Twelve tips for establishing a new medical school", Medical Teacher, 2019 Publication	<1%
21	C.P.M. Van Der Vleuten, A.J.J.A. Scherpbie. "Clerkship assessment assessed", Medical Teacher, 2009 Publication	<1%
22	Olle ten Cate. "Medical Education, Competency-Based", Wiley, 2014 Publication	<1%
23	Anthonio Adefuye, Matthew Benedict, Johan Bezuidenhout, Jamiu O Busari. "Students' Perspectives of a Community-Based Medical Education Programme in a Rural District Hospital", Journal of Medical Education and Curricular Development, 2019 Publication	<1%
24	van der Vleuten, C.P.M.. "The assessment of professional competence: building blocks for	<1%

theory development", Best Practice & Research Clinical Obstetrics & Gynaecology, 201012
Publication

25	John A. Dent. "AMEE Guide No 26: clinical teaching in ambulatory care settings: making the most of learning opportunities with outpatients", Medical Teacher, 2009 Publication	<1%

Exclude quotes	Off	Exclude matches	< 15 words
Exclude bibliography	Off		