



Australian Government
Office for Learning and Teaching

The **Australian Medical Assessment Collaboration**: developing the foundations for a national assessment of medical student learning outcomes

Final Report 2012

Lead Institution:

The University of Queensland

Partners:

Australian Council for Educational Research

Monash University

Project Leader:

Professor David Wilkinson (The University of Queensland)

Project Manager:

Dr Daniel Edwards (ACER)

Project Team members:

Associate Professor Hamish Coates (ACER)

Professor Ben Canny (Monash University)

Mr Jacob Pearce (ACER)

Dr Jennifer Schaefer (The University of Queensland)

Dr Tracey Papinczak (The University of Queensland)

Associate Professor Lindy McAllister (The University of Queensland)

Project website: <www.acer.edu.au/amac>



Support for the production of this report has been provided by the Australian Government Office for Learning and Teaching. The views expressed in this report do not necessarily reflect the views of the Australian Government Office for Learning and Teaching.



With the exception of the Commonwealth Coat of Arms, and where otherwise noted, all material presented in this document is provided under a Creative Commons Attribution 3.0 Australia licence (<http://creativecommons.org/licenses/by/3.0/au/>).

The details of the relevant licence conditions are available on the Creative Commons website (accessible using the links provided) as is the full legal code for the CC BY 3.0 AU licence (<http://creativecommons.org/licenses/by/3.0/au/legalcode>).

Requests and inquiries concerning these rights should be addressed to:
Office for Learning and Teaching
Department of Industry, Innovation, Science, Research and Tertiary Education

GPO Box 9880,
Location code N255EL10
Sydney NSW 2001

[<learningandteaching@deewr.gov.au>](mailto:learningandteaching@deewr.gov.au)

2012

ISBN 978 1 922125 33 0

Acknowledgements

The project team would like to thank all those in the sector involved in this project including participants in the each of our engagement forums, medical schools who contributed items to the project and staff in schools who facilitated participation in the pilots. The team would also like to thank staff at the ALTC/OLT for their support throughout the project, in particular Siobhan Lenihan. Acknowledgement and thanks is also given to the clinicians who took part in the item revision workshops at The University of Queensland and Monash University. Finally, thanks to the 500 or so final year medical students who donated their time to take part in the pilot testing and provide valuable feedback in relation to the assessment.

List of acronyms used

ACCLaiM	Australian Collaboration for Clinical Assessment in Medicine
ACER	Australian Council for Educational Research
AHELO	Assessment of Higher Education Learning Outcomes
ALTC	Australian Learning and Teaching Council Ltd
AMAC	Australian Medical Assessment Collaboration
AMC	Australian Medical Council
AMSAC	Australian Medical Schools Assessment Collaboration
IDEAL	International Database for Enhanced Assessments and Learning
LTAS	Learning and Teaching Academic Standards
MDANZ	Medical Deans of Australia and New Zealand
MCQ	Multiple Choice Question
OECD	Organisation for Economic Cooperation and Development
OLT	Office for Learning and Teaching

Table of Contents

Acknowledgements.....	3
List of acronyms used	3
Executive summary.....	5
Introduction	7
Project Outcomes and Impacts.....	8
Intended outcomes	8
Approach and methods	10
Outcomes and achievements	15
Dissemination	19
AMAC Assessment Framework	19
Engagement forums	20
Institutional involvement	20
Linkages.....	22
Evaluation	23
References	25
Appendix A: Project Reference Group.....	27
Appendix B: Institutions and organisations represented at AMAC engagement forums	28
Appendix C: Institutions involved in item provision and pilot testing.....	30
Appendix D: Evaluation Report.....	31

Executive summary

In late December 2010 the Australian Learning and Teaching Council Ltd (ALTC) provided a grant to The University of Queensland along with the Australian Council for Educational Research (ACER) and Monash University to develop foundations for a national assessment that evaluates the learning outcomes of later-year medical students in Australia.

The project, titled 'Developing the foundation for a national assessment of medical student learning outcomes' responds to the growing need to prove and improve the standards of medical education by establishing an Australian Medical Assessment Collaboration (AMAC). This project includes scoping work, wide-ranging sector engagement, development of an assessment framework, the compilation of assessment items, and the validation of items through pilot testing. The work provides a foundation for what will be the ongoing development and implementation of collaboration that will provide a sustainable and robust means of assuring the standards of medical education in Australia. A project website, www.acer.edu.au/amac provides further information about AMAC.

The ALTC grant for this project has enabled the establishment of AMAC. This report outlines the achievements of the project and provides operational detail on the establishment of the collaboration. Key outcomes of this project include:

- Development of an AMAC Assessment Framework;
- Engagement with all medical schools in Australia and New Zealand, and with key medical education stakeholders (nationally and internationally) through two engagement forums;
- Participation of nine medical schools in provision of draft items for the AMAC Assessment;
- Development of 120 final assessment items, verified and validated by clinicians;
- Participation of seven medical schools in pilot testing of the AMAC Assessment, involving more than 500 final year medical students;
- Exploration of pilot outcomes through psychometric analyses and examination of survey results from participating students and institutions;
- Development and distribution of institution-level and student-level benchmark reporting from the piloting of the AMAC Assessment;
- Enhancing and expanding discussion relation to common assessment in medical education, including helping to instigate the formation of a specific committee devoted to this in the medical dean's peak body; and
- The engagement with additional partner medical schools in the collaboration and planning for future sustainability of AMAC.

Through this initial project, a total of ten Australian medical schools are now formally involved in the collaboration and in the process of developing projects to build AMAC into a sustainable resource for medical schools into the future.

AMAC has substantially developed the discussion within the medical education community regarding common assessment across medical schools. While the idea of a common assessment of learning outcomes across institutions was an issue which seemed relatively

radical in the early stages of this project, there now appears to be almost uniform acceptance that there is a need for development of common assessments to enhance processes aimed at building continuous improvement. Key evidence of this is the formation during the period of this project of a Medical Deans of Australia and New Zealand (MDANZ) committee with a specific remit on developing policy relating to assessment and assisting in providing links for groups involved in these collaborations.

Introduction

In late December 2010 the Australian Learning and Teaching Council Ltd (ALTC) provided a grant to The University of Queensland along with the Australian Council for Educational Research (ACER) and Monash University to develop foundations for a national assessment that monitors the outcomes of later year medical students in Australia.

The project, titled 'Developing the foundation for a national assessment of medical student learning outcomes' responds to the need to prove and improve the standards of medical education by establishing an Australian Medical Assessment Collaboration (AMAC). This project includes scoping work, wide-ranging sector engagement, development of an assessment framework, the compilation of assessment items, and the validation of items through pilot testing. The work provides foundations for what will be the ongoing development and implementation of collaboration that will provide a sustainable and robust means of assuring the standards of medical education in Australia.

The ALTC grant for this project has enabled the establishment of AMAC. This report outlines the achievements of the project and provides operational detail on the establishment of the collaboration. Key outcomes of this project include:

- Development of an AMAC Assessment Framework;
- Engagement with all medical schools in Australia and New Zealand, and with key medical education stakeholders (nationally and internationally) through two engagement forums;
- Participation of nine medical schools in provision of draft items for the AMAC Assessment;
- Development of 120 final assessment items, verified and validated by clinicians;
- Participation of seven medical schools in pilot testing of the AMAC Assessment, involving more than 500 final year medical students;
- Exploration of pilot outcomes through psychometric analyses and examination of survey results from participating students and institutions;
- Development and distribution of institution-level and student-level benchmark reporting from the piloting of the AMAC Assessment;
- Enhancing and expanding discussion relation to common assessment in medical education, including helping to instigate the formation of a specific committee devoted to this in the medical dean's peak body; and
- The engagement with additional partner medical schools in the collaboration and planning for future sustainability of AMAC.

This report provides an overview of the processes employed in this project and the outcomes and impacts of the work. The following section outlines the intended outcomes for this AMAC project, details the approaches and methods used throughout and then provides an insight into the outcomes and achievements of the project. These detailed sections are followed by some brief exploration of the dissemination of project outcomes and materials, exploration of the linkages made through this work and an overview of the evaluation process. A full evaluation report by the project evaluator is included in Appendix B.

Project Outcomes and Impacts

Intended outcomes

The development of AMAC responds to the growing need for an evidence base on which to establish graduate capability, measure success and most importantly, facilitate continuous improvement in medical education in Australia.

The project was initiated to meet some important needs of the Australian medical education. Namely: collaboration in the area of common assessment; ability to benchmark outcomes across medical schools; and increasing capacity in development of quality assessments and assessment items – all with a particular emphasis on learning outcomes towards the end of the medical degree.

This project was specifically intended to have widespread impact on medical education in Australia. As detailed in the proposal for this project, core outcomes were intended to include:

1. The establishment of AMAC as a body to facilitate the ongoing implementation of a national assessment of learning outcomes in the medical field;
2. Medical education in Australia will have identified, in a collaborative manner, structures and approaches for assuring the quality of assessment;
3. Sector-wide involvement in an aspect of educational practice which is gaining serious momentum in the global landscape;
4. A model for the implementation of discipline-based objective testing of learning outcomes;
5. Framework and item development for the medical field to facilitate national benchmarking and evaluation of outcomes;
6. Initial psychometric evaluation of the results of the pilot testing of draft items; and
7. Development of an item bank, formed to provide the basis for the future development of a robust assessment instrument in the field.

The following section of the report outlines some of the context to the establishment of AMAC and the rationales behind developing a project with these intended outcomes. The Outcomes and Achievements section later in this report describes the extent to which the project team believes these intended outcomes have been achieved.

Context of the AMAC development

The need for the development of common assessment materials for medical education is fuelled by several pertinent contexts. The Australian Government has established the Tertiary Education Quality and Standards Agency (TEQSA) which has a broad remit to monitor educational outcomes. The Australian Learning and Teaching Council's (ALTC) Learning and Teaching Academic Standards (LTAS) project has developed a series of broadly discipline-focused threshold learning outcomes. Professional bodies in Australia and abroad are exploring and promoting core curricula and competencies that they expect graduates to achieve (for example, the competency-based occupational standards of the Speech

Pathology Association of Australia which formed the basis for the Carrick- and ALTC-funded COMPASS (competency assessment in speech pathology) (McAllister et al., 2006), and related benchmarking projects). Further catalysts are the *Review of Undergraduate Medical Education in Australia* (DEEWR, 2008), the Australian Medical Council's exploration of competency (2010), and a project on competencies in medical education conducted by the MDANZ (2011).

Growing internationalisation of the medical profession, increasing diversification of programs and curricula, and ever-growing pressure to improve academic standards heightens the need for robust assessment in medical education. In recent years there have been numerous calls for Australian medical students to be assessed against national competency standards in order to ensure that they have requisite core skills. In particular, recent discussion has been prompted by the Australian Medical Education Study that found some students graduate with "deficiencies in a number of clinical and procedural skills" as well as "inadequate knowledge of many of the basic medical science foundations" (DEEWR, 2008:14).

There is considerable growing national interest in and recognition of the policy importance of assessing learning outcomes through objective, system-wide testing. The OECD has seen this as important, currently facilitating an international study to examine the feasibility of initiating discipline-based assessments of learning outcomes across countries, known as the Assessment of Higher Education Learning Outcomes (AHELO) Feasibility Study. Australia is a key contributor to the study, with the Australian Government providing funding to facilitate the implementation of the project and ACER leading the multi-national consortium undertaking this important work (OECD, 2010).

Separately, but in response to many common drivers, interest in implementing international medical assessments is also gaining pace (Archer, 2009; Gorsira, 2009; Harden, 2009; Melnick, 2009; Van der Vleuten, 2009; Van der Vleuten et al., 2004).

Further adding to the contextual background in which AMAC has been developed are a range of complementary projects and stakeholders who share common ground with this project. The Australian Medical Schools Assessment Collaboration (AMSAC) involves a group of medical schools using shared assessment materials that are embedded in regular examinations held at the mid-point of the medical degree; the Australian Collaboration for Clinical Assessment in Medicine (ACCLAiM) project benchmarks clinical graduate outcomes across four medical courses; and the International Database for Enhanced Assessments and Learning (IDEAL) Consortium is an international collaboration of medical schools aimed at sharing of items for assessment in medical education. In addition, the Australian Medical Council (AMC) conducts standards-based program accreditations and develops tests for assessing the knowledge, clinical skills and professional attributes of overseas qualified medical practitioners seeking registration to practise medicine in Australia.

AMAC does not intend to replace or compete with any of these other projects in medical education. Instead it offers a complementary role, focusing on the final year of the medical degree and providing national benchmarks for learning outcomes while also promoting awareness and skills in the practice of item and assessment development. The development

and implementation of AMAC is intended to boost the capacity of medical schools to deliver high-quality education.

In this context it is also critical to note that AMAC does not intend to produce a national licensing exam, but rather a common assessment that medical schools can use to provide them with information about student learning outcomes and benchmark these across other medical schools.

Approach and methods

The approach of the AMAC project team throughout the project has been focused on collaboration and consultation. This section highlights the approaches to some key developments of the project. Throughout the discussion below, the theme of consultation and collaboration is prominent.

Development of the AMAC Assessment Framework

The AMAC Assessment Framework provides a structured conceptual understanding of the areas to be considered for assessment in the AMAC project. An assessment framework is similar to a curriculum framework, but more detailed such that it provides a robust roadmap of areas to assess. It does not specify what is to be taught or how to teach, rather, it specifies what is to be assessed. The framework articulates the learning outcomes to be attained by medical students after completing their regular medical training. It provides a structured conceptual understanding of the areas to be assessed and a reference system for assessment tasks to evaluate the coverage of assessment content. The framework provides substantive foundations for subsequent development, along with technical and practical considerations of what would be appropriate and feasible to assess.

The range of competencies expected of medical students by the time they reach graduation is substantial. These students will need to be able to demonstrate basic competencies in professional practice, professional behaviour and communication. They will also need to possess an integrated body of skills and knowledge.

The framework development was strongly informed by many national and international assessment frameworks and curriculum documents. The most notable of these are *The 2009 Framework for Undergraduate Medical Education in The Netherlands* (NFU, 2009); *The CanMEDS 2005 Physician Competency Framework* (CanMEDS, 2005); the *Australian Curriculum Framework for Junior Doctors* (CPMEC, 2008); the *Australian Medical Council Multiple Choice Examination Specifications Booklet* (AMC, 2011); and the Australian and New Zealand Medical Deans report, *Developing a Framework of Competencies for Medical Graduate Outcomes* (MDANZ, 2011). It is informed by the processes and practices of The Tuning Project (Medicine) – Learning Outcomes/Competences for Undergraduate Medical Education in Europe (Cumming, A., Ross, M., 2009); the AHELO project (OECD, 2010); and the AHELO assessment frameworks (OECD 2011; OECD 2011a).

In addition to this, the development of the framework has taken place over a period of 18 months of consultation with medical education experts, clinicians, stakeholders and students. Consultation has been sought specifically at the two AMAC engagement forums as

well as via email, through reference group meetings, during workshops with clinicians and within the project team itself.

The AMAC Assessment Framework is a key deliverable of this project and can be downloaded at www.acer.edu.au/amac/framework. The Assessment Framework is seen as a 'living document' in that it is anticipated that it may undergo further modifications and incarnations in the spirit of continuous improvement.

Development of the AMAC Pilot Assessment

Item collection and test construction of the AMAC Pilot Assessment was based on the Framework. The pilot AMAC instrument aimed to assess a few key areas of the Framework, rather than initially trying to cover the full spectrum which the Framework represents. In essence, assessing the full range of areas covered by the AMAC Framework is a task much greater than can be achieved in one single assessment. Building test items across the whole Framework is a long-term goal of AMAC. Developing an assessment for this particular project was intended primarily as:

- a proof of concept for utilising the AMAC Framework;
- a means of informing item development processes;
- a conduit for testing the practical implementation of the assessment in medical schools; and
- a mechanism for trialling and developing the reporting of AMAC results.

The AMAC Pilot Assessment specifically focused on developing items in two sub-domains of the Assessment Framework: the Clinical Problems and Conditions sub-domain of the Medical Sciences and Practice content domain, and the Cognitive process sub-domain. Other areas such as the Skills and Procedures sub-domain, the Professional Practice content domain and the Behavioural process sub-domain were not included in the pilot assessment.

The reasons for this are two-fold: firstly, due to funding and time constraints, it is more efficient to develop a test focusing on a few specific areas which can be operationalised using multiple choice questions (MCQs); secondly, in the spirit of feasibility and continuous improvement, the AMAC project team believes that the AMAC assessment instrument will be able to expand once it is instantiated.

The process of assessment item development began with a call to all medical schools in Australia and New Zealand to offer items from their existing assessments to the AMAC team for review, revision and possible incorporation into the AMAC Pilot Assessment. Once items were received, members of the project team from ACER reviewed items collected, mapped them to the Framework, and then re-wrote and revised items to increase clarity and consistency. The items were then scrutinised by panels of item developers at ACER where they were further revised. The revised items were then reviewed by medical clinicians from a range of different disciplines in two assessment revision workshops. During these workshops, items were accepted, revised or rejected. In total, 14 clinicians were involved in the workshops (seven at each workshop). The clinicians represented a wide range of medical disciplines including obstetrics, psychiatry, paediatrics, intensive care, surgery, general practice, gastroenterology and infectious diseases.

Following the workshops, remaining items were then mapped again to the Assessment Framework, and a core selection of items was made based on ensuring a spread across the sub-domains being assessed. The draft final selection was then reviewed by clinicians who participated in the workshops and an external consultant with expertise in medical assessment for coverage of the Framework and consistency.

This process resulted in:

- the collection of 420 items from nine medical schools;
- scrutiny of items in panel sessions by professional test developers and item writers at ACER;
- review, rejection and revision of 300 items in two workshops with clinicians from a range of medical disciplines (one workshop at The University of Queensland, the other at Monash University), resulting in 140 final items;
- review of final items by workshop clinicians and a medical education assessment expert outside the project team; and
- a final collection of 120 MCQs, each with one single correct answer and four distractors, mapped to the AMAC Assessment Framework.

The final 120 items were grouped into six sets of 20 items. From these items and sets, six different test permutations of 100 items were created. This allowed for a number of different versions of the AMAC Assessment to be implemented and for the team to trial a greater number of items than were required for a single test sitting.

The AMAC Assessment is designed for administration entirely online. Once items were completed and test rotations identified, they were inserted into an online assessment delivery platform used by ACER.

Table 1, Table 2, and Table 3 provide an overview of the basic distribution of items for the AMAC Pilot Assessment. The domains and components come specifically from the AMAC Assessment Framework. Therefore, these tables also provide an insight into the structure of these elements of the Framework and insight into the range of different ways in which the AMAC items can be categorised and reported. It is important to note that the percentages are based on the professional judgement of the project team, but that these figures are deliberately malleable. That is to say that it is possible to assess the components listed in the sub-domains below, but that the emphasis on any one of these components in any one assessment incarnation depends on the purpose of that specific assessment.

Table 1: Approximate Process Domain emphasis AMAC assessment instrument

Process Sub-Domain	Component	Percentage
<i>Cognitive</i>	Knowledge	25%
<i>Cognitive</i>	Understanding	75%
		<i>Total = 100%</i>

Table 2: Approximate Content Domain emphasis AMAC assessment instrument

Content Sub-Domain	Component	Percentage
<i>Clinical Problems and Conditions</i>		
<i>System</i>	<i>Respiratory</i>	15%
<i>System</i>	<i>Circulatory</i>	15%
<i>System</i>	<i>Digestive</i>	10%
<i>System</i>	<i>Nervous</i>	10%
<i>System</i>	<i>Musculoskeletal</i>	8%
<i>System</i>	<i>Endocrine</i>	8%
<i>System</i>	<i>Immune</i>	8%
<i>System</i>	<i>Lymphatic</i>	8%
<i>System</i>	<i>Reproductive</i>	8%
<i>System</i>	<i>Urinary</i>	8%
<i>System</i>	<i>Others</i>	10%
		<i>Total = 100%</i>
<i>Medical Speciality</i>	<i>Cardiology</i>	15%
<i>Medical Speciality</i>	<i>Gastroenterology and Hepatology</i>	15%
<i>Medical Speciality</i>	<i>Endocrinology</i>	10%
<i>Medical Speciality</i>	<i>Haematology</i>	10%
<i>Medical Speciality</i>	<i>Oncology</i>	8%
<i>Medical Speciality</i>	<i>Infectious diseases</i>	8%
<i>Medical Speciality</i>	<i>Neurology</i>	8%
<i>Medical Speciality</i>	<i>Urology</i>	8%
<i>Medical Speciality</i>	<i>Immunology</i>	8%
<i>Medical Speciality</i>	<i>Clinical Pharmacology</i>	8%
<i>Medical Speciality</i>	<i>Others</i>	10%
		<i>Total = 100%</i>
<i>Medical Context</i>	<i>Primary care</i>	45%
<i>Medical Context</i>	<i>Emergency department</i>	45%
<i>Medical Context</i>	<i>Others</i>	10%
		<i>Total = 100%</i>
<i>Demographic</i>	<i>Adult health</i>	50%
<i>Demographic</i>	<i>Women's health</i>	15%
<i>Demographic</i>	<i>Men's health</i>	15%
<i>Demographic</i>	<i>Paediatrics</i>	15%
<i>Demographic</i>	<i>Others</i>	5%
		<i>Total = 100%</i>

Table 3: Approximate Clinical Context emphasis AMAC assessment instrument

Clinical Context	Percentage	
<i>Making a diagnosis</i>	50%	
<i>Decision making</i>	15%	
<i>Medical testing</i>	15%	
<i>Medical knowledge recall</i>	10%	
<i>Others</i>	10%	
		<i>Total = 100%</i>

Implementation of pilot AMAC assessment

Two pilot implementations of the AMAC assessment were undertaken for the project. The initial aim of the project team was to conduct a single pilot, but the assessment development process was so well supported by institutions that the assessment was ready for implementation some months earlier than anticipated. As such, the first AMAC pilot, conducted in the two medical schools who were specific partners in this project, was undertaken in November 2011. The second involved seven medical schools and was conducted between April and June 2012.

The first pilot was implemented in order to provide an early trial of the system and methods devised by the project team for administration of the AMAC assessment. It also provided some indicative statistics relating to the items in the test and, through a survey taken by students following the assessment, an early indication of the relevance of the content from the perspectives of the students participating.

The second pilot involved a much larger cohort of students and a wide range of institutions. It was run in order to gather more substantial data for psychometric evaluation of the items, to engage other medical schools in the administration of the assessment, to gather feedback from a large cohort of medical students in relation to the assessment and to provide some insight into the possibility of benchmarking through a common assessment.

All medical schools in Australia and New Zealand were invited to participate in the piloting of the AMAC assessment. In total eight medical schools took part in one of the two pilots. Each participating institution was supported by the AMAC team in the administration of the pilot. Support came in the form of a detailed test administration manual as well as email and telephone support. The AMAC test administration manual included details on technical requirements (i.e. computer specifications, and details for test security), wording and information relating to the assessment that could be used in recruiting students, guidance on enabling students to log into the test system and a script for invigilators conducting the assessment. Each institution was supplied with unique student logins and passwords to access the test.

In essence, the implementation of the AMAC assessment involved institutions:

- Recruiting students;
- Organising access to university-administered computers and creating user 'profiles' to prevent anything other than the assessment being accessed during testing;
- Distributing login details; and
- Invigilating or overseeing students as they took part in the test (this was dependent on the level of security each institution was able to achieve on their computers).

Once students had completed the test, all data was automatically sent to the ACER servers and stored securely. All scoring is undertaken automatically in the online test platform used for AMAC. So no marking or coding was required once items were built into the online system.

Institutions were asked to complete a short questionnaire following the test administration in order to ascertain any problems that may have occurred during testing and to maintain a record of the methods of test administration employed.

Reporting and dissemination of pilot results

Once all participating institutions had completed testing, the test data was downloaded from the ACER servers and cleaned. Analysis was conducted to provide psychometric properties for all items as well as examining the test as a whole. Initial outcomes were presented at the second AMAC Engagement Forum in early June 2012.

Reports were generated for each participating student in the test, providing individual results on items, grouped into a range of categories and results at the national level were provided for comparison. Information about the project was also provided to students as part of the reporting. This information included clear explanations of the data provided and highlighted a number of key caveats to interpretation and comparisons that could be made using the information at the pilot stage of AMAC.

Each participating medical school was also provided with a report detailing the outcomes of their cohort and providing some benchmark figures (again containing caveats about interpretation and comparison). Results for institutions were disaggregated by a number of student characteristics. However, individual students were not identified and if certain groupings of participating students were small in some institutions (for example international students), no results were reported so as to protect anonymity of individuals.

Outcomes and achievements

The intended outcomes of this project are listed earlier in this report. In general, the project has met these intended outcomes. This section addresses each of the seven intended outcomes listed and provides an overview within this discussion of the achievements of the project.

1. The establishment of AMAC as a body to facilitate the ongoing implementation of a national assessment of learning outcomes in the medical field

The current project has provided the foundations for the ongoing implementation of assessment of learning outcomes in medicine. This has been achieved through the introduction of the AMAC idea to institutions through two engagement forums and through the involvement of medical schools in item contribution and piloting of the assessment. The name and acronym AMAC is now readily recognised in Australian medical education, evidenced by the substantial participation in the AMAC forums (involving all medical schools from Australia and New Zealand) as well as the participation of more than 500 students in the assessment. Evidence of the project reaching further into the knowledge of medical students can also be seen in the mentioning specifically of AMAC in an editorial of the *Australian Medical Student Journal* (Ingham, 2011).

The AMAC team believes that this project has facilitated achieving the 'proof of concept' of the general idea of a common assessment of learning outcomes for benchmarking purposes. The team is well advanced with plans for the expansion of the current AMAC partners and

the development of further processes and procedures to ensure AMAC is sustainable into the future.

2. Medical education in Australia will have identified, in a collaborative manner, structures and approaches for assuring the quality of assessment

This project has increased the number of conversations and discussions relating to quality in medical assessment in Australia. This has primarily been facilitated through the two engagement forums held as part of the project, at which medical educators and stakeholders have come together to discuss their ideas about collaborative assessment and openly voice their opinions about levels of quality. However, the forums are only one part of the AMAC project that has provided a platform for these discussions.

Institutions involved in the pilot testing of AMAC have also advised the project team that the implementation of an assessment of learning outcomes such as this project has itself raised awareness about the role of assessment and the critical importance of ensuring high quality assessment tools are used. This project has begun the process of identifying structures and approaches for ensuring quality assessment, through the collection, review and revision of hundreds of assessment items from nine medical schools. The ongoing pursuit of excellence in this area is proposed to continue through AMAC in a subsequent stage of the project.

AMAC has also provided a platform for peak groups to increase involvement in common assessment. Through Professor Wilkinson's role on its Executive, the issues being explored by AMAC have been kept high on the agenda of MDANZ. With a focus on collaboration and inclusion, MDANZ have subsequently formed a special committee to oversee developments in common assessment and the range of collaborative ventures that exist in Australian medical education.

3. Sector-wide involvement in an aspect of educational practice which is gaining serious momentum in the global landscape;

As noted throughout this report, consultation, engagement and collaboration have been key themes of the AMAC project that have been emphasised in all aspects. The involvement of eight medical schools in pilot testing, the donation of more than 400 assessment items from nine medical schools and the participation at engagement forums for the project of all Australian and New Zealand medical schools, along with key stakeholder groups including MDANZ, the AMC and members of other medical assessment consortia is evidence of the sector-wide engagement achieved through this project. In addition, global participation has been included in AMAC primarily through the first engagement forum, which was facilitated by an international expert in medical assessment and collaboration (Prof Cees van der Vleuten from Maastricht University), included presentations from the US National Board of Medical Examiners (NBME) and was attended by international guests from the Malaysian Health Department.

Through this project the AMAC collaboration has evolved from comprising two universities and a not-for-profit organisation into a collaboration involving a range of institutions and stakeholders. This has been facilitated on a number of levels. As has been highlighted in this report, the specific outputs of the project (i.e. the Framework, Assessment and forums)

have included medical schools right across Australia and New Zealand. However, another important aspect to which the project has contributed is at the policy level within the key policy-makers in medical education, specifically Medical Deans. Through the leadership of AMAC, the concepts of common assessment, benchmarking and building of evidence to support quality in medical education have been incorporated more closely in the discourse and activity of key stakeholders such as MDANZ.

The momentum offered by this project has also attracted sector-wide enthusiasm in being a formal part of the AMAC team. The proposed team for the next stage of AMAC is expected to formally involve up to ten medical schools in Australia, a substantial increase on the two medical schools who began discussing the project along with ACER in 2010.

4. A model for the implementation of discipline-based objective testing of learning outcomes;

The Approach and Methods section of this report provides an outline of the methods employed in the implementation of the AMAC assessment for this project. During the implementation the project team were able to hone the processes for development and administration of learning outcomes assessment. The lessons learnt from this project, the development aspects as well as the administration and reporting of results, will be used by the AMAC team to strengthen future iterations of the assessment and to build up a range of resources through future work with the collaboration.

5. Framework and item development for the medical field to facilitate national benchmarking and evaluation of outcomes;

This project has both completed an Assessment Framework and undertaken two pilots of the AMAC Assessment. Details of these are noted in earlier sections of the report, but in summary, the Assessment Framework has been developed over an 18 month period of consultation with medical schools and stakeholder groups. The AMAC Assessment was developed based on the provision of more than 400 items from nine medical schools, which were reviewed and revised by experts to result in 120 items mapped to the Framework and built into an online system for administration to students.

The AMAC Assessment has been piloted in eight medical schools and secured participation of more than 500 students. Student and institutional reports have been developed and distributed to all participants, detailing their outcomes from the piloting.

6. Initial psychometric evaluation of the results of the pilot testing of draft items;

Based on the results of 512 final year medical students who participated in the AMAC pilots, detailed psychometric analysis of all 120 items has been undertaken with the intention of revising and honing items and the structure of the assessment for future administration of the assessment. Analysing technical properties of assessment items offers multiple avenues for continuous improvement and further validation and calibration of assessment materials. Initial results of these analyses were presented at the second AMAC engagement forum. It is intended that further work be conducted on these results following the conclusion of the current AMAC project.

7. Development of an item bank, formed to provide the basis for the future development of a robust assessment instrument in the field.

This project has created 120 assessment items, which are mapped to a detailed Assessment Framework. These items have been used to create six different versions of a 100 item, two-hour instrument focused on the Clinical Problems and Conditions sub-domain of the Medical Sciences and Practice content domain, and the Cognitive process sub-domain of the Framework. The initial collection, review, revision and implementation strategies employed in this project offer important foundations for the continued development of high quality items for use in medical education in Australia. As articulated earlier in this report, the intention of the AMAC team is to expand development into other areas of the Assessment Framework through increased collaboration and cooperation with medical schools into the future.

Dissemination

A key aim of the AMAC project has been to engage medical schools, students and the higher education sector more generally in a conversation about learning outcomes and the use of assessment to facilitate improvements in outcomes. As such there has been a strong emphasis on the dissemination of ideas, findings and methods used throughout the duration of this project. Three key aspects of this dissemination are explored below: consultation relating to the AMAC Assessment Framework, the AMAC engagement forums, and the involvement of institutions in item provision, piloting and reporting. Many of these aspects have been covered already in this report, so the discussion here is relatively brief.

In addition to the specific items discussed below, the AMAC team have also developed a project website (www.acer.edu.au/amac) which provides information about the project and access to many of the documents produced for AMAC.

AMAC Assessment Framework

The development of the AMAC Assessment Framework has taken place over a period of 18 months of consultation with medical education experts, clinicians, stakeholders and students. Numerous opportunities to provide feedback on the Framework have been provided (to the medical education community in particular) throughout the project. In response, revisions to the document and the overall ideas of the project team have been incorporated on an ongoing basis.

The initial presentation of the AMAC Framework took place at the AMAC engagement forum in May 2011. All 53 forum attendees were sent the draft Framework via email, as well as the provision of hard copies at the event. A session presenting the overarching structure and content of the Framework was included in the forum and slides from the presentation posted on the project website. Forum participants were given opportunities to discuss the Framework at the forum and also to provide other feedback to the project team via email or discussions in the days and weeks following the forum.

Further important dissemination of the AMAC Assessment Framework took place as part of the item review and development process. At two workshops for item review, clinicians involved were given the opportunity to review the Framework and provide advice and feedback on the document. Importantly, this particular dissemination was helpful in gathering ideas from a group of people who had been involved in the practical implementation of matching items to the Framework. It offered the AMAC project team new insights into the structure, categorisations and application of the Framework.

The AMAC Reference Group also provided an important avenue for disseminating materials for the project, particularly the Assessment Framework. Members of the Reference Group not only provided their personal views in relation to the Framework, they also offered feedback based on the views of the organisations they represented. As detailed in Appendix A, this group comprises representatives from a number of important stakeholder groups. While not formally a member of the AMAC Reference Group, the AMC has also been closely

informed of the developments of the project. Mr Ian Frank, CEO of the AMC attended a face-to-face day-long meeting of the Reference Group.

The second AMAC engagement forum, held in June 2012 was another opportunity for dissemination of the Framework. By this stage of the project, the Framework had been substantially revised. Discussion at this forum offered another opportunity for review of the concepts and categorisations within the Framework.

The AMAC Assessment Framework is a 'living document', being under constant review and revision. A current version of the Framework is publicly accessible through the AMAC project website at www.acer.edu.au/amac/framework.

Engagement forums

The AMAC project team has held two engagement forums over the duration of the work, both hosted by The University of Queensland.

The first AMAC Engagement Forum was held on May 24, 2011. The forum was facilitated by an international expert in medical assessment and collaboration (Prof Cees van der Vleuten from Maastricht University). In total there were 53 attendees who comprised senior faculty from every medical school in Australia and some international schools, as well as representatives from the AMC, the Malaysian Health Department and the American National Board of Medical Examiners (NBME). It was a highly successful day, opening up a dialogue within the sector about assessment of university outcomes and how this might work in an Australian setting, while keeping in touch with developments across the world. The Forum allowed the project team to highlight the plans for this project and the ongoing establishment of AMAC, while also providing plenty of opportunity for participants to hear about what else was happening across the world and what they thought relevant to their institutions. More information about the forum is available at: <http://www.acer.edu.au/amac/engagement-forum-held-at-university-of-queensland-may-2011>.

The second AMAC engagement forum was held on June 5, 2012. This forum involved 33 attendees, representing most medical schools in Australia as well as key stakeholders such as the AMC, MDANZ and the OLT. The day provided the opportunity for updating participants on the progress of the project, dissemination of some of the results from the piloting, in-depth discussion of the AMAC Assessment Framework, and debate about what the future of AMAC might look like and what medical schools hope to gain from the collaboration.

A list of institutions and stakeholder organisations represented at each of the engagement forums is included in Appendix B.

Institutional involvement

Dissemination of project materials, ideas and outcomes has occurred in this project with individual institutions at a number of specific points throughout the project. Direct involvement with institutions has come through the engagement forums, but also through

the practical implementation of the project. The engagement forums described above generally facilitated discussion of the overall project, while interactions individually with institutions provided an opportunity for dissemination of methods and practical issues to occur. It is the nature of these individual interactions with schools that are outlined here.

As noted earlier in the report, nine medical schools voluntarily provided the AMAC team with exam items for review and possible inclusion in the AMAC Assessment (see Appendix C for a list of schools). The process of informing institutions about this and collecting the items was in itself an important avenue for disseminating ideas and aims of the AMAC project to individual schools. Initial requests were made to each Dean through individual letters explaining the project and the aim intended in collecting items. Later, interaction by the AMAC team with individual members of staff from these institutions (often assessment coordinators) was also an opportunity for discussing the aims and focus of the project and informing people within the medical schools about how the process would develop.

Seven institutions were also closely involved in the implementation of AMAC through the piloting of the AMAC Assessment (see Appendix C for a list of institutions). This process involved the dissemination of a detailed manual for institutions, explaining how to go about implementing the assessment. Importantly, it also put the AMAC team in direct contact to provide advice, talk through objectives and facilitate the piloting of assessment materials.

The AMAC Assessment pilot was also important in connecting with medical students and disseminating the ideas behind the assessment and the objectives of the project. More than 500 students volunteered to be involved in the assessment pilot and in doing so were informed about the project, provided with a student report and given the opportunity to explore the collaboration in more detail through visiting the project website and contacting the project team. The AMAC team received emails and phone calls from a number of medical students involved in the project offering support for the concept. Participating students also provided direct feedback about the AMAC Assessment through a short survey undertaken at the conclusion of the pilot testing session.

Linkages

This project has been involved in creating a number of important national and international linkages in the relation to the ALTC Strategic Priority Area 'Principles for Assessment and Assurance of Graduate Learning Outcomes'.

Importantly from the point of view of the project team, the AMAC work has created links with other collaborative groups within the medical education community. These have included the AMSAC, ACCLAiM and IDEAL groups who have all been involved in the AMAC engagement forums and have member schools involved in the piloting of the AMAC Assessment. These linkages have been important in discussing approaches and aims and highlighting the complementarities of the various projects and collaborations.

AMAC has also helped to substantially develop the discussion within the medical education community regarding common assessment across medical schools. While the idea of a common assessment of learning outcomes across institutions was an issue which seemed relatively radical in the early stages of this project, there now appears to be almost uniform acceptance that there is a need for development of common assessments to enhance processes aimed at building continuous improvement. Key evidence of this is the formation during the period of this project of an MDANZ committee with a specific remit on developing policy relating to common assessment and assisting in providing links for groups involved in these collaborations.

The AMAC team has also been successful in using this project to develop the involvement and interests of other medical schools. A proposal being developed for the OLT for a new project based around the foundations developed here sees the AMAC project team increase from two to ten medical schools, now representing a substantial number and variety of institutions in Australia.

Evaluation

The evaluation framework for this project closely followed the ALTC/OLT guidelines and focused on a multi-stage, multi-stakeholder approach allowing for ongoing evaluation, reflection and change throughout the project.

Associate Professor Heather Alexander (Griffith Institute for Higher Education) has been engaged to evaluate this project. Heather has been involved in project meetings, engagement forums and Reference Group meetings throughout the project. Her involvement has been extremely important for the project team and has involved ongoing feedback and contribution to the project. An Interim Evaluation Report was completed and presented to the project team in December 2011.

The Final Evaluation Report by Associate Professor Alexander is contained in Appendix D of this report.

Areas of focus of the evaluation framework are detailed in the table below.

Table 4: AMAC Evaluation Framework

Evaluation Criteria	Key questions addressed by evaluator	Possible evaluation methods
Project aims	<ul style="list-style-type: none"> • What are the aims and scope of the project and how have they changed over the course of the project? • To what extent has the project been successful in building consensus for a national medical student assessment? • To what extent has the project been successful in engaging Medical Schools, and staff within those Schools, across Australia? • Does the framework reflect the views and expectations of stakeholders? • What processes are in place to ensure high quality items? • Is the framework perceived as appropriate by medical schools and do the prepared items sufficiently cover the areas identified as important in the framework? 	<p>Review of documents and record of refinement of aims during project meetings</p> <p>Outcomes of discussions at Forum</p> <p>Evidence of involvement of medical schools through the project; survey at the end of the project</p> <p>Feedback from Framework Workshops (2011 and 2012); survey at the end of the project.</p> <p>Feedback from reviewers and evaluator (by observation) re: item review process</p> <p>Structured feedback from reviewers of submitted items and framework; feedback from review process; end of project survey/ interviews</p>
Impact on stakeholder engagement	<ul style="list-style-type: none"> • Has the project involved all relevant stakeholders? • Has the project sufficiently sought involvement and input from academic and support staff from medical schools? 	<p>Forum attendance and range of stakeholders represented</p> <p>Participation of Med Schools in forum / item contribution / test administration</p>

Effectiveness of dissemination strategies	<ul style="list-style-type: none"> • How effective was the Engagement Forum, consultations and workshops in providing further feedback, evaluating and refining the framework? • What has been the impact of the project outcomes, deliverables and dissemination strategy? 	<p>Outcomes of workshops at forum; end of project structured, short interviews or surveys of impact & activity in Med Schools</p> <p>End of project structured, short interviews or surveys with staff in Med Schools re use to date of framework, items, and future uses, needs for change etc</p>
Ongoing effectiveness of project processes	<ul style="list-style-type: none"> • What changes /amendments needed be made to ensure the project meets its intended aims? Were there any variations from the processes that were initially proposed, and if so, why? • What unintended benefits accrued from the project? • What factors helped/hindered in the achievement of the outcomes? 	<p>Record of major changes made during the project</p> <p>End of project structured, short interviews or surveys with staff in Med Schools</p>
Sustainability of AMAC	<ul style="list-style-type: none"> • Has a model been set in place for the effective and financially viable continuation of AMAC? • Are there sufficient processes in place to sustain and maintain the item bank created from this project? 	<p>External evaluator-use structured, short interviews or surveys with team members and staff in Med Schools</p>
Timeliness	<ul style="list-style-type: none"> • Were timelines managed appropriately? • What strategies were in place for risk management? 	<p>Feedback from Evaluator</p>
Budget	<ul style="list-style-type: none"> • Did the budget describe accurately the extent of time required to undertake the project? 	<p>Feedback from Evaluator</p>

References

- Aaron, S., (2009), Moving Up the Pyramid: Assessing Performance in the Clinic, *The Journal of Rheumatology*, 36(6): 1101-1103.
- ACGME, (2005), *The ACGME Outcome Project*, Chicago: Accreditation Council for Graduate Medical Education.
- AMA, (2010), *Position Statement: Competency-based training in medical education*. Australian Medical Association. Accessed 9 May 2011 from: <http://ama.com.au/node/5977>
- AMC, (2010), *AMC Accreditation Report*, University of Queensland, School of Medicine.
- AMC, (2010), *Competence-based medical education: an AMC Consultation Paper*, The Australian Medical Council. Accessed 3 May 2012 from: http://www.amc.org.au/images/publications/CBEWG_20110822.pdf
- AMC, (2011), *Multiple Choice Examination Specifications Booklet*, Accessed 3 May 2012 from: http://www.amc.org.au/images/publications/amc_exam_spec.pdf
- AMC (2012), *The List of Australian Recognised Medical Specialities*, Accessed 3 May 2012 from: <http://www.amc.org.au/images/Recognition/AMC-list-of-specialties.pdf>
- Archer, J., (2009), European licensing examinations – the only way forward, *Medical Teacher*, 31:215-216.
- CanMEDS, (2005), *The CanMEDS 2005 Physician Competency Framework*, Ottawa: The Royal College of Physicians and Surgeons of Canada. Accessed 12 May from: http://rcpsc.medical.org/canmeds/bestpractices/framework_e.pdf
- Carreras, J. (2008), *Assessment of Undergraduate Medical Education Competencies*, AMSE Annual Conference Barcelona 2008. Accessed 9 May 2011 from: <http://www.amse-med.eu/documents/Carreras.doc>
- Coates, H., Wilkinson, D., Glasgow, N., Richardson, S., (2011), Generalisable assessment of learning outcomes in Australian medical education: options for consideration. National Assessment Forum, The University of Queensland, Brisbane.
- CPMEC, (2008), *Australian Curriculum Framework for Junior Doctors*, Accessed 4 May 2011 from: <http://www.cpmecc.org.au/ACF-2010/index.cfm>
- Cumming, A., Ross, M., (2009), *The Tuning Project (Medicine) – Learning Outcomes/Competences for Undergraduate Medical Education in Europe*. Edinburgh: University of Edinburgh. Available online: <http://www.tuning-medicine.com/pdf/booklet.pdf>
- DEEWR, (2008), *What makes for success in Medical Education: Australian Medical Education Study*, Canberra: Department of Education, Employment and Workplace Relations.
- González J. and R. Wagenaar (eds.) (2008). Tuning Educational Structures in Europe. Universities' contribution to the Bologna Process. An Introduction. 2nd. Ed, Bilbao-Groningen.
- Gorsira, M., (2009), The utility of (European) licensing examinations, *Medical Teacher*, 31:221-222.
- Harden, R., (2009), Five myths and the case against a European or national licensing examination, *Medical Teacher*, 31:217-220.
- Ingham, A. I., (2011), The great wall of medical school: A comparison of barrier examinations across Australian medical schools, *Australian Medical Student Journal*, editorial, v2, 2:5-8.
- Laan, R., Leunissen, R., van Herwaarden, C., (2010), The 2009 Framework for Undergraduate

- Medical Education in The Netherlands. *Undergraduate Medical Education*, 29: 10-15.
- McAllister S., Lincoln M., Ferguson A., McAllister L., (2006) *COMPASS (competency assessment in speech pathology) assessment and resource manual*. Melbourne: Speech Pathology Australia.
- Melnick, D., (2009), Licensing examinations in North America: is external audit valuable? *Medical Teacher*, 31:212-214.
- Miller, G., (1990), The Assessment of Clinical Skills/Competence/Performance, *Academic Medicine*, 65, S63-7.
- MDANZ (2011), *Developing a Framework of Competencies for Medical Graduate Outcomes*, Accessed 3 May 2012 from: <http://www.medicaldeans.org.au/wp-content/uploads/Competencies-Project-Final-Report1.pdf>
- NFU (2009), *The 2009 Framework for Undergraduate Medical Education in The Netherlands*, Utrecht: Nederlandse Federatie Van Universitair Medische Centra. Available online: http://www.nfu.nl/fileadmin/documents/Raamplan2009engelstalige_versie.pdf
- OECD, (2005), *DeSeCo Project: - The Definition and Selection of Key Competencies: Executive Summary*, Paris: Organisation for Economic Cooperation and Development. Available online: <http://www.oecd.org/dataoecd/47/61/35070367.pdf>
- OECD, (2010), *Assessment of Higher Education Learning Outcomes: Feasibility Study*, Paris: Organisation for Economic Cooperation and Development.
- OECD, (2011), *Module B Assessment Framework (Economics)*, Paris: Organisation for Economic Cooperation and Development.
- OECD, (2011a), *Module C Assessment Framework (Engineering)*, Paris: Organisation for Economic Cooperation and Development.
- Rychen D. S. and Salganik, L. H. (Eds.). (2003). *Key Competencies for a Successful Life and a Well-Functioning Society*. Göttingen, Germany: Hogrefe and Huber.
- Van der Vleuten, C. (2009) National, European licensing examinations or none at all? *Medical Teacher*, 31:189-191.
- Van der Vleuten, C., Schuwirth, L., Muijtjens, A., Thoben, A., Cohen-Shotanus, C. & van Boven, C., (2004), Cross institutional collaboration in assessment: a case on progress testing, *Medical Teacher*, 26:8 719-725.
- Van der Vleuten, C., Schuwirth, L., Scheele, F., Driessen, E., Hodges, B., (2010), The assessment of professional competence: building blocks for theory development, *Best Practice and Research Clinical Obstetrics and Gynaecology*, 24:703-719.
- Wojtczak, A. (2002), Glossary of Medical Education Terms: Part 1, *Medical Teacher*, 24(2). 216-219. Available online: http://esmeprogramme.org/amee2006/pdf/Occasional_03.pdf

Appendix A: Project Reference Group

- Professor Richard Hays – representative of Medical Deans of Australia and New Zealand (MDANZ).
- Ms Siobhan Lenihan – representative of the Office for Learning and Teaching (OLT).
- Mr James Churchill – President of the Australian Medical Students’ Association (AMSA).
- Associate Professor Terry Brown – representative of the Confederation of Postgraduate Medical Education Councils (CPMEC).

Appendix B: Institutions and organisations represented at AMAC engagement forums

First engagement forum, May 24, 2011:

Australian National University
Bond University
Deakin University
Flinders University
Griffith University
James Cook University
Monash University
The University of Auckland
The University of Melbourne
The University of New South Wales
The University of Notre Dame Australia
The University of Queensland
The University of Sydney
The University of Western Australia
University of Adelaide
University of Newcastle
University of Otago
University of Tasmania
University of Western Sydney
University of Wollongong
International Medical University, Malaysia
Maastricht University
Australian Medical Council
Ministry of Health of Malaysia
National Board of Medical Examiners
Queensland Health
The University of Queensland Medical Society

Second engagement forum, June 5, 2012:

Australian National University
Bond University
Deakin University
Flinders University
Griffith University
Monash University
The University of Melbourne
University of New England
The University of New South Wales
University of Newcastle
University of Notre Dame Australia
University of Otago

The University of Queensland
The University of Sydney
University of Tasmania
University of Wollongong
Australian Medical Council
Australian Medical Students' Association
Medical Deans Australia and New Zealand
Office for Learning and Teaching

Appendix C: Institutions involved in item provision and pilot testing

Medical schools providing items to the AMAC Assessment:

Deakin University
Griffith University
Monash University
The University of Queensland
The University of Western Australia
University of Newcastle
University of Notre Dame Australia
University of Western Sydney
University of Wollongong

Medical schools involved in the pilot of the AMAC Assessment:

Monash University
The University of New South Wales
The University of Queensland
The University of Sydney
University of Notre Dame Australia, Fremantle
University of Notre Dame Australia, Sydney
University of Otago
University of Wollongong

Appendix D: Evaluation Report

Associate Professor Heather Alexander

1. External Evaluator role

The external evaluator has been closely associated with the project team throughout the project and has attended both National Forums, participated in teleconferences from July, 2011 and has observed one of the two item review sessions. An ACER student feedback survey was attached to the pilot assessments with data returned to ACER. At the time of writing this report, feedback from the student evaluation of those examinations was not available to the evaluator.

At the end of the project, the external evaluator invited all participating medical schools to be interviewed. One representative, who had been closely involved in the project in each of seven of the nine medical schools were interviewed. Two representatives of another school were interviewed to ensure coverage of roles in the project (administrative and academic). One school did not participate in the interviews. The questions asked of all interviewees are attached to the end of this report. The interviews were all recorded and analysed, and respondents were assured that their responses would be de-identified.

As the project team is planning further work, this report will comment on work to date as well as giving recommendations to consider in future work.

2. Project aims

The project team has achieved a great deal in a relatively short time frame and although they may not have completely met the aims as originally written with a broad view of graduate outcomes of medical programs and the need to assure standards in medical education, the team has built a strong foundation for achieving those aims with further work. Key challenges will be to move to the use of a summative use of the assessment (further detail is given below in the findings of stakeholder interviews) and broadening the assessment formats to those other than Multiple Choice Questions. Further work will be needed to meet the goal of 'establishing AMAC sustainable procedures'.

3. Deliverables and Timelines

Referring back to the timeline in the revised project proposal, the project team has met its timelines for Scoping foundations, Framework development and Item production and validation and has successfully run two pilot assessments consisting of multiple choice questions across the participating universities.

4. External evaluator feedback from observations of review sessions

The project team established multidisciplinary clinical panels to review the quality and accuracy of the questions and agreed to clinician final sign off on the test content. It was

beneficial for the external evaluator to sit in and observe part of the Queensland item review process. From this observation and participation in the project team teleconferences, it was clear that the review session ran well and covered a large number of questions. Offering additional distractors for each Multiple Choice Question from which reviewers can select replacements, was an effective practice to use. The case-based nature of questions is to be commended and is in line with the stated aims of the project. Although most questions were case-based, some questions were set at fact recall level. It was stated in the revised project plan that AMAC frameworks and items will “... *assess student ability to apply broad concepts to solve problems rather than being based on course curriculum and content*’ and that it would be pitched ‘*above content*’. The fact recall questions would not appear to be achieving this goal, although it should be noted that the external evaluator did not see the final examination paper.

It was observed that there was considerable variation in the length of the stem for each question which may need to be considered in the time allocation for the test. This aspect was also commented on by one respondent in the stakeholder interviews. The number of questions that can be reviewed in a day is an issue for consideration in planning for future sustainability. The discipline mix of the reviewers is an important component of the review process.

For this work to be scaled up in the future, consideration will need to be given to the security of distribution of materials for review and to the use of a paper-based process for review of questions.

5. Validity of the assessment and the Framework

After the first review session and pilot, it was evident that the framework needed review. The questions observed at the first review session seemed to have a heavy emphasis on diagnosis and very little emphasis on other clinical tasks (eg management). This could have been simply due to the mix of questions allocated to the first review panel and may not reflect the composition of the final test. However, the framework did not address this aspect and the test needed a blueprint that covered the nature and focus on various clinical tasks. The use of body systems only in the Item Specification to determine the mix of questions in the first pilot, without consideration of disciplines and without *simultaneously* identifying the clinical task involved, could affect both the content and construct validity of the test (depending on its defined purpose). The disciplines to be represented in the test, and the proportional representation of each, needed to be clarified.

The framework has subsequently been reviewed and the revised version tabled at the second National Forum, addresses many of these concerns. As the pilot assessments focussed on only certain sub-domains within the framework, the validity of the revised framework will need testing in future work.

6. Findings from stakeholder Interviews

Interviews with key staff at eight of the nine participating medical schools were conducted during the last half of June and early July. All staff interviewed had a close association with

the project. The interview questions were drawn from the evaluation framework, the principles of assessment and feedback from team meetings. The final set of interview questions were sent to the project team for feedback, prior to the interviews. A copy of the interview questions is included at the end of this report.

6.1 Aims

The stakeholders were asked to what extent the project had achieved its aims, described in the project proposal as: *The aim of AMAC is to set foundations of a national assessment to monitor the outcomes of later year medical students in Australia...This ALTC project will provide the foundation for what will be the ongoing development and implementation of an item library that will provide a sustainable and robust means of assuring the standards of medical education in Australia.*

The stakeholders commented positively on the achievements of the project in the time available. Most stakeholders felt that the project had laid the foundations for the full aims (as stated above) to be achieved in the future. The comments indicated that the project had been successful in demonstrating the feasibility of the process (collection and review of MCQs across institutions) but that more work would be needed to fully achieve the stated aim of developing an 'item library'. One respondent commented that full feedback on this could not be given until all results were available to the schools. One respondent indicated that an unstated aim had been achieved in that at the National Forum, "... people arriving with interest and curiosity and leaving having changed their perceptions to this exercise [saying] 'we want to be involved', 'how can we help?' and that this represented an attitudinal shift.

6.2 The process for collection/ development of items

Two respondents were not directly involved in this part of the process and could not comment. The remainder were either directly involved or had received feedback from staff about the process. The respondents indicated that the process, administered by ACER, went well, was effective and was straightforward. One respondent indicated that although they contributed questions to the trial they were not involved in the further question review process and they felt it would have been an improvement had they been involved.

6.3 Multidisciplinary Review

All schools strongly supported the process of multidisciplinary review of questions and some used this process in their own examination review.

6.4 Question Format/Quality

Two respondents commented that they had not had the opportunity to see the final paper while two indicated that they had seen a small number of questions while invigilating. One of those respondents indicated that the questions were of high quality, while another commented that while some questions had a very short question stem, with some being a single line only, most were longer stems using a case-based approach. One respondent indicated that the question format varied in format from their usual university format (in the number of distractors) but this did not cause any problems. One respondent indicated that the question wording of some questions required further review.

6.5 Timing of the Examination for Student / Student approach to the examination

Students in all institutions approached this as a voluntary, formative exercise for feedback on areas they need to study. Overall, students were positive about the opportunity for feedback, although one respondent mentioned that students would have preferred more immediate feedback. One respondent commented that even though the students approached the examination reasonably seriously, they would not have been as prepared as they would have been for final examinations, given the timing of the test in the early part of the year. Due to the timing of the trial, not all students had completed all of the required rotations and so would not have been fully prepared for the examination. Another commented that students with higher levels of ability tended to be the ones to take up the opportunity to sit the test. It was considered that these issues limited the potential of the exercise for benchmarking purposes.

6.6 Online administration

All universities approached this as a formative exercise and under this condition, the examination ran smoothly, with the organisation by ACER being praised. Most universities ran the examination on university-owned computers although one site allowed students to use their own computers. The feasibility of running this type of examination under summative conditions would be problematic due to: (1) student cohorts being spread over multiple locations; and (2) the lack of computer labs in some institutions, with security issues preventing the use of students' own laptops. One institution had concerns over the configuration of usernames and passwords for student access, and another mentioned difficulty with having the correct version of software. These concerns have been communicated to ACER by the institutions. One institution with considerable experience in running online examinations found that it ran smoothly although raised concerns should the examination contain images.

6.7 Impact on student learning

Respondents saw value in the formative nature of this trial, in identifying areas of strength and weaknesses for students and in helping to ease the stress that builds up ahead of final examinations.

The potential for the examination blueprint and content to impact on what is taught was recognised. One respondent suggested that having the shared examination as only part of overall student assessment would help to safeguard against this. Two institutions commented on the need to ensure the blueprint covered both knowledge-based and clinical examinations to drive student learning in desirable directions.

6.8 Analysis/Scoring

Five institutions had not seen the results or analysis at the time of interview and so were unable to comment. Others commented that the analysis was similar to that done in other collaborations of which they were part. One respondent commented on the future need to address standard setting being a big issue. Two respondents commented on the need for more detailed feedback to schools (ensuring data available at discipline level) and one commented on the need for summative assessment data if it was to be used for benchmarking.

6.9 The Framework

Most respondents did not provide detailed comments on the framework, through not being able to recall the detail or not having it to hand at the time of interview. This may reflect a lack of engagement of the schools with the latest version of the framework. Most commented positively on it in a general sense. One respondent commented on the limitations of the early version of the framework but the later version was acceptable as a knowledge-based framework. It was noted by one respondent that teaching in the later years of the medical program tended to be discipline-based rather than systems-based.

6.10 Value in this initiative for moderation / benchmarking

The respondents were mostly positive about the benchmarking potential for this initiative, with particular benefit for new schools. The current implementation as a formative, voluntary exercise was seen as limiting this potential for benchmarking. One respondent commented that an increase in participation by schools would be needed to ensure its value in benchmarking. One respondent commented that this iteration was of greater benefit to the students than for benchmarking schools as it was a voluntary although it laid the foundation for cooperation between schools. Another respondent commented on a significant cultural shift that has occurred and has led to broader acceptance of such an approach to sharing assessment. An additional benefit of sharing assessment was seen to be alleviating the pressure on time and resources within schools and to improve the overall quality of assessment items. One respondent felt that the initiative had potential for future research looking at the diverse range of programs and the effects on student performance.

6.11 Unintended Consequences

The level of collaboration and the development of goodwill and relationships between medical schools were seen very positively. One respondent said: *'although no one is opposed to it [collaboration]... sometimes you need an avenue for it to be facilitated and this was a great facilitation'*. One respondent hoped that this would extend to other areas of learning and teaching. Another respondent commented on the unexpectedly positive reaction of the students to the test. The trial of the online administration of the examination was seen as a positive by-product of this trial and would open the door for the school to undertake further online examinations. For one respondent, the limitations of the university IT systems was a surprise, as was the potential research interactions from such an initiative.

6.12 Future sustainability

Three of the seven respondents commented on the need for ongoing financial support for this process, necessitating the use of efficient processes. If it was to expand to an increased number of questions or to a summative assessment, financial support would be needed for the data analysis and management of the item bank. Two respondents indicated a sharing of costs by participating schools might be possible. One respondent commented that ease of access to shared assessment, for example through the use of a database, would be needed and that it would be important to embed an ongoing collaboration between schools rather than relying on one-off funding. That respondent felt that the process *'...needs a sustained organisation to oversee it.'* A further respondent suggested that the success of this work in the future would also be dependent on the careful handling of the future collaborations due to the breadth of diversity of the approaches to assessment, and that it would be preferable for medical schools to build that collaboration rather than having a

national test imposed externally. One respondent commented that the involvement of a commercial organisation external to the medical schools was problematic. While a commercial external organisation might be used to collate, cleanse and de-identify data, the respondent felt that the analysis, question development and blueprint development should be in the hands of the medical school as sufficient expertise was available.

6.13 Willingness to be involved in the future

All respondents indicated that they would consider participating further in the AMAC work as it develops and that there had been high interest generally in this project. One respondent indicated that their school would be happy to submit questions for formative assessments and that the issue of the timing of the assessment over the academic year would need to be considered.

6.14 Formats of assessment for future trials

There was general support for broadening the assessment to other formats. Four respondents indicated that MCQ or written assessments were the easiest to use for this purpose. Other formats suggested were:

Clinical / OSCE (6) (with one respondents commenting upon the wide disparity of rotations in the last year of medical school making this difficult)

EMQ (2)

Key features questions (1)

Short answer (1)

Script concordance (though needs further development) (2)

Workplace-based assessment (2)

7. Summary and Future considerations

As summary of the evaluator's observations and the feedback from the stakeholders and issues for consideration should further work be undertaken by the project team, are described below.

- There has been strong positive support for the collaboration evident in this project and for ongoing collaboration across medical schools
- The aims of this work need to be further clarified by the medical schools and MDANZ in terms of benchmarking
- While the value of this assessment as a benchmarking exercise is supported, it is dependent on its summative use rather than as a voluntary, formative assessment and there are a number of hurdles to consider in further work:
 - There is some concern amongst stakeholders that the item bank might not be large enough
 - The timing of the examination needs consideration in terms of diversity and timing of student rotations in the final year
 - For the pilot, some institutions had to run this assessment over several days
 - There is diversity in the timing of final examinations in medical schools

- Online administration:
 - The online administration limited the assessment to a formative assessment for this project, with universities sometimes lacking the facilities to run simultaneous summative assessments across multiple sites. One stakeholder felt that the need to run the examination at different times across multiple sites may preclude its use as a summative assessment.
- Involvement of stakeholders
 - For future trials, ensure all schools participating have the opportunity to participate in the multidisciplinary review panels and that the stakeholders have access to the final selected questions at some stage in the process
- Assessment formats
 - There is a need for further work to broaden the assessment formats, particularly to include some form of clinical assessment, to enable the assessment of the integration of knowledge and skills
- Framework
 - The schools do not appear to have engaged well with the latest version of the framework and this will need review if further work is undertaken.
- Analysis and Reporting
 - Given the timing of this report, some schools had not fully engaged with, or sometimes seen, the results and analysis from the trial. The usefulness of the data provided to schools will need further review.
- Standard Setting:
 - Future projects should address the issue of standard setting of the passing score for the examination
- Sustainability
 - Ongoing support for embedding the collaboration across schools and the sustainability of this process will need to be a focus of future work.

**Australian Medical Assessment Collaboration (AMAC)
End of Project Evaluation – Stakeholder Evaluation**

These questions will be used in semi-structured interviews with key staff from participating medical schools.

1. What has been your involvement with this project? (you personally and your school –interviewer to ask about which pilots this school was involved in and to check the role of the person being interviewed)
2. Please comment on the extent to which the project has met its original aims described in the project plan as:

The aim of AMAC is to set foundations of a national assessment to monitor the outcomes of later year medical students in Australia...This ALTC project will provide the foundation for what will be the ongoing development and implementation of an item library that will provide a sustainable and robust means of assuring the standards of medical education in Australia.

3. Please comment on the effectiveness of the **process for the collection/ development** of items.
4. The items were **reviewed** in multidisciplinary panels. Were you directly involved in this process? Do you have any comments on this process?
5. What feedback would you have on **the nature/ format/ quality of the final questions**?
6. What were your perceptions of the **reactions of students** to the examination?
7. What do you think the **impact** could be on **student learning**?
8. Please comment on the nature of the analysis/ scoring used.
9. Please comment on the **feedback** your school received from the examination.
10. Please comment on the framework used to construct the examination **at the time of the test** and the **latest version** of the framework.
11. Was there value in this initiative for your School in moderation/ benchmarking?
12. Have you identified any unintended outcomes (positive or negative) to date?
13. What are your suggestions for establishing this as a sustainable process?
14. Would you be prepared to continue to be involved should this work continue?
15. What formats of assessment should be included in any future work?
16. What is your view of the best future directions for this work?