Standards for RACP Assessment Programs

July 2016

Contents

Introduction

In November 2011, a panel of external experts was formed to review the RACP's assessment strategy. Their review was based on submissions from and consultation with RACP Fellows, trainees and staff across the Divisions, Faculties and Chapters in Australia and New Zealand.

Following the review, the panel made a series of recommendations as documented in the April 2012 *'Report to RACP, the External Review of Formative and Summative Assessment'.*

These recommendations were considered by a number of staff and Fellows with leadership roles in assessment, teaching and learning at the External Review of Assessments Planning Forum in November 2012. Forum participants identified the development of assessment standards for RACP training programs as a priority.

This paper outlines the development process for the assessment standards, background information on existing local and international standards for assessments, and the Standards for RACP assessment programs. These standards were developed through a consultative process that included consultation with the College membership and comprehensive review by the College Censor and College Assessment Committee.

The Standards for RACP assessment programs supplement the College-wide *Assessment Policy* to guide the development, implementation and evaluation of individual assessments and assessment programs related to RACP training. RACP training programs are required to adhere to the standards and policy. It is recognised that implementation of the standards may take some time, and that any changes to existing assessment practices will need to be carefully planned and managed through the annual revision of training program requirements.

Executive summary

What are standards for assessment?

- Standards for assessment are the guiding principles for organisations when setting assessment tasks.

- Standards for assessment are used to define how assessment tools are chosen, implemented and evaluated.

- Standards for assessment clarify expectations of trainees and assessors regarding the purpose and use of assessments within a training program, and how information from those tools is translated into defensible decisions.

Why does the RACP need standards for assessment?

- RACP standards for assessment will help to ensure that assessment methods, the results produced and the decisions arising are reliable, valid, cost efficient, acceptable, feasible, have the desired educational impact, and reflect the objectives of the training program.

- A common set of standards for assessment will help create consistency in the design and implementation of assessments across College training programs.

What are the principles for assessment?

The establishment of best-practice principles underpins all robust assessment processes. The following key principles have been identified based on a review of best practice from other specialist medical colleges and standard setting bodies both locally and internationally; and consideration of the literature on assessment design and implementation.

- Clear educational value and rationale

- Programmatic assessment and blueprinting

- Fair and transparent processes and decision making

- Sustainability

- Quality feedback

- Appropriate support

- Evidence informed and practice based

What do the standards for RACP assessment programs look like?

- The standards for RACP assessment programs are structured as 30 standards within seven subcategories, in a framework of three overarching themes. These standards are to be used for developing new and reviewing existing assessment practices.

Planning
1. Clear educational value and rationale
2. Programmatic assessment and blueprinting

Implementation
3. Fair and transparent processes and decision making
4. Sustainability
5. Quality feedback
6. Appropriate support

Evaluation
7. Evidence informed and practice based

- The standards underpin the College Assessment Policy, which requires adherence to the Standards for RACP assessment programs.

- Both the standards and policy are supported by governance and oversight through the College Assessment Committee and College Censor. The established governance processes ensure that the correct expert bodies are signing off on assessment development, conduct and decisions, and that reporting pathways are clear and adhered to.

Standards for assessment in medical education

Standards for assessment are the guiding principles for education providers when setting assessment tasks. Standards for assessment are used to define how assessment tools are planned, implemented and evaluated, and serve to clarify expectations of trainees and assessors regarding the purpose and use of assessments within a training program. Such standards also show how information from those tools is translated into defensible decisions about progression through training.

Many key bodies, both within Australia/New Zealand and internationally, have undertaken considerable work in defining standards for assessments. These bodies include the Australian Medical Council (AMC), the UK's General Medical Council (GMC), the Royal College of Physicians and Surgeons of Canada (RCPSC) and the Accreditation Council for Graduate Medical Education, USA (ACGME).

A number of recommendations relating to standards for assessments were made following the *RACP External Review of Formative and Summative Assessments* in April 2012. Commonalities exist between these recommendations and the published standards for assessment of the AMC, GMC, RCPSC, and ACGME, as evidenced in *Appendix 1 – Commonalities of existing standards for assessment.*

Australian Medical Council (AMC) and Medical Council of New Zealand (MCNZ)

The AMC's Standards for Assessment and Accreditation of Specialist Medical Education Programs and Professional Development Programs set out the standards for accreditation that must be met by specialist medical training colleges. These standards have been jointly agreed and applied by the AMC and the MCNZ.

A number of these standards apply to the assessment of learning, and cover areas such as curriculum and assessment alignment, performance feedback and assessment quality*.*

International standards for assessment

The GMC, RCPSC and ACGME have each published a set of standards relating to the assessment of the training programs that they oversee.

General Medical Council (GMC), UK

The Standards for Curricula and Assessment Systems (GMC, UK) lists detailed standards under the five headings of planning, content, delivery, outcomes and review. Each of the 17 standards lists a number of mandatory requirements further clarifying the standard.

Royal College of Physicians and Surgeons of Canada (RCPSC)

The General Standards Applicable to All Residency Programs: B Standards (RCPSC) provides one standard for assessment of resident performance, detailed in a number of points relating to the interpretation of assessment data.

Accreditation Council for Graduate Medical Education (ACGME), USA

The ACGME Program Requirements for Graduate Medical Education in Internal Medicine (ACGME, USA) lists the program requirements for Graduate Medical Education in Internal Medicine, and includes requirements for formative and summative evaluation of residents.

The GMC standards are comprehensive, with a number of standards stipulated for both curricula and assessments. In comparison, the RCPSC and the ACGME stipulate fewer standards for assessments. Regardless of these differences, a number of commonalities exist between these international standards (see *Appendix 1)*.

Goals of assessment at the RACP

College training programs have established programs of assessment that include both formative and summative components. The goals of assessment at the College are to:

- Motivate trainees to learn.
- Engage in an accurate, timely and fair process to generate information of trainee competence for the individual, their supervisor, the College and the broader community.
- Provide progressive feedback on performance to ensure that learning is ongoing.
- Maintain professional standards to promote the highest quality patient care and public health.

The need for RACP standards for assessment

In November 2011, a panel of external experts was formed to review the assessment strategy of the RACP. The three external reviewers were Dr Tim Allen (Canada), Prof David Boud (Australia), and Dr Simon Newell (UK).

Their review was based on submissions and consultation with a number of focus groups comprised of RACP Fellows, trainees and staff across the Divisions, Faculties and Chapters in Australia and New Zealand.

Following the review, the panel made a series of recommendations as documented in the April 2012 *Report to RACP, the External Review of Formative and Summative Assessment*.

Recommendations related to standards for assessment included:

- "College-wide standards should be adopted for the design and administration (including marking and pass/fail decisions) of all summative examinations, and the College should provide support and direction to those who need help meeting these standards.

- Use the Van der Vleuten utility index (or a derivative) to review the assessment program and to justify future development or changes.

- Without neglecting the other factors in the index, consider the main purpose of the assessment program to be the driver of learning towards the achievement of competence and readiness for expert practice."

Principles of good assessment practices

Underpinning all robust assessment processes is the establishment of good-practice principles. The following key principles have been identified based on a review of best practice from other specialist medical colleges and standard setting bodies both locally and internationally; and consideration of the literature on assessment design and implementation.

- **Clear educational value and rationale.** It is widely accepted that assessment drives learning. It is vital that the assessments reflect the purpose and what is valued within the educational program, given that assessment shapes what trainees will learn.

- **Programmatic assessment and blueprinting.** A program of assessment is the planned and deliberate use of assessments rather than an arbitrary selection of tools and content for assessment. It is planned so that an overall picture of the trainee's competence can be formed through repeat observations and assessments in varying contexts by multiple assessors. As well as providing evidence for high-stakes decisions about competence, programs of assessment are designed to maximise learning throughout the training program.

  Programs of assessment should be aligned and blueprinted against curriculum standards, teaching and learning activities and other program requirements. Each individual assessment should also be blueprinted against the curriculum standards to encourage coverage of as many domains of practice as possible.

- **Fair and transparent processes and decision making.** Generally assessments contribute to decisions regarding whether a trainee is able to continue in their training program and reach their career goals. Therefore it is vital that assessment standards and processes are fair and publicly available.

- **Sustainability.** The resources required to develop, implement and maintain assessments should be sustainable over time.

- **Quality feedback**. Trainees should be provided with feedback following assessments to allow them to change and improve their practice. Trainees should also have the opportunity to provide feedback about their assessment experiences.

- **Appropriate support.** Clear communication and appropriate resources should be provided to assessors and trainees to support the implementation of assessments, including suitable training and feedback for assessors.

- **Evidence informed and practice based**. Assessments should be informed by available evidence and subject to a process of continuous quality improvement. The design, implementation and evaluation of assessments should be a collaborative process, including consultation with all impacted stakeholders and consideration of local needs.

Governance of RACP assessment programs

Assessment practices and their underpinning principles are supported by established governance arrangements and clear reporting requirements. As part of the RACP education governance review, new arrangements for governance of assessment were implemented in 2016 and included the establishment of the College Assessment Committee and appointment of a College Censor, who oversee these Standards for RACP assessment programs.

Standards for RACP assessment programs

The development of the standards for RACP assessment programs has been informed by:
- a review of best practice from other specialist medical colleges and standard setting bodies both locally and internationally
- consideration of the literature on assessment design and implementation
- feedback from Fellows and trainees involved in the *External Review of Assessments Planning Forum* in November 2012
- feedback from Division Education Committees, Advanced Training Committees, examination committees and trainee committees
- feedback from Committees, trainees and Fellows during the consultation phase of the Assessment Policy in 2015
- a comprehensive review by the College Censor and College Assessment Committee

This background research and consultation feedback has been incorporated to create the standards for RACP assessment.

Structure of the standards
It is important to consider the structure and categorisation of standards in order to promote clarity and usability of the standards. This approach has been followed in the GMC's Standards for Curricula and Assessment Systems, which are separated into the five distinct categories of *planning, content, delivery, outcomes* and *review.*

A similar approach has been taken for the proposed RACP standards for assessment, with standards grouped into the three categories of *plan, implement* and *evaluate.* These themes align with the development cycle for programs of assessment, commencing with a period of planning, prior to implementation followed by evaluation. Following evaluation it is expected that the program of assessment may be modified, bringing the process back to the planning phase (see *Figure 1).*

Each of the three categories of the standards for RACP assessment programs contains a brief description outlining the purpose of the category and key areas for consideration; this is followed by the standards for that category.
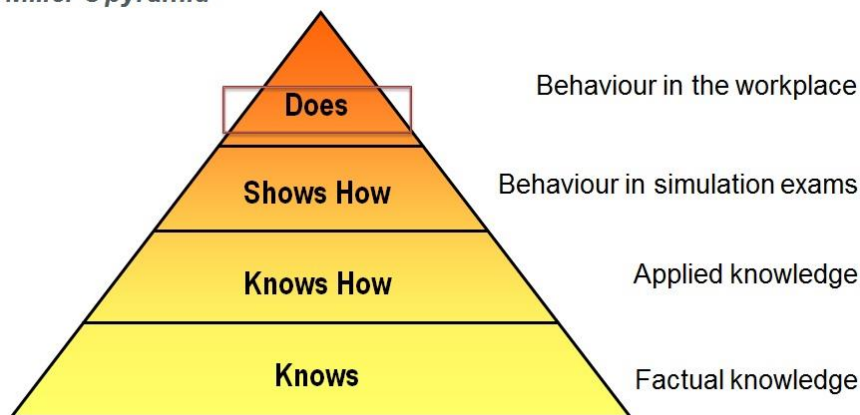
Figure 1.
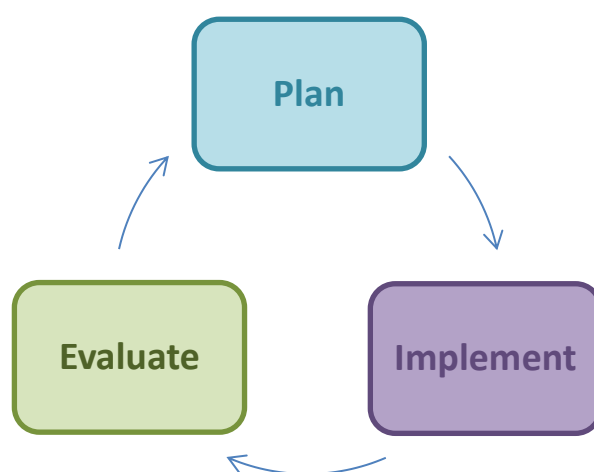
Planning

This section contains standards for the development of assessments, including determining the educational value and rationale of assessments, developing programs of assessment, ensuring that assessments are fit for purpose, and that the development of assessments is informed by evidence, and the context in which it will be used.

Miller's pyramid, shown below, provides a conceptual framework for the assessment of clinical competence.

Miller's pyramid

Does	Behaviour in the workplace
Shows How	Behaviour in simulation exams
Knows How	Applied knowledge
Knows	Factual knowledge

(Miller, 1990)

Assessment methods can be mapped against the various tiers of the pyramid:
- **Knows**. Tests factual recall, e.g. MCQs, EMQs, essay, oral exam
- **Knows how.** Assesses the application of knowledge within context to solve problems and make decisions, e.g. MCQs, EMQs, essay, oral exam
- **Shows how.** Assesses demonstration of competence, e.g., Long and short cases, OSCE, clinical simulations, standardised patient tests
- **Does.** Assesses actual performance in real life context, e.g. work-based assessments, entrustable professional activities.

Ideally assessments should evaluate performance in the context in which they are performed - assessing the 'does' level of Miller's pyramid. This is an ongoing area of focus for medical education research and is of particular importance to issues such as certification and revalidation of practitioners. Workplace-based assessment strongly aligns with the 'does' level, as it is authentic assessment of performance in the workplace context.

Clear educational value and purpose of assessments
It is widely accepted that assessment drives learning. To ensure that assessments direct the intended learning, the educational value and purpose of assessments should be clearly thought through when planning assessments. Users of assessments should also be aware of the intended educational value and purpose of assessments so that they can best reinforce that intention.

Programmatic assessment and blueprinting
A program of assessment is the planned and deliberate use of assessments rather than an arbitrary selection of tools and content for assessment. It is planned so that a whole overall picture of the trainee's competence can be formed. Whilst single assessments provide only a snapshot of the trainee's competence in a particular situation and point in time, repeat

observations and assessments in varying contexts by multiple supervisors help the supervisor to form an overall view of the trainee's competence (Van der Vleuten, Schuwirth, Driessesn, Dijkstra, Tigerlaar, Baartman & Van Tartwijk, 2012). 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.

Planning a program of assessment includes selection of a variety of assessment methods that sample as many situations as possible. 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 range of contexts and include different methods such direct observations, case discussions, written reflections. By collecting assessment information across the breadth of practice, a complete view of the trainee's performance can be formed.

As well as providing evidence for high-stakes decisions about competence, programs of assessment are designed to maximise learning throughout the training program. Supervisors will combine information that they gained from assessments and, together with the trainee, will use this to plan learning activities and learning goals. Information gained through assessments can be used to continuously measure progression towards the end goals of the learning program. This follows the concept of assessment *for* learning, as compared to assessment *of* learning (Schuwirth & Van der Vleuten, 2012).

Constructive alignment of curriculum standards, learning activities and assessments supports development of a program of assessment. While it is not feasible to assess each individual learning outcome within the curriculum, a program of assessment ensures that groups of curriculum standards are assessed. Ideally, each assessment should also provide feedback for a range of knowledge, skills and behaviours (Schuwirth & Van der Vleuten, 2011).

To create constructive alignment, a process of blueprinting is undertaken which involves reviewing each learning outcome and determining how it could best be learned and assessed (Holsgrove, Malik & Bhugra, 2009). Blueprinting is a tool used to define the parameters of an assessment task or program before it is developed, in order to determine in advance the information that will be assessed.

Planning – Standards

A program of assessment includes a mix of assessment activities, with methods that are matched to the required purpose or intent of the assessment, and implemented at an appropriate stage of training. Integrated assessment programs, aligned to desired curriculum standards, are important to gain a more complete picture of competence.

1. Clear educational value and purpose

Standard 1.1	The purpose of the proposed assessment program at each progression decision point is clearly stated and readily available, including how it relates to the learning required to progress through the next stage of training or practice.
Standard 1.2	The progression decisions, that will be made based on the assessment results, are clearly stated and readily available.

2. Programmatic assessment and blueprinting

Standard 2.1	The assessment methods are chosen in order to promote learning across the whole program.
Standard 2.2	A range of assessment methods are selected and blueprinted to ensure coverage of the depth and breadth of curriculum standards that are contained within the curriculum framework for each RACP training program.
Standard 2.3	The intended curriculum standards that may be sampled by each assessment method are identified and readily available.

Implementation

This section contains standards for the implementation of assessments, including fairness and transparency, sustainability of assessments, the provision of feedback, and implementation using appropriate methods of communication and training where necessary.

Fair and transparent process and decision making

In order for assessments to be successfully implemented, their organisation and conduct should be well planned. Trainees and assessors should be provided with adequate notice prior to the introduction of new assessments or changes to existing assessments to ensure that they are sufficiently prepared and able to make best use of the assessments. It is also important for the processes and policies surrounding assessments to be transparent and defensible to allow openness and accountability to regulators and the public.

Sustainability

Assessments need to be designed in such a way that the input of resources required to develop, implement and maintain them is sustainable over a period of time. Resources may include input from subject matter experts, for example assessment content from Fellows or assessment tool information from educationalists; costs to trainees or the College; and health system resourcing, such as arranging cover for Fellows or trainees completing College assessment work.

Quality feedback

The provision of feedback is an important aspect of both formative and summative assessments and should be provided according to performance on the task at hand, as well as overall progress through the curriculum and training program. Feedback should be used by trainees and their supervisors to plan for future learning, and amend their practice accordingly. Trainees should also have the opportunity to provide feedback about their assessment experiences.

Appropriate support

For successful implementation of assessments, stakeholders should have a high level of knowledge about the purpose and process of the assessments. This may involve training sessions or other appropriate forms of communication, such as emails or the availability of online resources. Assessors should be appropriately selected and trained.

Implementation – Standards

A number of supporting structures should be put in place when implementing a high quality program of assessment. This includes: the use of fair and transparent assessment processes and fair and transparent decision making; sustainability of assessments and assessment processes; the provision of feedback to trainees as a result of assessments; and the development of communication and training resources to engage stakeholders.

The process of implementing assessments also includes consideration of how changes will affect stakeholders and how assessments will be consulted on, implemented and evaluated, including process, educational rationale, examination construction, psychometric properties and examination outcomes.

3. Fair and transparent processes and decision making	
Standard 3.1	The required level of performance for each assessment is determined according to the standards contained within each curriculum.
Standard 3.2	Pass thresholds for examinations are criterion referenced where appropriate, or developed using recognised methodologies for standard setting.
Standard 3.3	A process of quality assurance is in place for all stages of the examination process, including the development of questions, construction of the examination, maintenance of security, data gathering, data collation and validation, and dissemination of results.
Standard 3.4	Program level blueprints, policies and criteria for progression through training are publicly available.
Standard 3.5	Trainees are provided with clear and accessible information about the purpose and processes of assessments.
Standard 3.6	Trainees and assessors are informed to whom assessment information will be provided and how it will be used.
Standard 3.7	Resources to support trainees and assessors are readily available prior to the implementation of new assessments or significantly revised assessments.
Standard 3.8	The process for Reconsideration, Review and Appeals is readily available.
Standard 3.9	Each training program documents its implementation of the College Special Consideration for Assessments Policy.
Standard 3.10	The College Impact Assessment Guidelines are used to determine the impact of any changes to assessments and the period of notice given prior to implementation.
Standard 3.11	High stakes decisions for trainees are made by an appropriately constituted panel of decision makers drawing on sufficient relevant information.

Standard 3.12	Measures employed to guard against bias in decision making are documented and readily available.
Standard 3.13	Decision making processes are subject to quality assurance measures.

4. Sustainability	
Standard 4.1	Logistical support for developing and maintaining assessments is appropriately resourced and managed.

5. Quality feedback	
Standard 5.1	Individual assessments generate feedback for trainees on their performance and progress in the particular areas being assessed in order for them to plan future learning.
Standard 5.2	Programs of assessment generate feedback for trainees on their performance and progress through the curriculum overall in order for them to plan future learning.
Standard 5.3	Feedback is provided by those who have undertaken appropriate training.

6. Appropriate support	
Standard 6.1	Assessors are appropriately selected, trained and supported, and their roles are clearly defined.
Standard 6.2	Designated College supervisors are provided with assessment performance information of the trainees for whom they have training responsibility.
Standard 6.3	Each training program adheres to the College Trainee in Difficulty Support policy and process.
Standard 6.4	There are procedures to inform the designated College supervisor and, where appropriate, the regulators, where patient or public safety concerns arise in any assessment. Disciplinary action in respect to employment or medical registration is a matter for the employer, the Medical Board of Australia (MBA) or the Medical Council of New Zealand (MCNZ)[1] as appropriate if there is evidence of serious breaches of care. Assessors are bound by mandatory notification requirements to the relevant medical board or council.

---

[1] http://www.medicalboard.gov.au/ (Australia), https://www.mcnz.org.nz/ (New Zealand)

Evaluation

This section contains standards for the evaluation and continuous quality improvement of assessments.

Evidence informed and practice based

The process of developing quality assessments involves two important sources of information. First, assessments should be developed on the basis of evidence of their effectiveness in promoting learning and quality assessment. Second, development of assessments should be guided by local content experts, such as Fellows and trainees who work in the settings in which the assessments are to be used, and who will be required to use the assessments. This involves careful piloting of new assessments in the planning phase and reference to the literature on assessment.

Regular monitoring and evaluation of assessments and assessment programs is essential to maintain quality assessments that sit well within the workplace setting. Evaluation should be conducted using published research and feedback from trainees, Fellows and other relevant stakeholders.

Assessments should be evaluated to ensure that the purpose aligns with the curriculum standards. For example, direct observations can be used to gain data on a specific encounter as it occurs in situ, while written examinations can be used to gain a broad picture of the breadth of clinical knowledge.

Other factors such as educational impact, reliability, acceptability and feasibility should also be considered when evaluating the usefulness of assessments and van der Vleuten's utility index (PMETB, 2007) is commonly used for this purpose. The index, shown below, is a conceptual model that defines utility (U) of an assessment as a product of several criteria:

$$U = E \times V \times R \times A \times F$$

- **Educational impact (E).** What is the educational purpose of the assessment? What are you aiming to assess?
- **Validity (V).** Did the assessment measure what it was intended to measure?
- **Reliability (R).** What is the quality of the results of the assessment? Are they consistent and reproducible?
- **Acceptability (A).** Is the assessment going to be accepted by the trainees and assessors? Will it drive learning or detract from learning?
- **Feasibility (F).** Can this assessment be implemented? What are the practicalities, e.g. cost, resources, availability of assessors.

The criteria in the model are weighted according to their relative importance in the assessment context. For example, the weights attached to each criteria in a high stakes summative assessment will be different to those used when the primary purpose of the assessment is to provide feedback in a work-based context (van der Vleuten & Schuwirth, 2005).

Evaluation – Standards

Regular evaluation of assessment tools and programs of assessment is essential to maintain quality assessments. Evaluation should be conducted using published research and feedback from trainees, Fellows and other relevant stakeholders.

Evaluation underpins the planning and implementation of assessments and programs of assessment.

7. Evidence informed and practice based

Standard 7.1	New programs or significant changes to programs of assessment include a period of consultation with relevant trainees, Fellows and additional stakeholders, prior to consideration by the College Assessment Committee.
Standard 7.2	Reporting pathways for both individual and aggregate assessment data are clearly established.
Standard 7.3	Assessments and programs of assessment are evaluated regularly, major reviews occur periodically, findings inform continuous quality improvement, and summaries are readily available.
Standard 7.4	A range of appropriate evaluation methods are used to review assessments and programs of assessment.

Related RACP policies and other documents

- Assessment Policy
- Reconsideration, Review and Appeals By-law
- Trainee in Difficulty Support Policy and Process
- Special Consideration for Assessments Policy
- Academic Honesty and Plagiarism Policy (to be superseded from Jan 2017 by the Academic Integrity in Training Policy and Process)
- Progression through Training Policy (revised version in effect from Jan 2017)
- Training Program Requirements Handbooks

Definitions

Assessment	The systematic process for measuring and providing feedback on the candidate's progress or level of achievement, against defined criteria.
Assessment program *(or program of assessment)*	An assessment program is where multiple measures of trainees' knowledge, skills and professional qualities over time are aggregated and synthesised to inform judgements and provide feedback about progress. It involves the planned and deliberate use of assessments to provide an overall picture of the trainee's competence over a specified period of time.
Assessor	Assessors are appropriately qualified and trained individuals who measure and provide feedback on a candidate's progress or level of achievement. College assessors may include but are not limited to: supervisors, examiners and committees.
Decision makers	A group of appropriately constituted people who draw on results of all relevant assessments in order to make well informed and justified high stakes decisions
Blueprinting	A blueprint defines the educational domains and related competencies covered by an assessment program. The blueprinting process involves aligning assessments with the intended curriculum standards and learning activities.
Designated College supervisor	Designated College supervisors are named supervisors on prospectively approved training periods, including but not limited to Directors of Physician Education, Educational Supervisors, and Advanced Training Supervisors.

References

- Accreditation Council for Graduate Medical Education (2013). *ACGME Program Requirements for Graduate Medical Education in Internal Medicine.* Retrieved July 30, 2013 from http://www.acgme.org/acgmeweb/Portals/0/PFAssets/2013-PR-FAQ-PIF/140_internal_medicine_07012013.pdf

- Allen, T., Boud, D., & Newell, S. (2012). *Report to RACP, the External Review of Formative and Summative Assessment, Final Report, April 2012.* Retrieved July 30, 2013 from http://www.racp.edu.au/index.cfm?objectid=AB250BCB-AADD-850D-2CB7F04F332332B1

- Australian Medical Council (2010). *Standards for Assessment and Accreditation of Specialist Medical Education Programs and Professional Development Programs by the Australian Medical Council 2010.* Kingston: Australian Medical Council Limited.

- Biggs, J. (2003). *Teaching for quality learning at university: What the student does.* (2nd ed). Maidenhead, UK: Open University Press.

- Epstein, R. (2007). Assessment in medical education. *The New England Journal of Medicine, 356:4,* 387-396

- General Medical Council (2010). *Standards for curricula and assessment systems.* Retrieved July 30, 2013 from http://www.gmc-uk.org/education/postgraduate/standards_for_curricula_and_assessment_systems.asp

- Hodges, B. (2013). Assessment in the post-psychometric era: Learning to love the subjective and collective. *Medical Teacher, 35,* 564-568.

- Holsgrove, G., Malik, A., & Bhugra, D. (2009). The postgraduate curriculum and assessment programme in psychiatry: the underlying principles. *Advances in psychiatric treatment, 15,* 114-122.

- Jackson, A. Jamieson & A. Khan (Eds), *Assessment in medical education and training: a practical guide* (pp. 11-26). Oxford: Radcliffe.

- PMETB. (2007). *Developing and maintaining an assessment system – a PMETB guide to good practice.* Retrieved November 6, 2012 from http://www.gmc-uk.org/Assessment_good_practice_v0207.pdf_31385949.pdf.

- Schuwirth, L. W. T, & Van der Vleuten, C. P. M. (2011). Programmatic assessment: From assessment of learning to assessment for learning. *Medical Teacher, 33,* 478-485.

- The Royal College of Physicians and Surgeons of Canada (2011). *General Standards Applicable to All Residency Programs, B Standards.* Canada: Author.

- Van der Vleuten, C. P. M., & Schuwirth, L. W. T. (2005). Assessing professional competence: from methods to programmes. *Medical Education, 39,* 309-317.

- Wass, I., Bowden, R., & Jackson, N. (2007) The principles of assessment design. In N. Jackson, A. Jamieson & A. Khan (Eds), *Assessment in medical education and training: a practical guide* (pp. 11-26). Oxford: Radcliffe.

Appendix 1 – Commonalities of existing standards for assessment

The following table shows commonalities between a range of existing standards for assessment and the recommendations from the External Review of Formative and Summative Assessments.

- *AMC* – Standards for Assessment and Accreditation of Specialist Medical Education Programs and Professional Development Programs by the Australian Medical Council 2010
- *GMC* – The Standards for Curricula and Assessment Systems
- *RSPSC* – The General Standards Applicable to All Residency Programs: B Standards (Royal College of Physicians and Surgeons of Canada)
- *ACGME* – The ACGME Program Requirements for Graduate Medical Education in Internal Medicine (Accreditation Council for Graduate Medical Education USA)
- *External review* – Recommendations from the External Review of Formative and Summative Assessments

Standard / Recommendation	AMC	GMC	RCPSC	ACGME	External review
Alignment between assessments and the goals and objectives of the curriculum	☑	☑	☑		☑
Systematic planning of assessments helps to collect assessment data	☑	☑	☑		
Performance standards for assessments are clear and identified	☑	☑	☑	☑	☑
Clear purpose for all assessments that is and available to trainees, educators, employers, professional bodies including the regulatory bodies, and the public	☑	☑			
Valid forms of assessment meaning that each different type of knowledge and skill is assessed using a different form of assessment (clinical, procedural, communication, professionalism assessed using methods such as written examinations, direct observation, multi-source feedback etc.)	☑	☑	☑	☑	
Measures of validity, reliability, feasibility, cost effectiveness, opportunities	☑	☑		☑	☑

Standard / Recommendation	AMC	GMC	RCPSC	ACGME	External review
for feedback and impact on learning used in the selection of assessments					
Use of lay and patient feedback in the development, implementation and feedback during assessments		☑			
Regular assessments provide relevant feedback to trainees	☑	☑	☑		☑
Feedback to supervisors on trainee's progress is facilitated through assessments	☑	☑	☑		☑
Documentation of assessment outcomes throughout the training program		☑	☑	☑	
Evidence from multiple sources is used to determine summative assessment decisions	☑		☑	☑	☑
Training for assessors, trainees and examiners in the use of assessments		☑			☑
Clear role of assessors/examiners that is used as the basis for recruitment and appointment		☑			
Trainees are informed when serious concerns exist with their training and given the opportunity to correct their performance	☑		☑		☑
Policies relating to disadvantage and special consideration for trainees during assessment are in place	☑				
A final report confirming sufficient competence for independent practice will be completed for trainees at the completion of their training			☑	☑	
Regular evaluation and improvement of education programs occurs	☑			☑	