

# CHAPTER 10

---

## Summary and conclusions.<sup>12</sup>

This thesis addresses the topic of student assessment as it pertains to the training of medical practitioners in South Africa. Four key themes, relating to the purpose and utility of assessment, form the focus of attention of the literature reviewed and the work presented. They are: (1) the use of assessment to measure professional competence, (2) the use of assessment to facilitate student learning, (3) the use of assessment to initiate and sustain curriculum change, and (4) the selection of assessment tools on the basis of their utility or fitness for purpose.

The literature review in Chapter 1 provides an overview of some of the most important assessment practice advances, relevant to the four key themes, implemented over the past 30 years. The review notes that most of the cited assessment practice advances have been implemented in, and have impacted upon, training programmes in developed world countries. The paucity of published data from developing world countries, and the consequent limited understanding of assessment practices in these resource-constrained settings, is highlighted. Chapter 2 provides a contextual framework for the work presented in the thesis. This includes a basic description of undergraduate and postgraduate medical training programmes in South Africa, as well as insights into the socio-political and economic factors that currently impact upon medical education in this sub-Saharan African developing country. Six questions exploring specific aspects of the four key assessment themes are outlined in Chapter 3, and the relevant research data are presented and discussed in the papers presented in Chapters 4-9. In this chapter, I return to the research questions to summarise my findings. Thereafter, I discuss some recommendations based on the work presented, and suggest directions for further research.

---

<sup>12</sup> In this chapter, references made to the published literature are not cited again. They are all cited elsewhere in this thesis.

## Summary of research findings

### *1. Are portfolio interviews a reliable measure of professional competence?*

Owing to their professional authenticity, both in terms of task and location, portfolios have gained considerable popularity in health professional training programmes in the past decade. Current portfolio assessment strategies, however, requiring examination times of up to 170 minutes per candidate, are prohibitive in resource-constrained settings. Secondly, the psychometric rigour of portfolio assessment requires improvement. Thirdly, careful review of the portfolio document, the usual current assessment method, may not provide real insight into trainees' clinical ability but simply show that they are good at writing about what they do. More appropriate ways of determining the ability of trainees to deal with complex professional tasks, requiring integration of the relevant cognitive, psychomotor and affective skills, are needed. The use of interviews has not been suggested in the literature.

Chapter 4 describes an evaluation of the reliability and educational impact of a single-examiner, structured portfolio interview strategy. Fourth year medical students (n=181) completing a 14-week medical clerkship collated a written portfolio of 25 patient encounters. Trained examiners conducted 30-minute interviews in which four randomly selected patient encounters were discussed using six questions that probed the ability of candidates to interpret and synthesise clinical data gathered during patient encounters. Case scores were assigned using a criterion-referenced percentage scale. Portfolio interview scores, the average across four cases expressed on a percentage scale, contributed to a composite assessment score comprising an in-course rating, a multiple choice question (MCQ) test, and a bedside oral examination (BOE).

The internal consistency (reliability) of the portfolio interview scores was determined by calculating Cronbach's alpha coefficient and the standard error of measurement (SEM). Pearson's correlation coefficients for the various component subtests were calculated and the number of portfolios containing extra entries, i.e. more than the required 25, was recorded.

The mean ( $\pm$  SD, 95%CI) portfolio interview score achieved was 67.5% ( $\pm$  10.5, 66-69.1). Cronbach's alpha coefficient was 0.88 and the SEM was 3.6. The correlation coefficients for the portfolio interview, when compared to the MCQ test and the BOE, were  $r=0.42$  and  $r=0.37$ , respectively. Portfolios containing extra entries were submitted by 45.3% of students.

This study showed that a 30-minute, single-examiner interview, using standardised questions and a criterion-referenced rating scale, provided a reliable assessment of portfolios in a vocationally authentic manner. The method required considerably less examination time per candidate than published data, and fostered desirable student learning behaviour in the clinical work environment.

## ***2. Do specialist certification examinations reliably measure competence?***

Postgraduate specialist certification and licensure is the responsibility of a large number of medical colleges, professional boards and other associations throughout the world. Given the high stakes nature of these examinations, certification bodies have a social and professional responsibility to ensure that they are robust, fair and defensible. There is a paucity of published data about the reliability of specialist certification examinations and objective methods for improvement. Published psychometric evaluations of these high stakes composite examinations are needed in order to foster the improvement of current assessment practices and sustain the international credibility of specialist certification processes.

The Colleges of Medicine of South Africa recently initiated a review of their Fellowship examinations. Chapter 5 describes the psychometric evaluation of Part II of the Fellowship examination of the College of Physicians (FCP), a composite examination that includes two written short-answer question (SAQ) tests, a data interpretation (DI) test and three real patients encounters (PE). Three cycles of examination results were analysed using multivariate generalizability theory to estimate component and composite reliability. Disattenuated correlation coefficients for the component subtests were also derived, and decision studies (D-studies) were done to identify feasible ways of improving examination reliability.

Of 79 candidates attempting the written tests, 69 (87%) passed. Of these invited candidates, 54 also successfully completed the clinical component of the examination (78.3% success rate). The mean (SD) percentage score for the overall exam (weighting strategy: DI=0.2, SAQ=0.2, and PE=0.6) was 58.2 (8.1). Mean percentage scores for component subtests were: SAQ= 58.6 (5.5), DI=57 (8.8), and PE=58.8 (10.3). Component subtest reliability coefficient estimates, standardised for one hour of testing time, were: DI=0.37, SAQ=0.2 and PE=0.32. Composite examination reliability, using the weighting strategy described, was 0.72. This could be improved to 0.8 by weighting component subtests equally, or increasing the number of PE cases to five or more. The disattenuated correlations between the test components were: DI – SAQ=1.0, PE – DI=0.88, and PE – SAQ=0.81.

This composite postgraduate specialist certification examination reliability coefficient of 0.72 could be improved to 0.8, desirable for high stakes examinations, by weighting component subtests equally or by increasing the number of PEs in the examination. Since examiners were unlikely to reduce the weighting of the PE test, ways of increasing the number of PEs were explored. Given existing resource constraints it was suggested that in-course PE assessment strategies, such as the mini-CEX, may be a feasible way of increasing the number of patient encounters contributing to the final examination score.

### ***3. Does formative assessment promote learning in clinical clerkships?***

Clerkships, the backbone of clinical training, may not provide sufficient problem solving opportunities for students because they encounter patients after assessment by the attending staff. One strategy that may remedy this situation is the use of “blinded” patient encounters (BPE), i.e. clinical encounters where students interview and examine patients without access to their clinical records. Clerkship learning may also be enhanced by feedback that motivates students, informs them of their progress and assists in identifying learning needs.

At the University of Cape Town clerkship students do not routinely receive feedback based on observed performance and mostly engage in patient encounters after consulting patient records. This was addressed by implementing a formative assessment strategy providing structured feedback based on directly observed BPE. Chapter 6 describes faculty and student opinion regarding the value of feedback and the feasibility, credibility and educational value, including the impact on learning behaviour, of bedside formative assessment (BFA).

Fourth year students (n=155) completing a 14-week medical clerkship and 17 of 36 (42.7%) bedside teaching clinician-educators participated in the study. Students engaged in BPEs during bedside teaching sessions, and faculty provided structured (9-point scale) feedback on observed tasks: patient interview, physical examination, and formulation of an appropriate investigation and treatment plan. The number of BFA events undertaken was recorded, and questionnaires elicited student and faculty opinion regarding the value of feedback and the feasibility, credibility and educational value of BFA, including its impact on learning behaviour.

Thirteen faculty (87.1%) and 135 students (87.1%) were surveyed. Students recognised the learning value of BFA (95.6%) and feedback (70%) and reported improved clinical reasoning skills (88.2%), more case-driven reading (71.9%) and regular use of the BPE strategy (69.6%). Students having more BFA exposure (mean and 95% confidence interval), 5.7 (5.3-6.1) vs. 3.2 (2.8-3.6), more frequently reported perceived fairness of BPE assessment, 83.5% vs. 65.7% (p=0.008), knowledge of their own competence, 79.4% vs. 61.2% (p=0.02), and desirable learning behaviour, i.e. reading (79.4% vs. 64.2%, p=0.05) and regular BPE use (80.9% vs. 58.2%, p=0.004). Faculty agreed that integrating BFA into bedside teaching was feasible and educationally valuable. They endorsed the assessment validity of the BPE method.

This study demonstrated successful integration of structured feedback, based on observed performance, into a bedside teaching programme. This BFA strategy, endorsed by faculty and students, promoted desirable learning behaviour in the clinical clerkship setting. Specifically, students exposed to more BFA events felt better informed of their competence, more frequently considered the BPE assessment method fair, and, most importantly, reported changes in learning behaviour more often than their peers.

#### ***4. Are the procedural skill competencies of medical graduates adequate?***

The ability to perform a wide range of diagnostic and therapeutic procedures competently is a core learning outcome of modern undergraduate medical curricula. While the need to acquire these basic patient care skills is widely endorsed, educational guidelines usually only provide broad educational objectives that do not stipulate the specific practical skills medical students need to master prior to graduation. Some training institutions have developed detailed list of expected procedural skills competencies to remedy this situation.

Guidelines detailing the procedural skills competencies expected of South African medical students upon graduation do not currently exist. To determine the need for such guidelines, an OSCE assessment of the basic procedural skills competence of South African medical graduates, on commencement of their internship, was undertaken. Chapter 7 describes an evaluation of the procedural skills proficiency of 58 graduates on entry to their pre-registration year. Each subject participated in a 7-station objective structured clinical examination (OSCE); 6 of these assessed individual competence in phlebotomy, intramuscular injection, female pelvic examination, bladder catheterisation, tracheal intubation and prescription writing, while competence in cardiopulmonary resuscitation was assessed in a seventh station in randomly allocated teams of 3 candidates. Candidates' perceptions of their own competence were sought by questionnaire.

There was a wide variation in competence between subjects and across the range of tasks studied. Mean scores ranged from 85.4% for phlebotomy to 55.3% for prescription writing. The average score across all stations was 67.5%, and no student obtained an overall cut score of 85% or more, which was established using a modified Angoff method. Subjects' assessment of their own performance was unduly optimistic; most believed that they had demonstrated competence despite clear shortcomings in technique. Objective scores for subjects who had been exposed to structured skills laboratory learning activities were not significantly higher than for those who had not, although their self-assessed performance was higher.

The study found that most of the participating South African medical graduates were unable to satisfactorily perform technical procedures appropriate to the house officer on entry to the pre-registration year. This is in line with the conclusions of the few studies published in other countries. Based on these findings it was suggested that the learning outcomes of undergraduate medical programmes should include an explicit statement of the competencies required for practice in the pre-registration year, and that these should be adequately taught and rigorously assessed before graduation.

### ***5. Does PBL impact on the performance of academically-at-risk medical students?***

Since 1995, race-based distribution of educational resources has been abolished in South Africa. Significant differences remain, however, and the majority of black South Africans continue to be educated in suboptimal circumstances. It is thus not surprising that when black students, emerging from these educationally disadvantaged circumstances, enter university they are at risk of performing poorly despite their best efforts and innate talents. Attempts to address the educational needs of these students have included the introduction of extended medical programmes at several South African universities. Such a programme, the Academic Development Programme (ADP), was implemented at the University of Cape Town in 1991. Over the past decade the ADP has graduated more than 100 students.

Upon implementation of a new problem-based learning (PBL) programme in 2002, the Academic Development Programme was discontinued and all students were entered directly into the new PBL programme. Students who demonstrated a need for additional academic support by the end of the first semester entered an intervention programme for one year before proceeding to the second semester of the PBL programme. Chapter 8 describes an interim analysis comparing the retention rates and academic performance of academically-at-risk students in the new PBL programme and the ADP.

The records of all academically-at-risk students entering the ADP (1991-2000) and the new PBL programme (2002) were reviewed. Retention rates for all years of study, and academic performance (final course score expressed on a percentage scale) in the 4<sup>th</sup> year clerkship courses, including Internal Medicine, Obstetrics, Public Health, Primary Health Care and Psychiatry, of the respective programmes were compared.

A total of 239 academically-at-risk students in the ADP and 43 at-risk students in the new PBL programme were studied. The median retention rates, per year of study, for at-risk students in the problem-based learning programme was significantly better than for at-risk students in the ADP ( $p < 0.02$ ). Academic performance of the at-risk students in all the 4<sup>th</sup> year clinical clerkship courses of the PBL programme was significantly better than the mean performance over 10 years for at-risk students in the same 4<sup>th</sup> year courses in the ADP: Internal Medicine  $p < 0.02$ , Obstetrics  $p < 0.03$ , Primary Health Care  $p < 0.00001$ , Public Health  $p < 0.001$ , and Psychiatry  $p < 0.001$ .

The study showed that the introduction of PBL at the University of Cape Town has not had a deleterious effect on the performance of academically-at-risk medical students. Interim analysis suggests that retention rates and academic performance in the PBL programme are better than those achieved in the extended traditional programme.

## ***6. How can assessment utility parameters be used to select assessment methods?***

Selecting suitable assessment methods remains a daunting task for African medical schools where attempts to revamp existing assessment methods have been hampered by serious resource constraints. Student performance measures in these settings are largely determined by the resources required and the overall feasibility of the method selected. Furthermore, assessment methods used by most African medical institutions have been empirically selected. This situation, possibly the consequence of difficulties in objectively evaluating factors that influence test selection, highlights the need for a framework to guide the rational selection of assessment methods, particularly in African countries where medical schools are transforming their learning curricula and, therefore, their assessment methods. A lack of guidance on the selection of assessment methods may result in institutions adopting testing methods that cannot be sustained.

The paucity of assessment literature appropriate to the resource-constrained African context, further compounded by adverse socioeconomic, fiscal and political factors, underscores the need for an objective way of selecting resource-appropriate assessment tools. Chapter 9 describes a model for selecting assessment methods suitable for use in resource-limited environments. The approach, based on a model used to make rational drug selection(s) in the management of human disease, focused on four factors influencing test selection: performance, cost, suitability and safety.

The paper describes the systematic evaluation of four factors that influence the selection and implementation of assessment methods. Six commonly used methods – essay questions (EQ), short-answer questions (SAQ), multiple-choice questions (MCQ), patient-based clinical examinations (PCE), problem-based oral examinations (POE) and objective structured clinical examinations (OSCE) – were evaluated by calculating a score weighted for performance, cost, suitability, and safety factors. In the model, the highest score identified the most appropriate method.

Selection of an assessment method was illustrated using two institutional models depicting: (1) an ideal situation in which the OSCE was identified as the preferred method and (2) the typical African scenario in which EQ or SAQ methods were preferred.

The paper highlights the need to recognise important assessment utility determinants and provides an objective way of accounting for them when selecting assessment tools in resource-limited settings. A model suitable for guiding African institutions in the selection of assessment methods, appropriate to both educational purpose and context-specific resource constraints, is proposed.

In summary then, these six questions have explored specific aspects of the four key assessment themes that frame the work included in this thesis. Firstly, the use of assessment to measure clinical competence is explored by addressing two questions: (1) determining the reliability (internal consistency) of a structured interview process recently implemented as a summative portfolio assessment method in the MBChB programme at the University of Cape Town (UCT), and (2) determining the component and composite reliability of a multi-component high stakes Fellowship certification examination, of the Colleges of Medicine of South Africa, using multivariate generalizability theory, including identifying feasible ways of improving the examination reliability using prediction studies based on the available data. Secondly, the use of assessment to facilitate learning is explored by determining the perceptions of UCT faculty and medical students regarding the feasibility, credibility and educational value, including the impact on learning behaviour, of a bedside formative assessment strategy using observed “blinded” patient encounters akin to the clinical work sampling and mini-clinical evaluation exercise strategies described in the literature. Thirdly, the use of assessment to initiate and sustain curriculum change is explored by addressing two questions: (1) the use of trainee performance data to determine the need for curriculum change is studied by evaluating the procedural skill proficiency of South African medical graduates commencing their internship in order to provide a cogent argument for the development of a national list of procedural skills proficiency expected of all South African medical graduates upon commencing their internship; (2) the use of trainee performance data to sustain curriculum change by evaluating the impact of problem-based learning (PBL) on the retention rates and academic performance of academically-at-risk medical students in order to demonstrate the educational benefits of PBL, as implemented at UCT. Finally, factors that determine the utility of assessment practices is addressed by proposing the use of an assessment selection model that accounts for the important determinants of assessment utility prevalent in the resource-constrained settings in which medical training programmes in sub-Saharan Africa are located.

## **Recommendations**

Considering the major assessment practice advances highlighted in the literature review in Chapter 1, and the stark realities of the resource constraints that significantly impact upon medical education in South Africa, and sub-Saharan Africa in general, as outlined in Chapter 2, what would be a pragmatic way of addressing the emerging medical education crisis brewing on the African continent? There are a multitude of possibilities, but I would like to propose six initial simple strategies that do not require considerable additional financial or human resources, have been shown to be feasible in the work presented in this thesis, and could be implemented



in other South African, African or developing world contexts with minimal adaptation as appropriate to the specific context. The basic principle of each strategy is briefly outlined.

### ***1. Use portfolios to structure clerkship learning***

Student clerkships form the backbone of clinical training and, hence, need to be well structured to ensure that clinical learning is optimal. The literature highlights the limitations of clerkship learning and identifies the lack of educational structure of learning activities as a major problem resulting in learning of a highly variable quality. In resource-limited settings, where supervision is considerably less optimal than well-resourced settings in the developed world, the importance of providing educational structure in the clinical workplace is of paramount importance. The use of learning portfolios may make a significant contribution in this regard. One of the most important factors limiting the use of this innovative educational strategy in developing world, however, is the availability of a resource-efficient, in terms of time and personnel, assessment strategy.

The work in this thesis suggests that 30-minute interviews conducted by single examiners using a set of standardised questions and a criterion-referenced rating scale, may be a feasible option, particularly if the portfolio-based assessment forms part of a composite assessment package limiting the need for extreme psychometric rigour. In addition, this simple assessment tool encourages desirable learning behaviour, i.e. students undertook more self-directed patient encounters.

Based on the successful implementation of this assessment strategy in the 4<sup>th</sup> year Internal Medicine course in the MBChB programme at the University of Cape Town, it currently forms part of the assessment package of all courses in the final three clerkship years of undergraduate medical training programme at UCT. In addition, portfolio interviews, as described in the paper in this thesis, have recently been implemented in a clerkship course in the MBChB programme at the University of Pretoria. It seems, therefore, that this learning and assessment strategy can be readily implemented within the South African context. Broader application may require some modification.

### ***2. Provide trainee observation and feedback in the workplace environment***

The importance of providing feedback based on direct observation of trainee performance (formative assessment) is highlighted in the literature review. Despite the importance thereof, the difficulty of achieving and sustaining this fundamental educational need, even in well-resourced settings, is well illustrated in Chapter 1. In resource-limited environments, therefore, the need to develop a feasible, sustainable formative assessment strategy is critical to its utility. Furthermore, given the shift of student learning from academic hospitals to community-based health care sites, there is an additional requirement – such

strategies need to be simple in order to ensure optimal use by less-experienced clinicians now additionally tasked with student teaching, training and supervision in overcrowded, poorly resourced facilities.

The paper in Chapter 6 suggests that one of the most plausible ways of ensuring trainee observation and feedback is to make the assessment strategy an integral part of the clinical teaching programme, particularly a bedside teaching programme. Indeed, inextricably linking learning and teaching to assessment is the key to the successful implementation of the formative assessment strategy described. Not only did students get regular feedback based on directly observed performance, but it resulted in a reported increase in desirable learning behaviour, a highly sought after outcome of any assessment strategy.

Based on student demand, this bedside assessment strategy is now used in a number of other courses in the MBChB programme at UCT. This observation in itself provides an important strategy – students will adopt any strategy perceived to facilitate their learning if the activities being undertaken are concordant with the summative assessment strategies in place. Since the bedside formative assessment in use, i.e. “blinded” patient encounters, is identical to the bedside oral examination (four real patient encounters of 15 minutes each), a part of the composite summative assessment package, the incentive to sustain the implemented strategy is largely being driven by the students rather than the senior clinician-educators providing the feedback. Using student enthusiasm to drive and sustain desirable faculty educational practice is preferable since it is a renewable resource!

### ***3. Use multivariate generalizability theory to improve assessment practices***

The value of using multivariate generalizability theory to determine estimates of the reliability of current assessment strategies, and ways to improve them, is poorly recognised in the context of medical education. This is highlighted in the literature review and the paper presented in Chapter 5. This is truly a surprising finding, given the utility of this statistical method and the drive for evidence-based medical education. While it is recognised that access to suitable software has been limited to a few users, familiar with a DOS-based programme, recent advances in improved access to user-friendly software should facilitate the more widespread use of this elegant method of selecting options for improving educational practice on the basis of objective data.

The work in this thesis demonstrates the potential of this statistical method to improve the credibility of high stakes national specialist certification examinations in South Africa. This is of critical importance, even if the relevance thereof is only recognised to a limited extent at the current time. While public and trainee pressure to demonstrate the credibility of assessment practices is only beginning to emerge, the situation is changing, including in developing

countries like South Africa. The availability of research-based evidence to support educational practice, in particular assessment practices, is going to drive review of assessment practices within medical education, even if clinician-educators are currently only marginally interested in the issue. When this situation arises, the interest in multivariate generalizability theory is likely to increase exponentially.

#### ***4. Evaluate curriculum relevance using student performance data***

The use of trainee performance data to determine the relevance of curriculum contents to clinical practice demands is essential so as to ensure delivery of a vocationally relevant curriculum. While this is of international importance, it assumes critically important levels in developing countries where (1) the resources spent on medical education make up a major part of tertiary education grants to government-funded universities, (2) further postgraduate training is expensive, scarce, and service demands are high, making the likelihood of further training less probable than in developed countries, (3) major changes in disease patterns have significantly altered the health care needs of communities, e.g. the emergence of the HIV /AIDS epidemic, the advent of highly active anti-retroviral therapy and the new problem of multi-drug resistant tuberculosis. It is essential that training curricula keep up with changing community needs. One of the simplest ways of demonstrating a curriculum change need is to survey recent graduates regarding knowledge or skills currently considered essential in the initial years of clinical practice.

The work presented in Chapter 7 identifies the significant procedural skills gap in South African medical graduates, despite the existence of skills training centres in a number of medical schools in South Africa. Two further points are emphasised: (1) the presence of training facilities does not automatically translate into trainee skills proficiency if the skills are not formally assessed, a practice not routinely undertaken (procedural skills assessment) in South Africa, and (2) the need for a nationally agreed-upon list of basic therapeutic and diagnostic procedures medical graduates should be competent at performing at the time of entering their internship. There is a third need: an efficient mechanism for in-course monitoring to ensure that specified procedural skills are learnt under supervision in the clinical workplace.

The Health Professions Council of South Africa recently stipulated the use of procedural skills logbooks in all undergraduate training programmes in South Africa. Currently individual institutions are free to determine the range of skills stipulated in these training logbooks. While this does not directly address the need for a national list of procedural skills, such as exists in the Netherlands, it does set the stage for developing a national skills training framework in South Africa. It is anticipated that a process of developing a national procedural skills list will commence in 2007; research funding is currently being sought to undertake the

project. The paper in Chapter 7, the only published data in South Africa, makes a significant contribution to initiating this necessary curriculum change process.

Similar studies in other developing countries are likely to demonstrate similar findings. The provision of substantial evidence of a local training programme deficiency is, however, a powerful lever for curriculum change that needs to be used more frequently. This is especially true of situations where the adage “Why fix it if it isn’t broke?” is offered as an unsubstantiated anecdotal argument against initiating change.

### ***5. Evaluate curriculum change using student performance data***

The use of trainee performance data to determine the early impact of curriculum innovations is an important strategy for providing objective evidence for sustaining change in the face of ongoing resource limitations. This is particularly true of expensive curriculum innovations like problem-based learning. The strategy is also highly relevant to developing countries where medical education expertise is generally scarce and most curriculum design, content and delivery issues are dictated by tradition and habit rather than modern educational practice. In the context of PBL, student retention rates in developed world programmes were a brief focus of attention in the 1980s. Given that no significant differences were observed, however, this simple early measure of curriculum change success has not been further pursued in the literature.

The work presented in Chapter 8 suggests that academically-at-risk students may be a specific subgroup of students benefiting from PBL. A recent paper from the Walter Sisulu University, a rural historically Black medical school in South Africa training predominantly academically disadvantaged students, supports the findings described in Chapter 8. These two papers suggest that there may be a specific subgroup of students where curriculum interventions like PBL may be of particular benefit. Given the national priority of addressing persisting racial educational inequalities in South Africa, at-risk students currently make up a significant component of all medical schools enrolments. Tracking their progress may serve as a useful indicator of the impact of PBL in other schools where the method has been implemented, to a greater or lesser extent.

Given the poor schooling profile of medical entrants in many other African countries, simple evaluation of the progress (retention rates) of these at-risk students may serve as a useful early indicator of the impact of curriculum innovation, be it PBL or any other educational advance. Once again, the provision of objective data to substantiate the ongoing use of curriculum innovations, unfamiliar to clinician-educators with little or no medical education expertise, is a powerful lever for sustaining change. This is especially true in resource-limited settings where innovations requiring more resources are less likely to be sustained.

## ***6. Evaluate the utility of assessment practices***

Major factors determining assessment utility, including resource issues, are well described in the literature. Owing to the well-resourced circumstances of developed countries, relative to developing countries, other assessment utility determinants, specifically assessment reliability and validity, have always been the major focus of attention. The extreme pressure to select assessment tools appropriate to existing resources, above all else, has never really been a widely described issue in published literature emanating from the developed world. The situation in developing countries is, however, dramatically different. Resource availability dictates almost all educational practice, particularly assessment practices. The need to recognise this truth and deal with it in an objective way, such that it does not impact upon assessment selection at the expense of educational appropriateness, is a priority issue in many developing countries, particularly in sub-Saharan Africa.

Chapter 9 proposes a simple model for taking all these factors into account when selecting assessment tools in resource-limited settings. The utility of this tool has not been tested in the field. It would be useful to do so in South Africa, and further afield in Africa, in order to determine the educational utility of the proposed assessment selection tool and obtain objective data about “on the ground” assessment practices in developing countries.

### **Further research**

The work presented in this thesis presents a diverse number of research issues that could be further explored. I have chosen to highlight three issues that focus on different research themes: (1) international medical education collaboration – the extent of current collaboration and its origin, the impact of collaboration on current practice and ways of improving future collaboration such that medical education practice in developing countries is enhanced, (2) developing an understanding of factors in PBL programmes that facilitate the learning of educationally disadvantaged students in developing world medical schools, and (3) the national need for an agreed-upon list of essential procedural skills at the time graduate commence their internship in South Africa. These questions address three international issues relevant to the 21<sup>st</sup> century: global expansion of medical education expertise, further advances in our understanding of human learning, particularly in students from educationally disadvantaged backgrounds, and the need for technical procedural expertise when medical graduates commence clinical practice.

#### ***Collaborative medical education research***

Collaborative research between developing world medical educators and international medical education experts has the potential to enhance educational practice in the developing

world. Indeed, the clinical practice literature is increasingly demonstrating the benefits of international collaborative research initiatives impacting upon medical practice in developing countries. Good examples are particularly relevant to infectious diseases including HIV /AIDS, malaria and tuberculosis. The limited extent to which this is currently taking place in the field of medical education, at the expense of educational advance in developing world regions, is hinted at by the paucity of published data from developing world countries in top-class medical education journals. Strategies for advancing such collaboration need to be actively explored, particularly by large international medical education units. Although many of these units conduct higher degree programmes in medical education, the impact of these on the institutions in the developing countries from which some of these candidates originate is not known. Furthermore, the extent to which these degree programmes promote ongoing collaborative research between the developed world institution and the developing world institutions from which the candidates come is also not known. This is a topic of further research that may shed light on important unexplored opportunities to further advance medical education practice in regions of the world where it is most needed.

One of the grave dangers of such collaboration, however, is an increase in the developing world “brain drain” crisis. Strategies to enhance collaborative research, without loss of expertise to developing world source countries, have not been addressed to any great extent. A notable exception is the Foundation for Advancement of International Medical Education and Research, a non-profit organization aimed at improving medical education expertise in the developing world by residents of the developing world. Strategies to successfully increase medical education expertise within developing countries, and retain it in these world regions, is a novel topic that also deserves further consideration.

### ***Reasons why problem-based learning benefits academically-at-risk students***

This fascinating question has only recently emerged in the literature. To date, it has only been demonstrated in two South African universities. Walter Sisulu University has suggested that PBL students learn to use a wider variety of learning strategies and, therefore, become more versatile learners. Further work, in particular the use of structural equations modelling, may provide additional answers to this very important question.

### ***Basic procedural skills competence of South African medical graduates***

The need for a national list of defined procedural skills, required of graduates entering their internship, has already been extensively discussed. Developing such a list, involving both university stakeholders and service providers, is a large-scale research project that requires urgent attention. The most important challenge of this project would be to achieve participation by all the universities, given the political historical legacy of their origins.