Medical Education Assessment for 21st Century Health Systems

AMC Workshop

**Workshop Report
1 November 2017**

Contents

Medical Education Assessment for 21st Century Health Systems Workshop

Workshop purpose

In this workshop we:

1. **Fundamentals of Programmatic Assessment** Gained an understanding of the fundamentals of Programmatic Assessment

2. **Common problems and innovations** Reviewed common problems and innovations in assessment across the medical continuum and at the AMC to understand alignment with programmatic assessment concepts and AMC standards

3. **Good ideas and burning questions** Had opportunities to share good ideas and ask burning questions about assessment from experts and peers

4. **Practical Strategies** Gained practical strategies for how to design and implement a programmatic approach to assessment

5. **Implementation – case studies** Shared information about a range of pilots of National and International innovations in assessment across the medical continuum and at the AMC relating to programmatic assessment

6. **Reflections on improving own training program and next steps** Reflected on how to further improve assessment practices in your own training program and future directions for review of AMC standards on assessment

Workshop themes

This workshop aimed to explore the following themes of assessment:

a. **Fit for Purpose:** ensuring that methods and approaches focus on best practice medical education evidence and assess skills and behaviours relevant to 21st century health needs and systems.

b. **Feedback and culture:** improving feedback and learning cultures to ensure learners are empowered and supported.

c. **Discrimination of performance:** improving methods of identifying good and poor performance through longitudinal assessment systems. This is a shift from more traditional norm referenced assessments, which are more focused on comparing candidates against each other rather than of their achievement of standards, and assessment for learning methods to ensure each individual doctor is the best they can be.

d. **Benchmarking, technology enabled reporting and research and evaluation:** ensuring equity in decision making, and feasibility and quality of innovation of assessments through benchmarking assessment methods across the continuum, conducting research and evaluation and enabling decision making with use of technologies for monitoring progression and quality programs.

e. **Implementation**: ensuring high quality, feasibility, sustainability, cost effectiveness, and acceptability of assessment systems. Successful implementation includes:

   i. robust communication and management of change process

   ii. effective governance and champion model of ongoing support and leadership

   iii. operational management to ensure smooth delivery of the assessment

   iv. supervisor/assessor training to ensure consistency of decision making and quality of feedback

   v. plans for evaluation and quality control to inform future quality improvements to assessments

   vi. ongoing resource and stakeholder management to ensure continuity of support.

Schedule

8.30		Morning tea on arrival
9.30		**Forum open**
9.30		Setting the Scene
		Welcome from President
		Presentation: Setting the Scene – Assessment at the AMC
9.50		Programmatic Assessment and the Medical Continuum
		Presentation: Programmatic Assessment Fundamentals – history of assessment from psychometrics to progression and subjectivity. Q & A
		Group Activity 1 and Plenary Discussion: Where are we with programmatic assessment concepts – summary of current state
11.00		Programmatic Assessment – What Problems are we Solving?
		Group Activity 2: Good Ideas and Burning Questions
11.30		**Morning Tea Break**
		Q and A Good Ideas and Burning Questions (cont.)
		Presentation: Problems Assessment Systems are Solving
13.10		**Lunch Break**
14.00		Focus on Implementation
		Group Activity: Five Case Studies of successful implementation of Programmatic Assessment concepts
		(Participants attend 3 case studies – one in each of the three sessions – see back of name badge for details)
		Tea/Coffee Available
15.20		Where to Next for Assessment?
		Panel Discussion: Next Steps in Medical Education Assessment for Health Systems of the 21 Century
		Next steps
16.15		**Meeting close**

Participants

Facilitators and Presenters	Institution	Role
Professor Cees van der Vleuten	School of Health Professions Education, University of Maastricht, Netherlands	Scientific Director
Professor Lambert Schuwirth	Flinders University	Strategic Professor for Medical Education and Director of the Flinders University Prideaux Centre for Research in Health Professions Education
Ms Christine Cook	GPEx and ModMed	Chief Executive Officer
Professor Kichu Nair AM	School of Medicine and Public Health / HNE Local Health District	Professor of Medicine and Deputy Dean (Clinical Affairs) / Director, Centre for Medical Professional Development
Dr Beth Mulligan	Launceston General Hospital	Director of Clinical Training
Professor Liz Molloy	Melbourne Medical School, the University of Melbourne	Professor in Work Integrated Learning in the Department of Medical Education
Professor Pete Ellis	The University of Otago	Head, Department of Psychological Medicine
Professor Liz Farmer	AMC Prevocational Standards Accreditation Committee	Chair
Professor David Prideaux	Prideaux Centre for Research in Health Professions Education, Flinders University	Director and Chair of Assessment Committee, AMC

Participant	AMC Committee	Role
Professor Jill Sewell	AMC	President
Associate Professor Kim Rooney	AMC	Director
Mr Ian Civil	Specialist Education Accreditation Committee (SEAC)	Committee member
Ms Jacqui Gibson	Specialist Education Accreditation Committee (SEAC)	Committee member

Mrs Helen Maxwell-Wright	Specialist Education Accreditation Committee (SEAC)	Committee member
Mr Rob Thomas	Medical School Accreditation Committee (MedSAC)	Committee member
Participant	**Health Department**	**Role**
Dr Linda MacPherson	NSW Ministry of Health	Medical Advisor
Dr June Song	Department of Health and Human Services - Tasmanian Health Service, Southern Region	Director of Medical Education & Training
Ms Megan Crawford	Queensland Health, Prevention Division	Director - Office of the Chief Medical Officer
Ms Rachel Hoffman		Manager - Office of the Chief Medical Officer
Dr Jeanette Young		Chief Health Officer & DDG

Participant	**Specialist College**	**Role**
Mrs Anna Kaider	Australasian College for Emergency Medicine	Education Development Projects Lead
Mr Bernhard Liedtke		Education Systems Development Manager
Dr David Burdon-Jones	Australasian College of Dermatologists	Chief Examiner
Dr Catherine Drummond		Chief Examiner Elect
Mr Tony Moore		Senior Academic Support Officer
Mr Brett O'Neill		Director, Education Services
Dr David Bolzonello	Australasian College of Sport and Exercise Physicians	Chair of Training
Dr Corey Cunningham		Chief Censor
Dr Meredith Craigie	Australian and New Zealand College of Anaesthetists	FANZCA, FFPM, Staff Anaesthetist
Mr Maurice Hennessy		Learning & Development Facilitator
Mr Olly Jones		General Manager, Education Unit

Associate Professor Kersi Taraporewalla		Program Quality and Development Lead, Education Unit
Ms Stacey Walker		Quality Officer, Education Unit
Dr Jennifer Woods		FANZCA, Staff Anaesthetist
Dr Neroli Chadderton		Supervisor
Associate Professor David Campbell	Australian College of Rural and Remote Medicine	Censor in Chief
Mrs Karen Connaughton		Assessment Manager
Professor Tarun Sen-Gupta		Director of Medical Education, JCU (on behalf of ACRRM)
Ms Charlotte Denniston	College of Intensive Care Medicine of Australia and New Zealand	Education Advisor
Mr Phil Hart		Chief Executive Officer
Dr Ian Graham	Royal Australasian College of Medical Administrators	Fellow; CEP Coordinator
Dr Lynette Lee		Dean of Education
Ms Anna Lyubomirsky		National Education and Training Program Manger
Associate Professor Pooshan Navathe		Chair, Education and Training Committee
Ms Valerie Ramsperger		Manager, Training Operations
Ms Melanie Saba		Chief Executive Officer
Associate Professor Alan Sandford		Censor in Chief
Mrs Genevieve Foster	Royal Australasian College of Physicians	Senior Executive Officer
Associate Professor Nicola Spurrier		Member of RACP Assessment Committee
Mr Adrian Anthony	Royal Australasian College of Surgeons	Chair, Board of Surgical Education & Training
Mr John Batten		President
Ms Kathleen Hickey		Director, Education Development & Assessment
Mrs Zaita Oldfield		Manager, Education Development & Research

Associate Professor Stephen Tobin		Dean of Education
Dr Ruth Ferraro	Royal Australian and New Zealand College of Ophthalmologists	Deputy Chief Executive Officer & Head of Education
Dr Margaret Aimer	Royal Australian and New Zealand College of Psychiatrists	Board Director, Education
Dr Anita Bhatt		Manager, Assessments
Ms Elaine Halley		Executive Manager, Education & Training
Dr Ronald McCoy	Royal Australian College of General Practitioners	Education Strategy Senior Advisor
Ms Leah Bloomfield	Royal College of Pathologists of Australasia	Curriculum & Assessment Development Officer
Associate Professor Margot Lehman	Royal Australian and New Zealand College of Radiologists	A/Director Dept of Radiation Oncology, Division of Cancer Services, Princess Alexandra Hospital
Mrs Pamela Spoors		Head of Specialty Training
Dr Alex Tan		Senior Radiation Oncologist, Townsville Cancer Centre
Dr Alexandria Taylor		Director of Training, RMH
Dr Meredith Thomas		Deputy Chief Censor
Associate Professor Dinesh Varma		Acting Program Director, Radiology and Nuclear Medicine, The Alfred Health and Monash University

Participant	Medical School	Role
Dr Brendan Condon	Deakin University	Year 4 Coordinator
Dr Karen D'Souza		Senior Lecturer in Medical Education
Professor Colin Bell		Professor of Public Health
Ms Mary Lawson		Senior Lecturer in Medical Education
Dr Dominique Martin		Senior Lecturer In Health Ethics And Professionalism
Dr Janet McLeod		Course Director

Dr Anita Phillips		Deputy Director of Clinical Studies
Mr Stephen McManis	Flinders University	Incoming president, Flinders Medical Students' Society
Dr Lisa Kruck	Griffith University	Academic Lead for DLEPP / Senior Lecturer in Medical Education
Professor Gary Rogers		Deputy Head of School, Learning & Teaching Program Director,
Professor Ray Tedman		Director of Medical Studies
Dr Mary Dalton	University of Newcastle	Lecturer
Dr Amanda Dawson		Senior lecturer in Surgery, Clinical Dean Central Coast Clinical School, Co-chair phase 2 Curriculum Development Committee
Dr Conor Gilligan		Senior Lecturer, Phase 1 Lead
Professor Brian Jolly		Director, MEU
Associate Professor Lisa Lampe		Year 2 Coordinator, JMP BMedSci/MD program
Dr Ben Walker		Senior Lecturer, Curriculum Design & Implementation
Associate Professor Frank Bate	University of Notre Dame Australia	Director, Medical Education, School of Medicine Fremantle
Ms Hayley Harris	Macquarie University	Program Manager, Education & Faculty Initiatives
Dr Mark Lee		Unit Convenor
Professor Joanne Lind		Associate Dean, Learning & Teaching
Dr Claire Harrison	Monash University	Curriculum and Assessment Lead, General Practice
Dr Julia Harrison		Clinical Skills Lead
Dr Diane Kelly		Adjunct Senior Lecturer, CICM Primary Examiner

Ms Jennifer Lindley		Senior Academic Lead (Medicine course curriculum)
Dr Ian Presnell		Senior Lecturer / Psychiatrist
Dr Narelle Mackay	University of Queensland	Clinical Assessment Lead
Dr Helen Wozniak		Academic Lead Assessment
Professor Jane Bleasel	University of Sydney	Co-Director, Sydney Medical Program
Associate Professor Deborah O'Mara		Assessment Lead
Ms Catherine Zhao		Associate Lecturer (Assessment & Evaluation)
Dr Dan Dumbrell	Medical Deans Australia and New Zealand	Project & Research Officer
Ms Carmel Tebbutt		Chief Executive Officer

Participant	Intern Accreditation Authority	Role
Associate Professor Katrina Anderson	Canberra Region Medical Education Council	Chair
Dr Claire Blizard	Health Education and Training Institute	Medical Director
Ms Marilyn Bullen	Postgraduate Medical Council of Victoria	Education Manager
Ms Carol Jordon		Chief Executive Officer
Ms Carmen Crawford	South Australian Medical Education and Training	Senior Project Officer, Education and Online Services
Associate Professor Alison Jones		Director – Medical Education and Research

Participant	Professional Association	Role
Mrs Jodie Atkin	Australian Orthopaedic Association	AOA 21 Project Team Leader
Dr Ian Incoll		Dean

Participant	GP Training Provider	Role
Associate Professor Jill Benson	ModMed	Medical Education
Ms Stephanie Clota		Executive Officer
Dr Nyoli Valentine		Lead Medical Educator,

Participant	Regulator	Role
Associate Professor Stephen Adelstein	Medical Board of Australia	Practitioner member
Professor Anne Tonkin		Practitioner member
Ms Michelle Wright		Community member
Ms Marina Fidanza	Australian Health Practitioner Regulation Agency	Policy & Project Manager, Medical
Ms Sarah Harper		Policy Manager, Medical
Dr Jo Katsoris		Executive Officer, Medical

AMC staff	Role
Mr Ian Frank	Chief Executive Officer
Ms Theanne Walters	Deputy Chief Executive Officer
Ms Chrissy Arnaoutis	Program Administrator
Ms Susan Buick	Program Director, Development & Quality Assurance
Dr Julie Gustavs	Manager, Strategy and Specialist Training Support
Ms Katie Khan	Accreditation Assistant
Mr Carl Matheson	Director, Assessment & Innovation
Ms Jane Porter	Manager, Specialist Training and Program Assessment
Ms Karen Rocca	Accreditation Policy Officer
Ms Sarah Vaughan	Manager, Prevocational Standards Accreditation

Summary of discussion questions

Programmatic Assessment and the Medical Continuum		
Presentation: Professor Cees van der Vleuten *Programmatic Assessment Fundamentals – history of assessment from psychometrics to progression and subjectivity.* Q & A	Page 22	*~30 minutes*
Where are we with Implementation of Programmatic Assessment Concepts?		
Small groups		*~20 minutes*
1. Consider the key concepts underpinning programmatic assessment 2. Discuss your current state and proposed plans for assessment in your training program 3. Consider where you are at with the design and implementation of programmatic concepts in your training program. 4. Record the key points of your group on the posters provided. Refer to pages 25-28 of this workbook for further information about the key criteria.		
Plenary		*~20 minutes*
5. On a scale of 1-10 (where 1 is not on the radar and 10 is fully implemented and accepted), where is your training provider in adoption of programmatic concepts of assessment?		
Programmatic Assessment – what problems are we solving?		
Good Ideas and Burning Questions	Pages 25-28	*~30 minutes*
Small groups		*~15 minutes*
1. What are your burning questions related to a key theme in assessment? 2. What are your good ideas related to a key theme in assessment?		
Plenary		*~15 minutes*

3. Review the burning questions and good ideas that other groups have created. 4. Using your red dots – select those questions, which you are most keen to hear the experts answer in the next session.		
Plenary		*~40 minutes*
5. Review the burning questions and good ideas that other groups have created.		
Presentation: Professor Lambert Schuwirth *What Problems does programmatic assessment solve?* Q & A	Page 30	~30 minutes

Focus on Implementation	
Case Studies on Implementation of Programmatic Assessment	*~80 minutes* *This includes 3 X 20 minute case study vignettes.*
Case study participants reflect on: 1. What are some smart design principles in this case study? 2. How closely does this model and assessment in the case study align with AMC standards? 3. What are some implementation strategies you can use in your own training program?	

Where to Next for Assessment?		
Panel presentation	Page 33	*~30 minutes*
Panel members to give a brief update on their thoughts about: 1. Next Steps in Medical education Assessment for health; and 2. Their ideas about international collaborations about assessment, sharing of best practice and AMC assessment standards review.		

Presenters

Professor Cees van der Vleuten

Scientific Director of the School of Health Professions Education

University of Maastricht, Netherlands.

Cees van der Vleuten, PhD, has been at the University of Maastricht in The Netherlands since 1982. In1996 he was appointed Professor of Education and chair (until 2014) of the Department of Educational Development and Research in the Faculty of Health, Medicine and Life Sciences. Since 2005 he has been the Scientific Director of the School of Health Professions Education. This graduate school offers master and PhD degrees in health sciences education to a wide variety of international students. He mentors many researchers in medical education and has supervised more than 80 doctoral graduate students. His primary expertise lies in evaluation and assessment. He has published widely in this domain, holds numerous academic awards, including several career awards. In 2005, he received John P. Hubbard Award for significant contribution to research and development of assessment of medical competence from the National Board of Medical Examiners in the US. In 2010, he received a Dutch royal decoration for the societal impact of his work and in 2012 the Karolinska Prize for Research in Medical Education. He serves frequently as a consultant internationally. He holds honorary academic appointments in the Department of Surgery and Internal Medicine, University of Copenhagen in Denmark, Department of General Practice, Radboud University Nijmegen in The Netherlands, School of Medicine, Flinders University, Adelaide in Australia, University of the Witwatersrand, Johannesburg in South Africa and the Uniformed Services University of the Health Sciences in the US. A full CV can be found at: www.ceesvandervleuten.com

Professor Lambert Schuwirth

Strategic Professor for Medical Education and Director of the Flinders University Prideaux Centre for Research in Health Professions Education,

Flinders University, South Australia.

Lambert Schuwirth obtained his MD from Maastricht University. In 1991, he joined the Department of Educational Development and Research there, taking up various roles in student assessment: Chairman of the Inter-university and the Local Progress Test Review Committee, the OSCE Review Committee and the Case-based Testing Committee. Since the early 2000s, he has been Chair of the overall Taskforce on Assessment. He has been advisor on assessment to medical colleges in the Netherlands and the UK. In 2010, he chaired an international consensus group on educational research, the results of which were published in *Medical Teacher*. Since 2007, he has been a full-professor for Innovative Assessment at Maastricht University – currently as Adjunct. In 2011, he was made a Strategic Professor for Medical Education at Flinders University in Adelaide, Australia and is the Director of the Flinders University Prideaux Centre for Research in Health Professions Education.

Setting the Scene

Assessment Standards at the AMC

Assessment is one of the areas of focus in the Prevocational, Primary medical program and Specialist medical program accreditation standards.
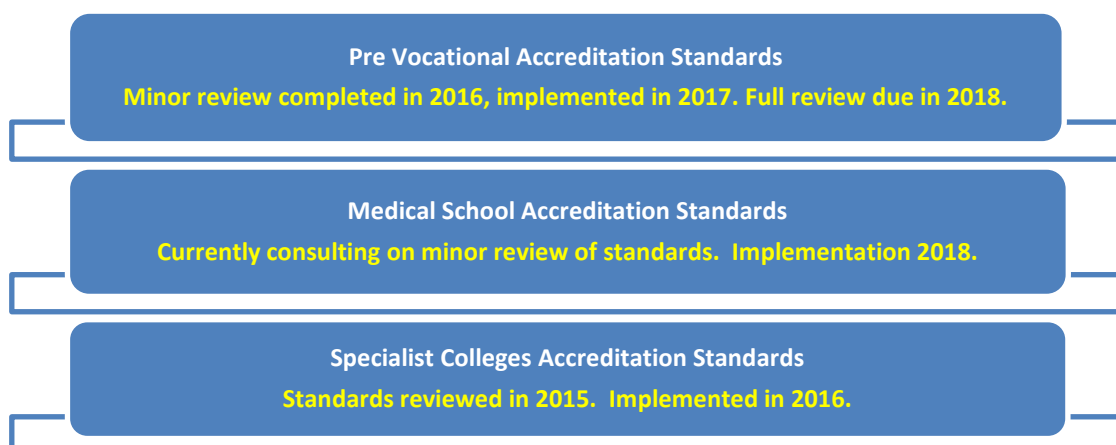
Key Concepts

The key concepts underpinning AMC standards on assessment for medical programs across the continuum are:

- **Assessment approach**

  The assessment program is aligned with learning outcomes, with requirements clearly documented and easily accessible to staff, supervisors and students/trainees/interns.

- **Assessment methods**

  The program contains methods that are fit for purpose, has a blueprint to guide assessment through each stage and uses validated methods of standard setting.

- **Assessment feedback**

  The provider/program facilitates regular feedback to students/trainees/interns to guide their learning, gives feedback to supervisors on assessment performance and has processes for underperforming students/trainees/interns and implementing remediation.

- **Assessment quality**

  The provider regularly reviews its program of assessment to ensure the validity and reliability and scope of its practices, processes and standards is consistent across teaching sites.

**Timeframes for Review of AMC Accreditation Standards**

AMC standards are reviewed on a five-year cycle – status is as follows:

> **Pre Vocational Accreditation Standards**
> Minor review completed in 2016, implemented in 2017. Full review due in 2018.

> **Medical School Accreditation Standards**
> Currently consulting on minor review of standards. Implementation 2018.

> **Specialist Colleges Accreditation Standards**
> Standards reviewed in 2015. Implemented in 2016.

Further information can be found at http://www.amc.org.au/accreditation

AMC Assessment Innovation

The AMC has developed expertise in Assessment over a significant period of time. Core responsibilities include the Assessment of International Medical Graduates (IMGs), the creation of a world-class National Test Centre in Melbourne and research and innovation to ensure methods and assessment approaches remain at the forefront of good practice in medical education assessment.

IMG Assessment	National Test Centre	Research and Innovation
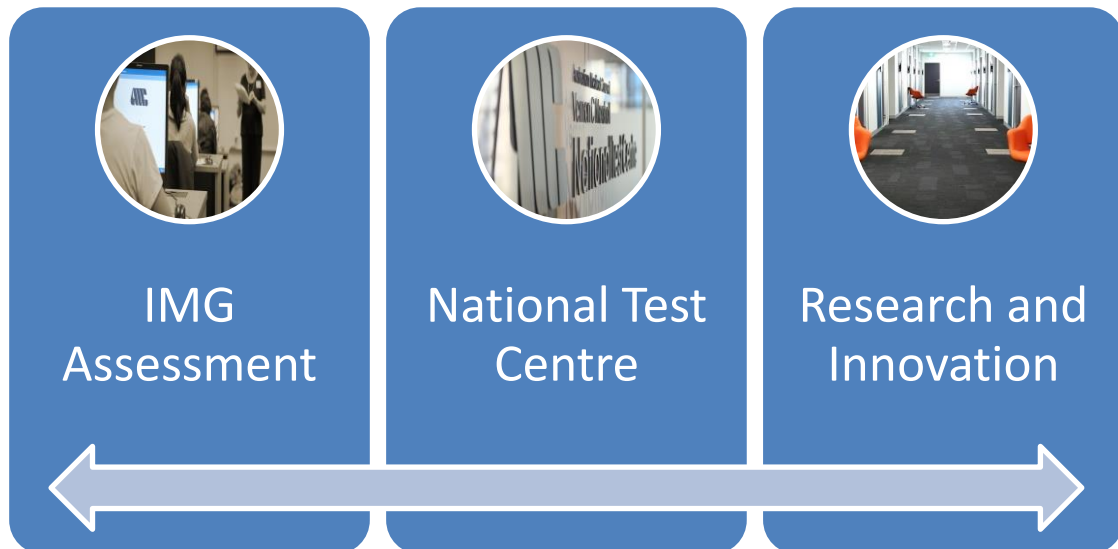
IMG Assessments:

Since 1986, the Australian Medical Council has been responsible for IMG examinations in Australia. Since July 2010, the examination, leading to general registration for IMGs has been conducted under the provision of the Health Practitioners Regulation National Law Act 2009. The first is a knowledge test in the form of a best practice computer adaptive test (CAT) examination. The second involves an Objective Structured Clinical Exam (OSCE) type clinical examination conducted at the National Test Centre and other selected venues. A small number of candidates undertake a highly successful Workplace Based Assessment program (WBA) in lieu of the clinical examination.

National Test Centre:

The AMC Vernon C Marshall National Test Centre (NTC) officially opened in 2013, and is a state of the art facility located in Melbourne to undertake the assessment of clinical skills in medicine utilising the latest technology and best practices. The NTC was established with support from the Australian Commonwealth Government (Health Workforce Australia). To learn more about the centre watch the YouTube video: search for AMC National Test Centre.

Research and Innovation:

The AMC is committed to research and innovation to ensure its methods of assessment and key approaches are leading practice. Current research projects focus on technology enabled systems, benchmarking items and expertise and assessment indicators.

Further information can be found at http://www.amc.org.au

At a Glance: Recent Trends in Assessment

Assessment is integral to education programs across the continuum in medical schools, the prevocational years, specialist training and in the assessment of International Medical Graduates (IMGs).

Why is assessment important?

Assessment is the mechanism by which the education provider determines the ability of individual students/trainees to meet specific milestones of the training program and ultimately measures readiness for unsupervised practice. Assessment is also fundamentally a learning process in itself. It has long been recognised that assessment drives learning but increasingly assessment *for* learning is emphasised. Assessment should promote learning.

Why are approaches to assessment changing?

In recent times, the field of medical education assessment has undergone significant change. This change is linked to the adoption of competency-based approaches to medical education whereby supervisors are required to make decisions about the learner's competence across a range of pre-determined standards (Ten Cate O and Scheele F 2007). Supervisors require a large amount of information to support these important decisions about competence and progress. The emphasis on assessment for learning has highlighted the shortcomings of assessments based solely on high stakes examinations. Such examinations do not provide the nuanced information required to have full confidence in the accuracy of assessment decisions, particularly on the assessment of professionalism and actual real world ability (Rethans J, Norcini J, Báron-Maldonado M, et al. 2002; Creuss et al 2006). This has seen an increased emphasis on work based learning and assessment (Norcini J, Blank LL, Arnold GK, et al. 1995; Govaerts MJB, Van der Vleuten CPM, Schuwirth LWT, et al. 2007). It also features new thinking about how multiple data points from formal exams and regular work-based low stakes assessments can be synthesised as a program of assessment to make progression and high stakes decisions on performance and work readiness (Van der Vleuten CPM, Schuwirth LWT. 2005; Van der Vleuten CPM, Schuwirth LWT, Driessen EW, et al. 2012).

What is a program of assessment?

A program of assessment is the planned and deliberate use of assessment rather than the arbitrary selection of tools and content of assessment. The planning of assessments includes selection of a variety of assessment methods that sample as many situations as possible. A program of assessment ensures that supervisors have clear guidelines and a framework to use as a reference point to guide their individual assessment decisions - therefore improving consistency across settings (Van der Vleuten CPM, Schuwirth LWT, Driessen EW, et al. 2012; Van der Vleuten CPM, Schuwirth LWT, Driessen EW, et al. 2015.)

What is the link between assessment and learning?

Newer thinking about assessment has also focused on the link between assessment and learning (Cilliers FJ, Schuwirth LWT, Adendorff HJ, et al. 2010; Cilliers FJ, Schuwirth LWT, Herman N, et al. 2012.) and feedback (Ericsson KA. 2007; Boud, D and Molloy, E 2012). This acknowledges that assessment is a powerful way to improve performance and this is best achieved through support rather than punitive means. Assessments should ideally provide feedback on a variety of aspects of practice, such as medical knowledge, communication and quality and safety. Assessments should also be undertaken across a broad range of contexts and include different methods such as direct observation, case discussions, and opportunities for reflection. It is through multiple biopsies of a learner's performance and ongoing feedback that a complete and more accurate picture of their level of ability can be formed and learning is consolidated (Schuwirth LWT, Van der Vleuten 2011).

How do we determine the quality of assessments?

The field is also marked by new ways of thinking about how the quality of assessments can be determined. It has seen a shift from purely psychometric concerns of assessment focused on statistical analysis of validity and reliability (Norcini et al 1985) to the use of qualitative measures, which are more aligned to the recognition of the subjective nature of assessment decision making (Hodges, B 2014). Van der Vleuten (1996) for example argues that the utility of an assessment tool is the product of its validity, reliability, educational impact, feasibility and acceptability. Newer thinking about determining the quality of assessment also highlights the question of the role of the learner in assessing their own performance, supervisors and other stakeholders including other health professionals, employers and consumers. Also part of the movement towards more contemporary evidence-based decision making in assessment is the use of technology enabled reporting to assist with the storage and interpretation of assessment data (Moonen-van Loon, J.M.W., Overeem, K., Donkers, H.H.L.M. et al. 2013). A further feature is the need for more transparent benchmarking of assessments across providers (Schuwirth LWT, Van der Vleuten CPM. 2011), and standard setting (Weller JM, Misur M, Nicolson S, et al. 2014; Cook DA, Kuper A, Hatala R, et al. 2016).

How can we get assessment to work well on the ground?

Given the scope of these changes there is also a recognition that improved implementation is paramount to the success of assessment innovation. This includes incorporation of change management strategies include co-design, broad consultation, communication and supervisor and assessor training.

Programmatic Assessment, first proposed by leading medical educators Profs Cees van der Vleuten and Lambert Schuwirth is a useful term, which encapsulates the key concepts, underpinning newer ways of thinking about medical education assessment. For key papers and further reading on this topic, please see references on page 36 of this workbook.

A key component of programmatic assessment is the separation of data from decisions, that is, not all assessment episodes need to be accompanied by a summative decision. Instead, high stakes decisions are made only after a sufficient number of observations of a trainee's performance have been gathered and synthesised.

Programmatic Assessment and the Medical Continuum

~30 minutes

Presentation: Professor Cees van Der Vleuten

*Programmatic Assessment Fundamentals – history of assessment from psychometrics to progression and subjectivity.*

Prof Cees van der Vleuten presenting at the AMC Workshop on Assessment 2017.

Programmatic Assessment and the Medical Continuum –
Prof Cees van der Vleuten

Current state of Assessment:

Medical Education has been dominated by the creation of a wide range of assessment tools. Miller's pyramid has been particularly influential in helping us to make the transition from fact-oriented assessment to focusing assessment on measuring what doctors actually do.
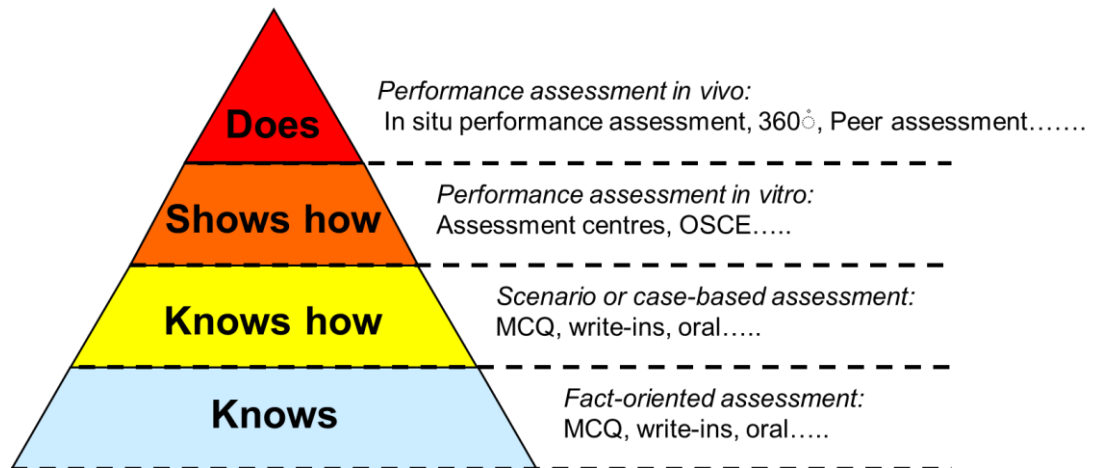
Group discussion at the AMC Workshop on Assessment 2017.

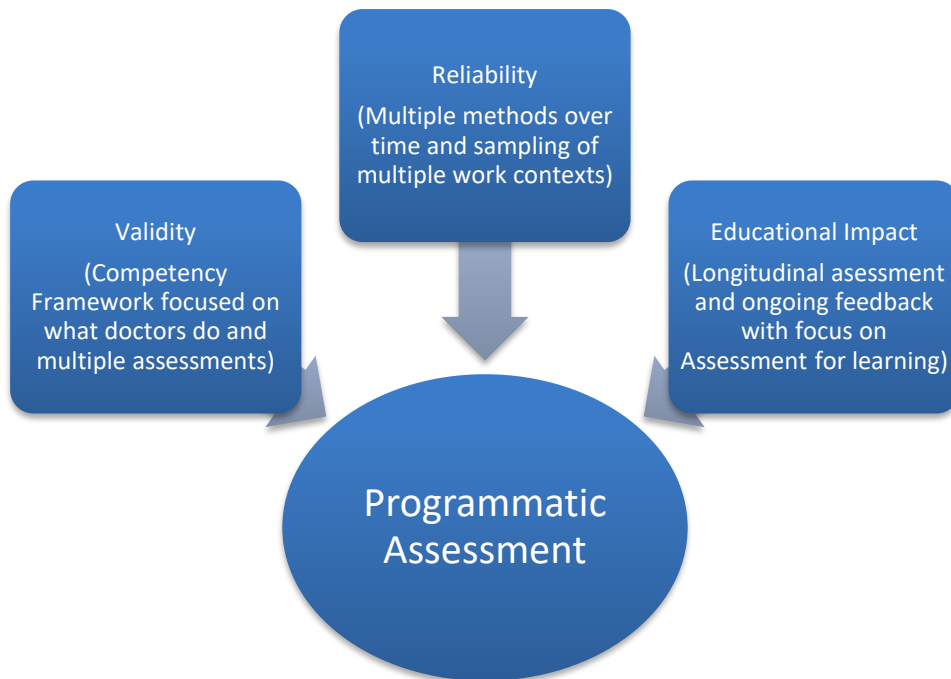**Instruments for determining the quality of Assessments:**

1. **Validity:** Focuses on the question of – *what are we assessing*. The focus of validity in medical education has changed over time:

    o   **From time served to outcomes:** We have seen a transition from input methods (defining hours and time served in training) to outcomes-based criteria. This has seen a transition from haphazard learning, to integrated objectives, to end objectives and now generic competencies. Beyond medicine, this movement has become a feature of school and professional education. Within medicine, there has been a great deal of consensus about what a doctor should do. CanMeds, as defined by the Royal College of Physicians and Surgeons of Canada, has been particularly influential in determining additional contemporary ways of defining the work of doctors. This framework has been developed with a lot of stakeholder engagement and consensus about what a doctor should do. CanMeds are based on the premise that most of the competencies move beyond the knowledge domains. They are complex skills and behaviours. The important observation with the CanMeds framework and which is mirrored in many other competency frameworks is the considerable attention they pay to the skills beyond the mere medical technical ones. Of course, the pure medical technical skills and abilities play a central role but so do the 'softer' abilities, such as collaboration, communication, advocacy and professionalism. There is good support for this broader view on medical competence, especially given that 80 percent of consumer complaints can be attributed to communication errors. These competencies develop longitudinally, so if we take this seriously, we need to assess performance or behaviours at the top of Miler's pyramid.

    o   **Teacher oriented to self-directed:** There has been a significant shift in theories and methods of learning, which have had an impact on what and how we assess. This transition has been marked by increased focus on the agency of the learner in determining how and what they need to learn to support them to succeed.

    o   **Messages from validity research:** No single method can do it all. We need a mixture of methods. We need both standardized and non-standardized assessment methods. For standardized assessment, the quality control around test instrument development and administration is vital. In non-standardised assessments, the utility is provided not in the instrument but in the interaction between humans – the supervisor and the learner. The quality of the assessment is provided through the feedback after the encounter – that is the value of the non-standardized assessment. This means that '*we don't need to sharpen the instruments. We need to sharpen the people – they need to understand what they are doing and why'*.

Prof Cees van der Vleuten presenting at the AMC Workshop on Assessment 2017.

2. **Reliability:** Focuses on the question of – *how stable is the assessment decision over-time and encounters.* Research into the reliability of assessment in medical education has shown:

   o **Shorter tests are unreliable:** Any sort of performance is contextually bound. As soon as you change the context, you change the outcome. Given this, in shorter tests we see that there are a lot of false positive outcomes and false negative outcomes (i.e., students who pass, who are in fact not sufficiently competent, or students who failed, who were in fact sufficiently competent) based on the noise of the measurement. *One single measure is no measure.* Research into the method reliability of testing time (Van der Vleuten and Schuwirth 2005) shows that if we assess for long enough (four hours+) we can select any type of assessment, as we will gain similar high quality and reliable scores of 0.8+. Therefore, reliability is not necessarily a function of standardisation, and adding structure to the assessment but much more a function of sampling.

   o **Sampling:** If we wish to assess complex behaviours for which we need to rely on some form of professional judgement, the lessons from reliability research is that many of these judgements can build up a sufficiently robust picture. Sampling allows the assessor to combine information across multiple sources. Research has shown that generally 7-11 observations are enough to obtain a sufficiently reliable score. There is also an effect of aggregation of score across methods, which means that if we combine the results of various methods the samples can be even lower.

3. **Educational Impact:** Focuses on the question – *how does assessment influence learning and performance.* Traditional assessment research has focused much on psychometrics so educational impact has not been of the agenda at the top. The relationship between assessment and learning is complex:

   o **Assessment for learning:** Traditional testing approaches to assessment clearly drove student learning but were unlikely to result in long-term retention. Traditional testing approaches typically feature informal workplace-based assessment, high levels of reductionism in progression decisions leading to a grades culture, poor feedback and tick box assessment practices. The most dominant learning induced by the traditional testing system, is a cram-and-dump style of learning and assessing, whereby a large proportion of what is learnt is not retained longer than two weeks.

   o **Longitudinal assessment:** To counteract such ineffective and inefficient learning, approaches have been developed with longitudinal assessment, whereby there is a continuous model of assessment throughout training, supported by ongoing constructive feedback.

Programmatic Assessment: At a Glance.

Validity (Competency Framework focused on what doctors do and multiple assessments)	Reliability (Multiple methods over time and sampling of multiple work contexts)	Educational Impact (Longitudinal asessment and ongoing feedback with focus on Assessment for learning)

Programmatic Assessment

12 Tips		
	1.	Develop a master plan
	2.	Adopt a robust system for collecting information
	3.	Develop examination regulations that promote feedback orientation
	4.	Assure that every low-stakes assessment provides meaningful feedback for learning
	5.	Provide mentoring for learning
	6.	Ensure trustworthy decision-making
	7.	Organise intermediate decision-making assessments
	8.	Encourage and facilitate personalised remediation
	9.	Monitor and evaluate the learning effect of the program and adapt
	10.	Use the assessment process information for curriculum evaluation
	11.	Promote continuous interaction between stakeholders
	12.	Develop a strategy for implementation.

Discussion

Small groups   ~20 minutes

1. Consider the key concepts underpinning programmatic assessment.
2. Discuss your current state and proposed plans for assessment in your training program.
3. Consider where you are at with the design and implementation of programmatic concepts in your training program.
4. For those participants who are from a regulatory organisation, jurisdiction, or other key stakeholder of medical education, consider your observations of the current state and desired future state of medical education assessment.
5. Record the key points of your group on the posters provided. Refer to pages 25-28 of this workbook for further information about the key criteria.
6. Prepare to report back, including where you are at on the scale (see page 24).

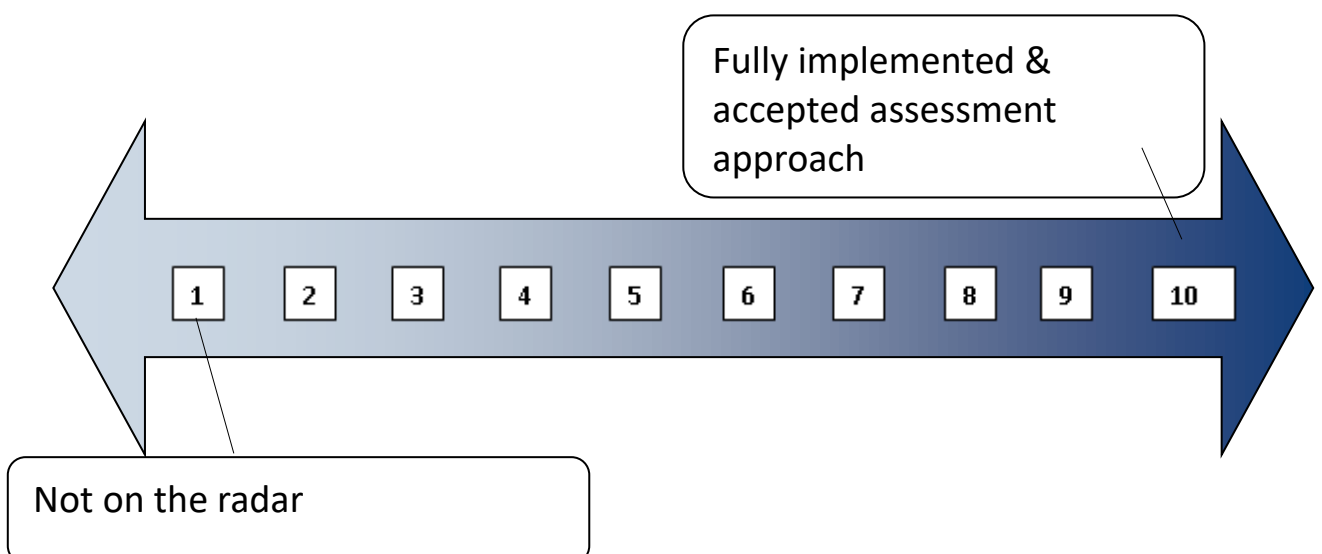| Programmatic Assessment Themes	Current State	Future State
a. **Fit for Purpose:** ensuring that methods and approaches focus on best practice medical education evidence and assess abilities and behaviours relevant to 21st century health needs and systems. b. **Feedback and culture:** improving feedback and learning cultures to ensure learners are empowered and supported. c. **Discrimination of performance:** improving methods of identifying good and poor performance through longitudinal assessment systems. This is a shift from more traditional norm referenced assessments which are more focused on comparing candidates against each other rather than of their achievement of standards and assessment for learning methods to ensure each individual doctor is the best they can be. d. **Benchmarking, technology enabled reporting and research and evaluation:** ensuring equity in decision making and feasibility and quality of innovation of assessments through benchmarking assessment methods across the continuum, conducting research and evaluation and enabling decision making with use of technologies for monitoring progression and quality programs. e. **Implementation:** ensuring high quality, feasibility, sustainability, cost effectiveness, and acceptability of assessment systems. Successful implementation includes: i. robust communication and management of change process ii. effective governance and champion model of ongoing support and leadership iii. operational management to ensure smooth delivery of the assessment iv. supervisor/assessor training to ensure consistency of decision making and quality of feedback v. plans for evaluation and quality control to inform future quality improvements to assessments vi. ongoing resource and stakeholder management to ensure continuity of support.		

7. On a scale of 1-10 (where 1 is not on the radar and 10 is fully implemented and accepted), where are we at in the adoption of programmatic concepts across the medical education continuum?)

Prof Liz Molloy facilitates the plenary discussion on Programmatic Assessment Across the Medical Continuum.

Fully implemented & accepted assessment approach

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

Not on the radar

Plenary Summary of Five Themes of Programmatic Assessment

> **Fit for Purpose**

Current and future state - Where we are at across the medical continuum and where are we going?

1. **Assessment Mapping.** Assessment mapping is being undertaken with clear rubrics to get reliable and valid assessments but we have a long way to go. Currently, assessment focuses on medical knowledge but not adequately on professional domains. Many work-based assessments have been implemented but they are not well aligned with summative assessments. At the program level, assessment modalities are used in isolation, each leading to an individual outcome only to be combined afterwards, which means that there might be a thought through assessment program but not a fully-fledged.

2. **Behavioral Markers.** Excited about potential use of behavioral markers as structure for feedback process and enabled through technology, rather than reliance on paper-based methods.

3. **Self-evaluation.** There is an opportunity to use technology for students to self-evaluate and then receive external cues and discuss.

4. **Sustainability.** It is vital to think through assessment from a sustainability point of view to ensure that it is doable. There is a high clinical and patient workload. In implementing further change to assessment, we need to consider the resourcing and impact on health systems.

5. **Adaptability and alignment.** We need to work on developing meaningful and authentic assessments, which are adaptive; in that, they can be customized to the needs of learners – lighter touch for those that are work ready. Assessments need to be implemented within a framework rather than being an array of discrete tools. GPs are using assessments for multiple purposes, and aligned with different outcomes. Trainees do case write-ups, and then they are asked to produce questions collated into formative assessments. They then identify which questions they find most contentious then they discuss in peer groups and come up with consensus. In this model of assessment, trainees learn, gain feedback, critique own assessment and discuss. This forms a community of practice around assessment and feedback thus through consensus building new information is built into the system based on authentic assessment. This assessment practice is aligned with David Boud's concepts of sustainability of assessment.

> **Feedback and culture**

Current and future state - Where we are at across the medical continuum and where are we going?

1. **Feedback.** Not doing feedback sufficiently well. There is a lack of understanding of what constitutes feedback, and it is of variable quality, tick-box and still too teacher driven. There is still a high proportion of 'failure to fail', which needs to be improved through supervisor training. In creating training for supervisors, we need to focus on developing skills to ask trainees and students what they need to improve, rather

than giving answers or building a culture whereby trainees and students are overly dependent on being told what to do. Such a change of culture and training of supervisors may be investment but it is an investment in efficiency because in the end, any supervisory activity that does not optimally lead to student learning is not a good use of time. Training in providing feedback and/or in teaching leads to increased educational expertise and thus do more efficiency and effectiveness. For example, through questioning there is an opportunity to build a more empowered and self-accountable trainee workforce, which is currently high on the national agenda. We also need to address improvements to feedback at a systemic and cultural level, whereby feedback becomes a part of everyday practice embedded in high and low stakes assessment practices and teaching and learning rather than a summative add-on or after thought.

2. **Questioning.** There are parallels with clinical practice in terms of the effectiveness of clinician communication techniques, such as motivational interviewing with patients and improved patient adherence to medication and trainee and medical student feedback approaches which focus on the supervisor asking questions rather than giving answers.

3. **360-degree feedback.** There is an opportunity to use 360-degree feedback effectively at the beginning and end of training as a diagnostic and evaluation tool. We need to incorporate the feedback of multiple stakeholders including patients, professional health teams and self-evaluation.

Progression and discrimination of performance

Some issues:

Current and future state - Where we are at across the medical continuum and where are we going?

1. **Criterion Referenced.** Need to further develop criterion-referenced assessments and adopt a more longitudinal approach to assessment, whereby the performance of the individually is tracked and monitored over time with multiple biopsies of evidence of learning.

2. **Pass/Fail Grades.** Students don't see the relevance of grades beyond Pass/Fail because it does not necessarily relate to how good they are as a doctor or how much they know. For this, narratives are more important than grades, because narratives contain more information/feedback for students as to how to improve, which grades don't.

3. **Work-based Assessment.** Work-based assessments are valuable. There still needs to be work done to ensure that we are using assessments appropriately to determine high stakes decisions.

4. **Feedback.** More important to focus on providing strong qualitative comments and feedback to strengthen assessment for learning.

Benchmarking, technology enabled reporting and research and evaluation

Some issues:

Current and future state - Where we are at across the medical continuum and where are we going?

1. **Reporting on qualitative data.** Technology enables reporting on qualitative data. In the future, Artificial Intelligence will most likely further support decision making.

2. **Behavioral Markers.** Excited about potential use of behavioral markers as structure for feedback process and enabled through technology rather than reliance on paper-based methods.

**Implementation**

Some issues:

Current and future state - Where we are at across the medical continuum and where are we going?

1. **Change management.** Robust communication and management of change process to build a collaborative and non-hierarchical training workforce and breakdown silos.

2. **Governance and leadership support.** Effective governance and champion model of ongoing support and leadership.

3. **Operational management.** Operational management to ensure smooth delivery of the assessment.

Programmatic Assessment – What Problems are we solving?

Burning questions and good ideas

In this section of the workshop, we added to our list of burning questions and good ideas to work towards avoiding some of the pitfalls of assessment in training programs, and more broadly impact positively on the growing of a quality medical workforce.
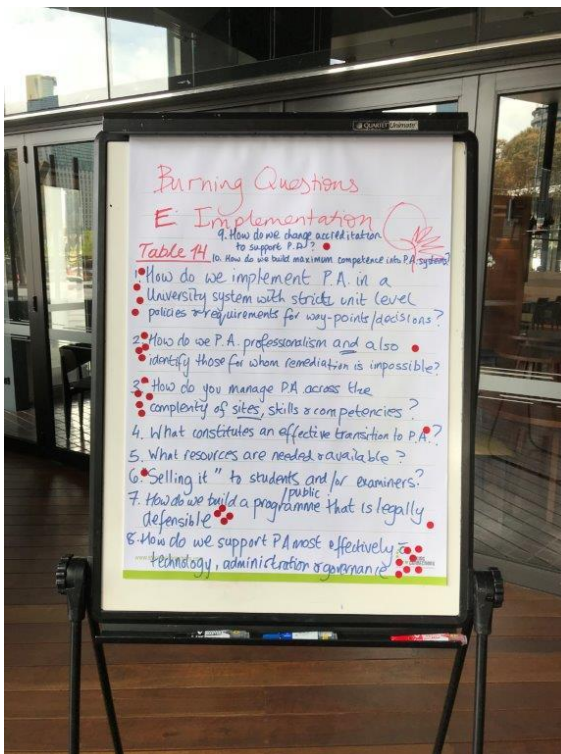
The following list of burning questions has been themed as a first step to establishing common issues and ideas about quality improvement to assessment.

In the New Year with our experts and with your support and expertise, we will seek to answer these questions and will circulate these as a common set of FAQs.

Discussion

*Small groups*   ~15 minutes

1.   What are your burning questions related to a key theme in assessment?
2.   What are your good ideas related to a key theme in assessment?

Plenary   ~15 minutes

3.   Review the burning questions and good ideas that other groups have created.
4.   Using your red dots – select those questions, which you are most keen to hear the experts answer in the next session.
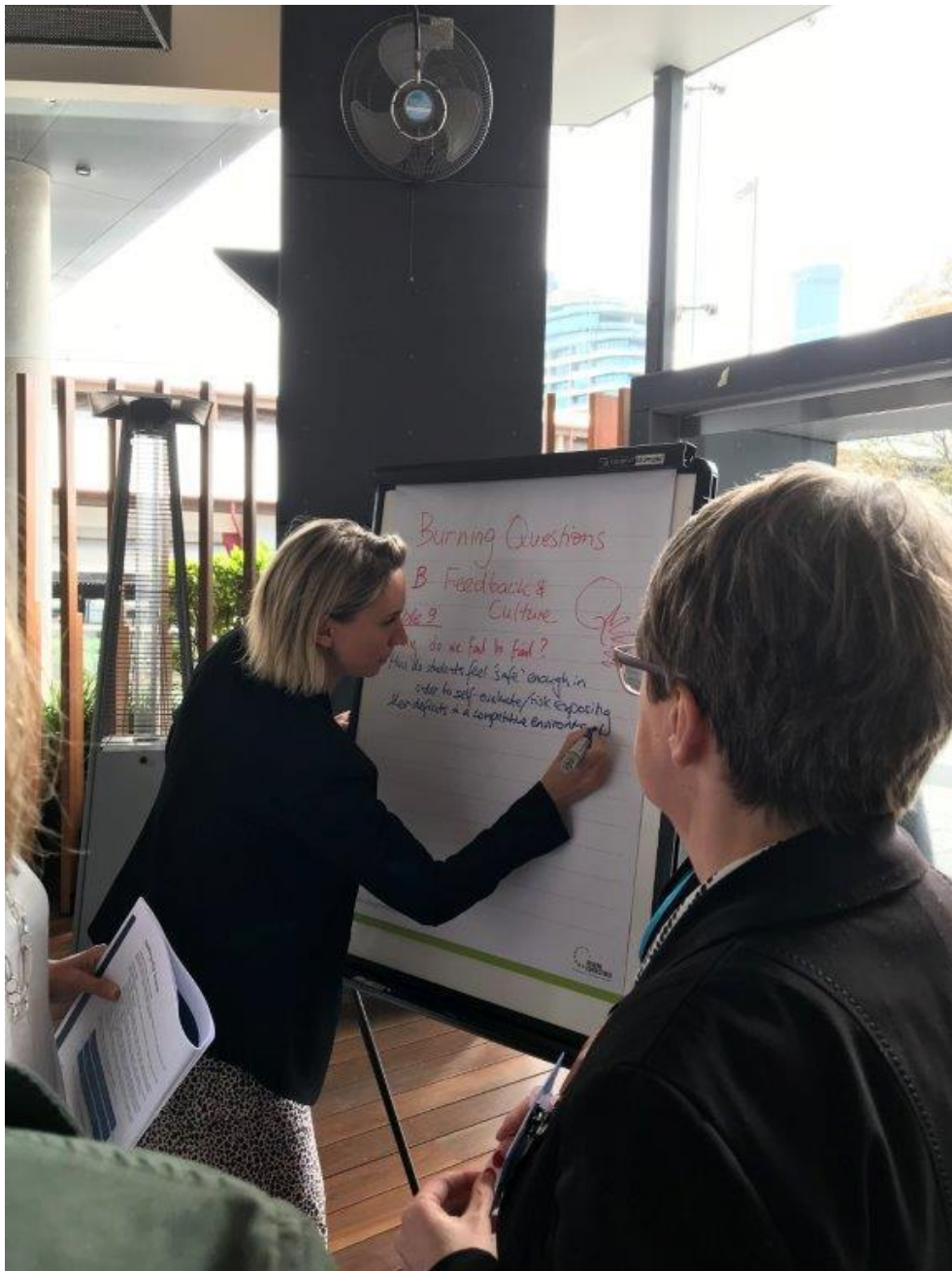
Prof Liz Molloy scribes some burning questions on feedback and culture.

Groups brainstorm burning questions for the experts to answer.

Fit for Purpose

Some issues:

- **Patient safety and appropriate assessment load** ensure assessments do not compromise patient safety or pose a risk to the wellbeing of doctors
- **Assessing important health priorities** ensure that the assessment aligns with assessment of health priorities including Indigenous Health and Cultural Competence, Professionalism, Patient-centred care and Inter-professionalism
- **Curriculum design and framework** reflective of best practice in terms of patient-centred efficient care, not just physiology and pathophysiology (i.e. how to be a good doctor)
- **Coverage of the curriculum** ensure that the assessment aligns with the curriculum and thus provides transparency about what needs to be learnt
- **Assessment of professional behaviour** emphasise the importance of aspects of the profession that has traditionally been part of the "hidden curriculum" of medical education, such as self-care and ethics
- **Assessment for learning** achieve a balance between detail and high level concepts and drives good practice in lifelong learning i.e. should not be focused on minutiae or esoteric "factual recall" – "facts" that will be revised / changed in a year or two

- **Health system and community needs** ensure assessment aligns with community and health system needs

Prof Cees van der Vleuten answering a burning question – Session chaired by Prof David Prideaux.

Burning Questions: Fit For Purpose			
1.	**Purpose**	How do you define 'the purpose' at the beginning of the continuum, e.g. medical school?	ooooo
2.	**Curriculum Alignment**	How to ensure alignment of assessment types at curriculum?	ooooo
3.	**Longitudinal Assessment**	How can each individual data point be different?	
4.	**Expertise**	How do we clone Cees?	
5.	**Change Management**	How do you raise awareness of need for change? Alternatively, the need to think about it?	oooooo
6.	**Large-scale Change**	How to implement with huge numbers?	ooo

Good Ideas: Fit For Purpose	
1.	**Portfolio:** specified contexts.
2.	**Tools:** custom designed tools: alignment with curriculum objectives.

Some issues:

- **Lifelong learning** design the assessment program to avoid engendering a tick-box approach to meeting training requirements
- **Agency and learner empowerment** enabling students to focus on learning and develop skills. The focus of summative assessment needs to be on how learners demonstrate their ability to identify the specific needs for their learning through awareness of their own skills, level of performance, current gaps and future learning goals and to formulate learning goals, which they then subsequently make happen.
- **Competition** assessment systems may drive a hyper-competitive environment between learners and impact the support they provide to each other as peers. We need to ensure that the assessment system supports peer learning. The fundamental principle is that the assessment is not trying to tell whether one student is better than another but for each individual student whether they are optimally better today than they were yesterday stop
- **Wellness** assessment may impact the health and wellbeing of doctors. We need to ensure that assessment load is appropriate, non-punitive and free from bullying and harassment. This doesn't mean that assessment needs to be soft, warm and fuzzy - learning often requires you to be pushed out of your comfort zone – but it will have to be respectful and meaningful.
- **Burden of assessment** the assessment load may be too high – causing too much stress and may be too high stakes without an appropriate balance with lower stakes assessment. Therefore, it is important that assessment moments and decision moments disentangled, so there will be many assessment moments but only few decision moments. Each assessment moment is namely a data point which contributes to a certain decision moment
- **Remediation** lack of remediation and negative effects on self-worth and treatment by the broader cohort and culture
- **Feedback** lack of feedback on exams and performance assessments
- **Gifted and talented** limited guidance to high performing candidates
- **Interprofessionalism** the avoidance of siloing of professions
- **Quality standards** avoid giving the idea that changes to the assessments will result in "dumbing down" the expectations of trainees. If we are embarking on multiple points or work-based assessments and reducing emphasis on high stakes examinations there is a danger that this will be seen as reducing standards.

Burning Questions: Feedback and Culture			
1.	**Shared Purpose**	How to create culture where supervisors and registrars are on a learning journey together?	OOOOOOOO
2.	**Learning culture**	How do students feel 'safe' enough in order to self-evaluate/risk exposing their deficits in a competitive environment?	OOOOOOOOOO
		Why are we still holding on to feedback information (prevents 'feedforward')? (Counter to PA.)	OO
		How can we prevent burnout of educators/feedback providers?	OOOOOOOOOOOO
		Why do we fail to fail?	OOOOOOOOOOO
		How do we tailor feedback to individuals in a timely way (to enable them us use feedback)?	OO
3.	**Self-directed learning**	How do we get learners to seek feedback?	OOOOOO
4.	**Large-scale change**	Large cohort: a. How do we provide meaningful feedback to all? b. How do we gather, collate and give feedback on multiple points – assessment/supervisors/peers/administrative staff? c. Any IT solutions? d. How big is too big?	OOOOOOOOOOOOOOO
5.	**Change Management**	What is our risk appetite / can we change?	OOO
		Busy clinical practice – how do we build culture over pure service provision? How do we make teaching/education/feedback a priority?	
		How do we turn the system on its head (? Our risk culture)	OOOO

Good Ideas: Feedback and Culture	
1.	**Feedback:** Ask students what feedback they want; Provide context – purpose of assessment and feedback; patient engagement in feedback; train supervisors to give feedback; specific actionable feedback.
2.	**Student Engagement:** Student engagement in assessment processes.
3.	**Transparency:** Clear transparent guidelines, processes and decision making concerning assessment.
4.	**Appropriate tone:** Non-judgmental.
5.	**Mentors:** Stronger, formalized mentorship.

Progression and discrimination of performance

Some issues:

- **Journey mapping** provides a high level visual of the key milestones of a training program drawn from the perspective of what the trainee/student needs to learn and how they will be assessed. It shows how a trainee will progress through the stages of training from start to conclusion of the program. For the assessment component of the journey map, it would provide clear guidance about:
  - o how assessments look in a practical sense including the specific assessment tools to be used
  - o the location of decision-making points for aggregated assessment
  - o specific progression decisions and the teaching and learning support students/trainees can access to support their progression including strategies for dealing with the student in difficulty and in need of remediation.

- **Professional behaviour** need to ensure that the assessment of professional behaviours is sufficient to *enable* or *disable* progression through training. There are two key approaches in the literature concerning assessment of professional behaviour in medical education; one – European and Canadian approach – focuses much more on professional behaviour as an observable characteristic which than can be evaluated and judged. The other – more the US approach – focuses more on professionalism as a person personality trait and therefore has to be measured.

- **Competence and time** what happens to time-based requirements? How long will training be – can it be shorter? Can students/trainees progress earlier than annually?

- **Core and non–core** Can we remove 'core' and 'non-core' references in the competency-based program?

- **Required evidence** how will we assess EPAs? – What level of evidence is required? Should all EPAs be assessed each year? Who should assess each EPA and "sign" a trainee off at the end of their training?

- **High stakes decisions** need to be made by combining "multi-point" assessment tasks, as well as examinations

- **Excellence** should we be assessing and promoting *excellence* as well as *competence?* If so, do we need to establish criteria for the assessment of excellence?

- **Failure to fail** how can we design the assessment program to avoid some of the "failure to fail" pitfalls.

Burning Questions: Discrimination of Performance			
1.	**Culture of Excellence**	Systems value for teaching training feedback.	o
		Motivation beyond "passing" or "acceptable": a. ASPIRATIONS beyond passing. **b.** CULTURE beyond passing.	oooooooooooooooooo
2.	**Competency-based Medical Education**	(Determination of Performance Standard)	ooo
		What is the best method of determining competence?	0000
		Focus on individual against standard.	0000
3.	**Levels and stages**	What is the definition of "performance" i.e. levels, stages?	oooo
4.	**Decision Making**	What are we trying to "discriminate"?	
5.	**Longitudinal Assessment**	What is the experience of standard tolerance and longitudinal assessment?	
		How do we record and manage the performance data?	
6.	**Assessment Scales**	Are scale ratings a good thing?	
		Is everything aligned to move from pass/fail?	
7.	**Change Management**	Standardisation - How to get assessors across training sites to have agreed standards/shared competencies?	ooooooooooooo
		Engagement - How do we get clinicians and hospital administration to take assessment seriously?	oooooooooo

Good Ideas: Discrimination of Performance	
1.	**Feedback:** both ways – trainee to organization and organization to trainee; regular face-to-face meetings of organisation and educators with supervisors gathering multiple data points to make global assessment of need for remediation or input.

> **Benchmarking, technology enabled reporting and evaluation/research**

Some issues:

- **Technology systems** avoid clunky design of online systems, which do not record or represent assessment information easily. Technology Systems need to support the experts and not to replace the expert. Systems that are clunky will have a negative effect on the reputation of the educational change.

- **Decision Analytics** record assessment information in ways that supports decision making and reporting on performance and work readiness

- **Sustainability of innovation** share assessment items and consider creation of a pool of assessment items for use across Medical Education Providers

- **Evaluation plans** for evaluation and quality control to inform future quality improvements to assessments

- **Research on health impacts** and innovation analysis of performance across the system – indicator of candidate comparability and limited writing up of local innovations in assessment

- **Evidence-based design** reference to evidence on assessment

Burning Questions: Benchmarking, technology enabled reporting and research and evaluation			
1.	Focus of Assessment	Improving standards of all vs detecting 'bad apples'.	
2.	Standard setting	Who sets 'the bar' with programmatic assessment – the benchmark may be fluid and depend on perspective/culture/societal norms	ooo
3.	Collaboration across the continuum	Are there efficiencies across colleges and universities for benchmarking, reporting and research (avoid reinventing the wheel!).	oooo
1.	Comparative research	Are there any future predictive studies using traditional assessment methods vs programmatic methods.	ooooo
2.	Reporting	Streamlining - How to integrate all data into a single reporting form? Qualitative - How can you effectively and efficiently report on qualitative data?	o ooooooooooooo
3.	Implementation	How do we overcome practical constraints in health services to embrace programmatic assessment?	ooooo

Good Ideas: Benchmarking, technology enabled reporting and research and evaluation	
1.	**Improvements**: Agreement to standardised measures to enable quality improvements (not league tables).
2.	**Data Systems:** capture assessment information, quantitative and qualitative data to facilitate learning at individual and organization level – mindful of privacy considerations.

Implementation

Some issues:

- **Change management** robust communication and management of change process
- **Governance and leadership** support effective governance and champion model of ongoing support and leadership
- **Operational management** to ensure smooth delivery of the assessment
- **Supervisor/assessor training** supervisor/assessor training to ensure consistency of decision making and quality of feedback and How do we convince supervisors that the "new and different" assessment tools are not enormously time consuming and demanding, but can be incorporated into day to day work?
- **Stakeholder management** ongoing resource and stakeholder management to ensure continuity of support

		Burning Questions: Implementation	
1.	**Change Management**	Strategic alignment - How do we implement programmatic assessment (PA) in a university system with strict unit level policies and requirements for way-points/decisions?	oooo
		What constitutes an effective transition to PA?	o
		How do we "Sell it" to students/public and/or examiners?	oo
		How do we build maximum competence into PA systems?	
		How do you educate students about assessment and reflections and action plans?	ooooooo
		How can PA be matched with a body focused on licensing?	
2.	**Managing complexity**	How do you manage PA across the complexity of <u>sites</u>, skills and competencies?	ooooo
3.	**Resourcing**	What resources are needed and available?	
5.	**Governance and Technology**	How do we support PA most effectively with technology, administration and governance?	ooooooo
		How do you get a platform (e-portfolio) for collection of information (quality and quantity). What do you do with the information?	O
6.	**Competency**	How do we assess professionalism in PA <u>and</u> identify those for whom remediation is impossible?	
8.	**Legal requirements**	How do we build a programme that is legally defensible?	ooooo

9.	Assessor training	How do you train assessors (100's) over multiple sites?	ooooooooooo
10.	Accreditation	How do we change accreditation to support PA?	-
11.	Evaluation	How do you evaluate the PA?	oooooooooo

Good Ideas: Implementation	
1.	**Assessors**: Funded training of assessors (theory/evidence based, practical); protected time for assessors;
2.	**Feedback**: Training students/trainees to receive feedback.
3.	**Training**: Better use of resources – shared training; Strong support from the training administration body.
4.	**Outcomes**: Longitudinal study of final outcome: medical school – practice – co-operation between universities – health service – specialist colleges.
5.	**Quality**: quality leadership driving quality training and assessment.

Presentation: Professor Lambert Schuwirth

~30 minutes

What problems does programmatic assessment solve?

Prof Lambert Schuwirth presenting – What Problems are we Solving?

Programmatic Assessment – What Problems are we solving?
Prof Lambert Schuwirth

Key Messages:

1. **Social Issues:**

    - **Educational World is changing dramatically.** There are no longer discrete phases where you know everything you need – there are constant transitionary phases where lifelong learning is paramount and embedded in everyday practice. Artificial Intelligence will have a huge impact on all professions including medicine. Technology is already highly influencing what and how we learn.

    - **Definition of what is a safe, independent practitioner is changing.** Communication and collaboration is more important than in the past. Assessment in Medicine previously focused on technical elements of the discipline. This has shifted to a firmer grounding on assessing medicine as a humanistic discipline. Medical education has put professionalism high on the agenda because there are most concerning societal health issues, which require different assessments.

    - **Agency and empowerment of the learner.** Wellbeing and burnout is an issue. We need the medical learner to be empowered and to search for feedback and manage their learning.

    - **Reward and punishment.** The considerable body of research – for example the interesting study by Gneezy et al. 2004 – demonstrates what is likely to happen when we change the rules and institute a monetary (punishment) in childcare pick-up has salient lessons for medical education. This research shows that more parents violate the pick-up rules following a punishment rather than previously when it was less about fee for service and more about the social contract with peers and the childcare workers – not wanting to let them down. This study highlights the importance of thinking through the unintended and intended consequences of the impact of assessment design on learners and assessors behaviour. What is a stick for some is a carrot for others. This study challenges the assumption that there is a simple relationship between punishment and reward. If we think of formal summative exams – it drives cram and dump behaviour where research has shown there is close to no retention of knowledge gained post exams. Feedback for learning provides tangible rewards and builds an ongoing culture of learning and ongoing dialogue about performance in the workplace.

    - **Government Review.** Breach of social contract. The Government has a deep scepticism of the professions. To some extent it has grown impatient in the ability of the professions to pull together to minimize risk, build efficiencies and innovate in the health system.

2. Technical Issues

- **Reductionism.** Why do we reduce everything? If you look at the normal scoring process, even in a simple multiple-choice test, we reduce the information all the time. At the outset there is still information about what particular mistakes students made when they answered questions, but then each item is scored and that qualitative information about what mistakes were made is reduced to which items were answered incorrectly. This is further reduced by calculating the total score, because now we only know how many items were answered incorrectly. In the end, this is reduced to a pass-fail decision, which leaves a binary result, while at the start there was information rich data. So in fact all the thinking about statistics and psychometrics is about reducing – how to throw away information in the most sensible way.

  o Often reductionism is arbitrary – which makes it less valid.

  o Reduction without reason is meaningless.

  o Without meaning, there is no feedback.

  o Without feedback there is inefficient learning.

- **Our remit is not education.** If not giving feedback then doing nothing to change the prior probability. Even if only regulation – feedback will help efficiency in re-sits. If you see an examination as a diagnostic test for the disease 'dyscompetence' than the prior probability, - the 'prevalence' of the 'disease'– determines the positive and negative predictive values. Even now, feedback is given and the learner has not remediated and takes the test again this prior probability is not changed the value or the usefulness of the test is not improved.

- **Arbitrariness.** Cut off scores often based on implicit assumptions.

- **Learning is complex.** There are multiple theories of how we learn. Evidence shows that continuous little bits rather than large amounts of cramming leads to better learning.

- **Prediction.** Assessment is not about trying to predict what students did in the past, but about trying to predict what they will do in the future. A single point measurement will not be as good predictor as longitudinal assessment in which a trajectory is plotted. This prediction can be even strengthened further by giving the student feedback and giving the student control over where they may want to end up.

- **Oversimplify complexity.** Competence like 'health' is a multifaceted approach. You would never tell a patient that they are 45.7% healthy and actually neither should you strive at telling a student that they are 45.7% competent. Numbers do play a role in healthcare (for example lab values) and so should they in assessment, but like you would not try to treat the health of the patient purely on lab values you would also not try to determine competence solely on numerical outcomes. In healthcare numbers are used in conjunction with narratives for example pathology reports. Both are reliable, but the numbers' reliability is based on statistics and population data and the pathology reports' reliability is based on the training and expertise of the pathologist, the clarity of the clinical question, and the clarity, coherence and plausibility of the pathology report.

- **Summary problems trying to solve**. Changing world, a focus on safe and independent doctors, learner's agency and empowerment, government actions, reductionism, arbitrariness, learning effects, predictive power, and complexity.

Focus on Implementation

Case Studies on Implementation of Programmatic Assessment   **~80 minutes**

Case study participants reflect on:

1. What are some smart design principles in this case study?
2. How closely does this model and assessment in the case study align with AMC standards?
3. What are some implementation strategies you can use in your own training program?

Case study presenters

Professor Cees van der Vleuten   Programmatic Assessment in the Netherlands

Professor Lambert Schuwirth   Programmatic Assessment Pilots at Flinders University

Ms Christine Cook   Programmatic Assessment in GP training

Professor Kichu Nair

Programmatic Assessment for IMGs

Dr Beth Mulligan

Professor Liz Molloy   Programmatic Assessment at Melbourne University

See attached case study booklet for summary.

Prof Kichu Nair and Dr Beth Mulligan presenting their case studies.

Profs Lambert Schuwirth and Cees van der Vleuten presenting case studies.
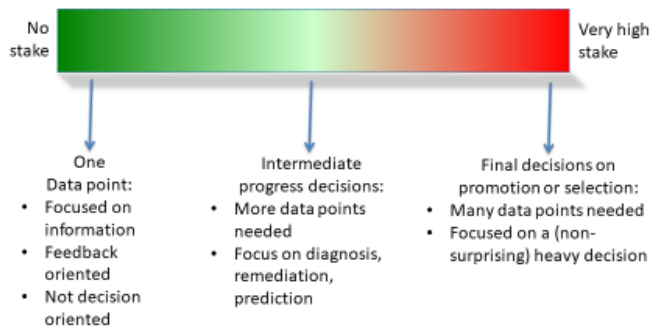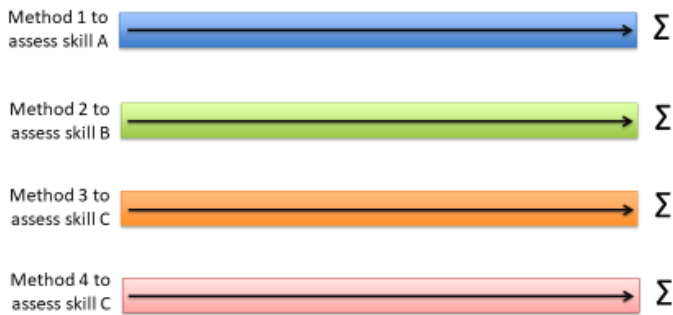
Prof Liz Molloy presenting a case study.

Christine Cook presenting a case study.
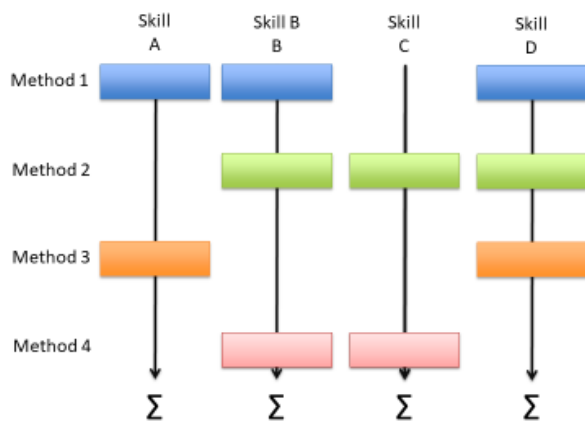
Top Tips: Implementation of Programmatic Assessment

Continuum of stakes, number of data point and their function

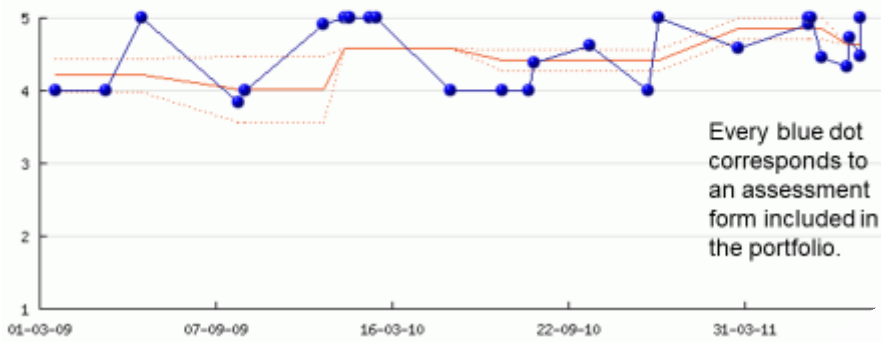No stake ——————————————————————→ Very high stake

One Data point:
- Focused on information
- Feedback oriented
- Not decision oriented

Intermediate progress decisions:
- More data points needed
- Focus on diagnosis, remediation, prediction

Final decisions on promotion or selection:
- Many data points needed
- Focused on a (non-surprising) heavy decision

Classical approach to aggregation

Method 1 to assess skill A ——————————→ Σ

Method 2 to assess skill B ——————————→ Σ

Method 3 to assess skill C ——————————→ Σ

Method 4 to assess skill C ——————————→ Σ

More meaningful aggregation

	Skill A	Skill B	Skill C	Skill D
Method 1				
Method 2				
Method 3				
Method 4				
	Σ	Σ	Σ	Σ

Maastricht Electronic portfolio (ePass)

1: Medical expert

Table view

Every blue dot corresponds to an assessment form included in the portfolio.

A model for programmatic assessment fit for purpose

C. P. M. VAN DER VLEUTEN[1], L. W. T. SCHUWIRTH[2], E. W. DR
D. TIGELAAR[3], L. K. J. BAARTMAN[4] & J. VAN TARTWIJK[5]

[1]Maastricht University, The Netherlands, [2]Flinders Medical School, Aust
The Netherlands, [4]Utrecht University of Applied Sciences, The Nethe

Abstract

We propose a model for programmatic assessment in action, wh
assessment for decision making about learner progress. This model
from empirical research. It specifies cycles of training, assessme
intermediate and final moments of evaluation on aggregated asse
are maximised for learning and feedback value, whereas high-st
Expert judgement plays an important role in the programme. F
with the inevitable subjectivity of this type of judgement. Bi
derived from criteria for qualitative research. We discuss a n
One of its prime virtues is that it enables assessment to m
individual instruments, towards a systems approach to ass

TWELVE TIPS

Twelve Tips for programmatic assessment

C.P.M. VAN DER VLEUTEN[1], L.W.T. SCHUWIRTH[2], E.W. DRIESSEN[1], M.J.B. GOVAERTS[1] &
S. HEENEMAN[1]

[1]Maastricht University, Maastricht, The Netherlands, [2]Flinders University, Adelaide, Australia

Abstract

Programmatic assessment is an integral approach to the design of an assessment program with the intent to optimise its
function, its decision-making function and its curriculum quality-assurance function. Individual methods of asse
purposefully chosen for their alignment with the curriculum outcomes and their information value for the learner, the teac
the organisation, are seen as individual data points. The information value of these individual data points is maximised by
feedback to the learner. There is a decoupling of assessment moment and decision moment. Intermediate and high
decisions are based on multiple data points after a meaningful aggregation of information and supported by ri
organisational procedures to ensure their dependability. Self-regulation of learning, through analysis of the assessment inform
and the attainment of the ensuing learning goals, is scaffolded by a mentoring system. Programmatic assessment-for-learnin
be applied to any part of the training continuum, provided that the underlying learning conception is constructivist. This p

Where to Next for Assessment?

Panel presentation ~30 minutes

Panel members to give a brief update on their thoughts about:

a) Next Steps in Medical education Assessment for health; and

b) Their ideas about international collaborations about assessment, sharing of best practice and AMC assessment standards review.

Panel members

	Prof Cees Van Der Vleuten	A lot has happened in Australia in Medical Education in the last few years – we have a real chance…
	Prof Lambert Schuwirth	The most important thing is that we have people from across the entire medical continuum in the one room. The only way to move forward is for everyone to work together and exchange experiences – all have different contexts and need own design approaches. In Belgium, there is a famous seafood restaurant – one enters through the kitchen. Let us all be like that – encourage others to enter through the kitchen and learn from our mistakes…
	Professor David Prideaux	I have used this metaphor before of Programmatic Assessment being like a Family photo album, which has studio portraits combined with some less good quality snapshots. Regardless of the differences in quality of the individual snapshots, we judge it overall as an album. Studio portraits – high stakes exams - do have a place in programmatic assessment just as part of the picture. To use another metaphor of the iceberg underneath that is the culture of feedback, mentoring and engagement that is critical. Implementation issues are tricky we have to think about all the stakeholders and how they might be part of it.

		Take home messages: there is a lot more to programmatic assessment than combining methods – it is a culture. Implementation – have to think about engagement of stakeholders.
	Professor Liz Farmer	The future, one important thing – this is a brave new world and new way of doing things. AMC standards need to move with the times in partnership with stakeholders, accreditation needs to be in line with the evidence and contemporary practice. Example of where the accreditor is being a thought leader and be part of the team. If standards are responsive to contexts then programmatic assessment should be responsible to the provider and it's individual.
	Ms Christine Cook	Align organization with the change, make sure everyone understands the change and is on board. Perhaps remember to walk through your kitchen yourselves. Culture of your organization will underpin everything you do.
	Prof Pete Ellis	I came today wondering what the problems were trying to fix. I heard lot of news (fake or real) about students concern about intern work readiness. The AMC survey results for the survey on work readiness will be finalised within the next few weeks. This has me thinking - Why are not they all work ready – I think we asking a lot more of them. We are focusing a lot on professionalism in assessment but perhaps not how we teach it. Professional training as socialization – you aren't socialized unless you are engaged with person you are learning with. Students are in a series of encounters. They have lost their connection with people. Programmatic assessment is promoting a more meaningful membership of individuals. It could go some way to addressing these difficulties and could remove students from the perspective there is no HD – always more to learn and they need to be self-motivated to get there.

Next Steps

- Circulation of a report summarising findings of day including review of key burning questions by working with experts post workshop.
- Opportunity to establish community of practice to work together on assessment innovations across the continuum and develop best practice guides to support assessment innovation.
- Review and updating of AMC standards on Assessment and Notes.

Further Reading

Biggs J. *Teaching for Quality Learning at University: What the Student Does.* Maidenhead UK: Open University Press 2003 (2nd ed).

Govaerts MJB, Van de Wiel MWJ, Schuwirth LWT, et al. Workplace-based assessment: raters' performance theories and constructs. *Advances in health sciences education* 2013;18:375-96. doi: DOI 10.1007/s10459-012-9376-x.

Harrison CJ, Könings KD, Dannefer EF, et al. Factors Influencing Students' Receptivity to Formative Feedback Emerging from Different Assessment Cultures. *Perspectives on Medical Education* 2016;5:276-84.

Heeneman S, Oudkerk Pool A, Schuwirth LWT, et al. The impact of programmatic assessment on student learning – theory versus practice - Medical Education 2015;49(5):487-98.

Schuwirth I, van der Vleuten C. Programmatic assessment: from assessment of learning to assessment for learning. Med Teach 2011;33:478-85.

Schuwirth LWT, Van der Vleuten CPM, Durning SJ. What programmatic assessment for learning in medical education can learn from healthcare. *Perspectives on Medical Education* 2017:1-5. doi: 10.1007/s40037-017-0345-1.

Schuwirth LWT, Van der Vleuten CPM. Programmatic assessment and Kane's validity perspective. Medical Education 2012;46(1):38-48. doi: 10.1111/j.1365-2923.2011.04098.x

Schuwirth LWT, Van der Vleuten CPM. Programmatic assessment: from assessment of learning to assessment for learning. Medical Teacher 2011;33(6):478-85.

Van der Vleuten C. The assessment of professional competence: developments, research, and practical implications. Adv Health Sci Educ 1996;**1** (1):41-67.

Van der Vleuten CPM, Schuwirth LWT, Driessen EW, et al. A model for programmatic assessment fit for purpose. *Medical Teacher* 2012;34:205-14. doi: 10.3109/0142159X.2012.652239.

Van der Vleuten CPM, Schuwirth LWT. Assessing professional competence: from methods to programmes. *Medical Education* 2005;39(3):309-17.