

Twelve tips for improving the quality of assessor judgements in senior medical student clinical assessments

Bunmi S. Malau-Aduli, Richard B. Hays, Karen D'Souza, Shannon L. Saad, Helen Rienits, Antonio Celenza & Rinki Murphy

To cite this article: Bunmi S. Malau-Aduli, Richard B. Hays, Karen D'Souza, Shannon L. Saad, Helen Rienits, Antonio Celenza & Rinki Murphy (2023) Twelve tips for improving the quality of assessor judgements in senior medical student clinical assessments, Medical Teacher, 45:11, 1228-1232, DOI: [10.1080/0142159X.2023.2216364](https://doi.org/10.1080/0142159X.2023.2216364)

To link to this article: <https://doi.org/10.1080/0142159X.2023.2216364>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 26 May 2023.



Submit your article to this journal [↗](#)



Article views: 1073










View related articles [↗](#)



View Crossmark data [↗](#)

Twelve tips for improving the quality of assessor judgements in senior medical student clinical assessments

Bunmi S. Malau-Aduli^{a,b} , Richard B. Hays^b , Karen D'Souza^c , Shannon L. Saad^d , Helen Rienits^e , Antonio Celenza^f  and Rinki Murphy^g 

^aSchool of Medicine and Public Health, University of Newcastle, Newcastle, Australia; ^bCollege of Medicine and Dentistry, James Cook University, Townsville, Australia; ^cSchool of Medicine, Deakin University, Geelong, Australia; ^dRoyal Prince Alfred Virtual Hospital, Sydney, Australia; ^eGraduate School of Medicine, University of Wollongong, Wollongong, Australia; ^fSchool of Medicine, University of Western Australia, Perth, Australia; ^gMedical Program, University of Auckland, Auckland, New Zealand

ABSTRACT

Assessment of senior medical students is usually calibrated at the level of achieving expected learning outcomes for graduation. Recent research reveals that clinical assessors often balance two slightly different perspectives on this benchmark. The first is the formal learning outcomes at graduation, ideally as part of a systematic, program-wide assessment approach that measures learning achievement, while the second is consideration of the candidate's contribution to safe care and readiness for practice as a junior doctor. The second is more intuitive to the workplace, based on experience working with junior doctors. This perspective may enhance authenticity in assessment decisions made in OSCEs and work-based assessments to better align judgements and feedback with professional expectations that will guide senior medical students and junior doctors' future career development. Modern assessment practices should include consideration of qualitative as well as quantitative information, overtly including perspectives of patients, employers, and regulators. This article presents 12 tips for how medical education faculty might support clinical assessors by capturing workplace expectations of first year medical graduates and develop graduate assessments based on a shared heuristic of 'work-readiness'. Peer-to-peer assessor interaction should be facilitated to achieve correct calibration that 'merges' the differing perspectives to produce a shared construct of an acceptable candidate.

KEYWORDS

Clinical assessment; assessor judgement; senior medical student; objective structured clinical examination; clinical workplace system

Introduction

Medical programs aim to graduate competent doctors able to work within healthcare teams that provide safe clinical care. These programs are designed and implemented by medical education faculty to ensure that learners achieve graduate outcomes agreed by regulators, employers, and the wider profession. Ideally, there is constructive alignment of learning outcomes, curriculum content, clinical placement experience, and assessment tasks (Biggs 1996), as this modifies learning behaviour towards achieving graduate outcomes (Wang et al. 2013). Preferably, assessment should be systematic, comprehensive, program-wide and used for both assessment *for* and *of* learning (Norcini et al. 2018). This improves the scope of assessment across all domains, providing both narrative and quantitative information as evidence of learning (de Jong et al. 2022). During the earlier years, students are taught and assessed mainly by academic faculty in preparation for clinical learning. During more senior years, where program leaders have less influence over the learning environment, the clinical exposure, learning opportunities, and supervision vary according to the ebb and flow of clinical service delivery and clinical faculty are relied on for both teaching and assessment.

While it is often assumed that the assessment of senior medical students is focused on a common benchmark – how well a candidate has achieved the agreed learning outcomes, recent research reveals that there may be two ways of interpreting this. The first is the academic system that focuses on assessment of learning outcomes across domains throughout the program and at all levels of Miller's pyramid (Miller 1990). There are several progression points, many assessment methods and assessment is both *for* and *of* learning. A final clinical assessment often includes an OSCE, usually at the 'Shows How' level in a limited number of standardised, time-constrained encounters. The second system is the clinical workplace, where calibration may be influenced by experience as a supervisor and senior colleague of many junior doctors. This provides expectations of how well a candidate would fit into the clinical team as a potential colleague (Malau-Aduli et al. 2021). Here, the application of those outcomes in practice is assessed at both the 'Does' level of Miller's pyramid and the 'Is' level of professional identity, later added by Cruess et al. (2016). This appears to be a more intuitive, global judgement for clinical assessors, based on a heuristic or mental image that represents an ideal, or at least a safe, junior doctor. Such judgements include consideration of 'hard to measure' professional attributes such as teamwork, reliability,

CONTACT Bunmi S. Malau-Aduli  bunmi.malauaduli@newcastle.edu.au  School of Medicine and Public Health, University of Newcastle, Newcastle, Australia

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

trustworthiness, teachability, insight, and 'safety' (Malau-Aduli et al. 2022). While these attributes may depend on achieving knowledge, skills, and behaviours in the academic system, the focus is more on their application in the real world, combining qualitative and quantitative information. This resembles the 'legitimate subjectivity' referred to in judging an entrustable professional activity (EPA), an increasingly popular approach in postgraduate and specialty training (Ten Cate 2013; Ten Cate and Regehr 2019).

Many clinical assessors are busy clinicians with varying degrees of engagement with program design and delivery and yet must balance both academic and workplace perspectives when scoring candidate performance (Malau-Aduli et al. 2022). Differences between the two systems may produce a mismatch in the wording of checklists and rating scales, particularly for clinical assessors who are more familiar with the workplace perspective, possibly explaining some of the assessor variation encountered in clinical assessments. Assessment of senior students may be more complex because of imminent and potentially conflicting role changes of both assessors and candidates. Assessors may have contributed to clinical experiential learning of candidates, are now contributing to high-stakes progress decisions and may soon become senior colleagues who will rely on them as contributors to team function. Assessing senior student may be a final opportunity to assess work-readiness expectations. From the perspective of future employers and senior colleagues, getting this right is of the utmost importance.

Combining workplace expectations in assessment processes that enable decisions aligned with the real-world practice of clinical medicine, would aid incorporation of assessor judgements that draw on global concepts regarding professional attributes and behaviours, adding a salient, fair, and authentic dimension to these decisions (Govaerts and van der Vleuten 2013; Valentine et al. 2022). This approach requires assessment tools to 'ask the right questions' of the 'right people', particularly in workplace-based assessment (Crossley et al. 2011). Furthermore, as schools incorporate a more programmatic approach to assessment to optimise progression decisions, enhance feedback, and further drive professional lifelong learning, there is an identified need for the incorporation of expert judgements to support this (Boursicot et al. 2021). As 'experts' in clinical practice, specific inclusion of their perspectives may improve the authenticity of clinical assessment in OSCEs (Van der Vleuten and Schuwirth 2019).

As the scholarly understanding of authentic assessment judgements in the professional clinical environment evolves, medical schools have the important opportunity to enhance assessment processes aligned with best available evidence. This article presents 12 tips for how medical education faculty might improve clinical assessment by capturing workplace expectations to develop a shared heuristic of 'work-readiness'. This conforms to the evidence that improving assessor judgement is the priority in clinical assessment (van der Vleuten et al. 2012). While the underpinning theory was explored in final-year OSCEs, the suggestions may also be relevant to workplace-based clinical assessments.

Tip 1

Engage clinicians in designing a comprehensive program of clinical assessment

Overall clinical assessment design should be embedded in workplace expectations of recent graduates through collaboration with stakeholders from the clinical workplace (Norcini and Zaidi 2018). Different components of clinical competency can be assessed in diverse ways and awareness of the contribution of OSCE assessments may improve confidence that learning is assessed comprehensively. Observation in skills labs, simulated encounters such as in an OSCE, encounters with real patients (Mini-CEx, Case-based discussions, etc.), and longitudinal supervisor reports all contribute. A systematic approach to triangulation of such assessment methods may improve the utility of clinical assessment and potentially improve constructive alignment (van der Vleuten and Schuwirth 2019). 'Constructive' refers to the type of learning and what the learner does, while, 'alignment' refers to what the teacher does. The explicit connection between teaching, assessment, and learning outcomes helps to make the overall learning experience more transparent (Biggs 2003).

Tip 2

Design clinical assessment tasks that reflect the roles of first-year graduates

Include assessment tasks that align with imminent workplace roles of first year graduates' clinical case management. This may be more about *how*, *when* and *why* (or *why not*) than *what* to do. Examples include responding to a ward call; recognising the worsening clinical condition of a patient; conducting a clinical hand-over; referring a patient to another clinical service; communicating with a colleague; responding to an abnormal result; and writing a prescription for IV fluids or medication. A focus on 'doing' and 'being' in demonstrating application of clinical knowledge and skills is more authentic and meaningful (Ajjawi et al. 2020).

Tip 3

Design mark sheets that reflect both academic and clinical workplace perspectives

These may be two different perspectives, but the goal should be shared. The application of knowledge or skills (academic system) should explicitly consider reliability, safety, trustworthiness, and teachability (workplace system), better reflecting the intended roles of new graduates (Malau-Aduli et al. 2022). The latter will be more intuitive (or 'legitimately subjective') and more difficult to assess in typical OSCE stations (Valentine et al. 2022). The case content and marking rubrics should be designed to capture these elements while remaining user-friendly for clinician assessors who are less engaged with teaching.

Tip 4***Facilitate peer-to-peer assessor interaction to develop a shared understanding***

All assessors should be engaged in discussion about what is more or less important within each task, combining both academic and workplace perspectives. This helps assessors to discuss and reach a shared understanding of how to use marksheets and achieve correct calibration that ‘merges’ the differing perspectives and facilitates development of a shared construct of an acceptable candidate (de Jonge et al. 2017). This may be more effective in face-to-face interactions, although using remote streaming software may be sufficient for some aspects. Solitary, self-paced, on-line calibration may be less effective as it does not explicitly include interactive discussions (Sturman et al. 2018).

Tip 5***Engage clinicians in setting performance standards for each assessment task***

The same clinical assessors may be involved in assessment of learners at several stages of undergraduate and post-graduate training. Expectations are different at each level, and these should be explicit in setting standards and calibrating judgements. Global ratings are likely to reflect the ‘prototypical’ candidate heuristic of the clinical workplace, so making this more overt may capture more useful information. Wording overall rating scales as ‘how ready is this candidate to work in the clinical team as a new graduate?’ may more directly gather this information (Malau-Aduli et al. 2022).

Tip 6***Build consideration of patient safety into assessment tasks and marking sheet design***

Patient safety is a crucial consideration and is relevant in most clinical encounters. Criteria that indicate safety and assign marks for demonstrating safety can be included in several OSCE stations. Examples include prescribing incorrect doses of medication; not responding appropriately to abnormal investigation results; and not seeking assistance when it is needed. These attributes may be more suitable for workplace-based assessments, but inclusion where relevant in OSCE stations confirms its importance.

Tip 7***Include professionalism in assessment tasks and marking sheet design***

‘Correct’ professional behaviour is valued highly by clinical assessors in prospective members of their clinical team (Wilkinson et al. 2012). Aspects of professionalism may be included in several assessment tasks in both workplace-based assessment and OSCE stations. Examples include honesty about limitations in knowledge and skills; awareness of scope of practice of a new graduate and calling for help appropriately; respectful communication with a patient or colleague; demonstrating patient confidentiality;

and observing legal requirements when advising about driving following a seizure.

Tip 8***Decide in advance how to manage indications of poor professionalism***

When patient safety and professionalism are included within a clinical case or assessment task where application of knowledge and skills is the main focus, consistency of application of marking rubrics is important (Yepes-Rios et al. 2016). If most criteria are achieved but with a lapse in professionalism, does this automatically mean *fail* at that station or even of the whole assessment? The answer may depend on the precise nature of the task and the professionalism laps, but in general, use of ‘critical errors’ is not recommended. One alternative is to include professionalism and safety in several assessment tasks and indicate a yellow flag for each professionalism lapse (Yates 2011). All candidates with one or more yellow flags are discussed separately and other assessment data sought to determine if there is a pattern of poor professional behaviour. If assessors know that their yellow flag does not mean automatic failure, they may be more likely to use the mechanism, which would also provide the ability to generate better feedback for candidates.

Tip 9***Include ‘teachability’ in assessment tasks and marking sheet design***

Some gaps in knowledge and skills and experience may be forgiven, so long as candidates demonstrate self-awareness and effort to improve. This may also be a more intuitive judgment, related to a ‘capacity to change’ concept (Hays et al. 2002). Best considered in workplace-based assessment, some elements can be designed into OSCE stations. Examples include recognising lack of knowledge or confidence and seeking assistance; checking therapeutic guidelines and doses of medications; and willingness to accept feedback and advice. However, be aware of the situations where a candidate would ask/check everything, including what should be considered ‘working knowledge’.

Tip 10***Ensure that clinical assessors understand that data from several clinical assessments will contribute to assessment decisions***

Individual clinical assessor judgments are not the sole source of clinical assessment data as results are combined with information from other assessment events. In particular, OSCEs are not appropriate as the sole means of assessing learning outcomes that require longitudinal observation, such as reliability, trustworthiness, safety, and teachability (Malau-Aduli et al. 2022). If assessors understand that other assessments (ideally workplace-based) are included in progression decisions, combining formative, and summative purposes (Lockyer et al. 2017), they may be more likely to use the marksheets as intended rather than

think that this may be the last opportunity to prevent progression of weaker candidates.

Tip 11

Minimise assessor fatigue through workload management

Assessing candidates, particularly when performance is borderline, is mentally fatiguing (Malau-Aduli et al. 2021). While the evidence is conflicting in medical education, assessing a large number of OSCE candidates or assessing for more than one session per day may increase risk of assessor fatigue (Humphris and Kaney 2001; Byrne et al. 2014). Rest stations or rotating (but calibrated on the day) relieving assessors may reduce this risk. Fatigue may also be a factor in workplace-based assessments as they are often squeezed into busy clinical workloads, so scheduling time to prepare, conduct, and reflect on these assessments is important.

Tip 12

Ensure that all assessors attend an 'on the day' briefing

Memories of prior assessor training may fade over time and each assessment task is different. Assessors should familiarise themselves with the task and briefly clarify the calibration, ideally just before the assessment, so that the shared construct is fresh (Harasym et al. 2008). OSCE assessors arriving too late for the briefing should not start assessing 'cold'. A better strategy is to ensure that a small number of emergency reserve assessors who can fill in until late examiners are briefed by an experienced examiner. For workplace-based assessments, OSCE-style briefings may not be possible, but assessors should take time to read the marking sheets and patient notes.

Conclusions

Clinical assessments for senior students should align with both formal graduate outcomes and the expected roles and responsibilities of commencing professionals in a dynamic clinical environment. Experienced clinical assessors may approach assessment from the clinical workplace perspective – 'how well would this candidate fit in as a junior member of my clinical team?' – particularly when candidates perform at the borderline. These judgments are based on impressions of reliability, trustworthiness, patient safety, and teachability, in addition to knowledge application and skills. While there may be elements of subjectivity in these qualitative judgements, assessors should have opportunities to discuss and share their expectations so that any subjectivity reflects the genuine roles of graduates in the workplace. While 'snapshot' assessments are not the best way to assess these more intuitive judgements, mark sheets should include elements of behaviours within the workplace perspective and contribute to assessment of that attribute and/or domain. This article presents 12 tips for medical education faculty when preparing final clinical assessments, based on exploration of assessor thinking,

aiming to increase the alignment and utility of assessment at the transition from students to junior doctors.

Acknowledgements

The authors acknowledge members of the Australasian Collaboration for Clinical Assessment in Medicine (ACCLAIM).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

Notes on contributors

Bunmi S. Malau-Aduli, BSc (Hons), MSc, PhD, is a Professor of Medical Education at the School of Medicine & Public Health, University of Newcastle, Australia and Adjunct Professor at the College of Medicine & Dentistry, James Cook University, Townsville, Australia. She is Co-Chair of ACCLAIM. Her main research interests include quality assurance of assessment, cultural competency, and program evaluation.

Richard B. Hays, MBBS, PhD, MD, FRACGP, FACRRM, is a Professor of Remote Medicine and Health at the College of Medicine & Dentistry, James Cook University, Townsville, Australia. His main research interests are curriculum and assessment design and evaluation.

Karen D'Souza, MBBS (Hons), currently holds the roles of Student Progression and Welfare Lead, and Co-ordinator of Years 3 and 4 for the Doctor of Medicine program at Deakin University, Australia. She is Co-Chair for both the Australasian Collaboration for Clinical Assessment in Medicine (ACCLAIM) and the Medical Deans of Australia and New Zealand Student Support Network. She has a professional and research interest in clinical assessment, clinical and communication skills, professionalism and professional identity formation, academic support and wellbeing.

Shannon L. Saad, BSc (Hons1), MBBS, MMedEd, FRACGP, is a General Practitioner and Staff Specialist in Virtual Healthcare in Sydney, NSW. She is an Executive member of ACCLAIM. She has extensive experience in teaching and research in medical education, including the fields of healthcare communication, clinical skills, and clinical assessment.

Helen Rienits, MBBS, DRANZCOG, Grad Cert Med. Ed, PhD (Medical Education), is a General Practitioner at the South Coast NSW, Academic Leader for Clinical Skills and Theme Lead for Clinical Competence at the Graduate School of Medicine, Faculty of Science, Medicine and Health (SMAH), University of Wollongong. She is an Executive member of ACCLAIM. Her research interests include medical education and assessment of clinical performance in medicine.

Antonio Celenza, MBBS MCLinEd FACEM FRCEM, is a Professor and Chair of Emergency Medicine at The University of Western Australia and a specialist Emergency Physician at Sir Charles Gairdner Hospital in Perth, Western Australia. He has previously had roles as Professor of Medical Education and Director of the UWA MBBS/MD programs. He is an Executive member of ACCLAIM. His research track record includes the areas of clinical education, skills teaching and assessment, emergency health systems, cardiac arrest outcomes, and acute aged care.

Rinki Murphy, MBChB, FRACP, PhD, is a Professor of Medicine and the MBChB Year 4 clinical skills assessment chair/convenor at the Medical School, University of Auckland, New Zealand. She is an Executive member of ACCLAIM. Her research interests are mainly in physiology of diabetes/obesity and clinical skills assessment.

ORCID

Bunmi S. Malau-Aduli  <http://orcid.org/0000-0001-6054-8498>
 Richard B. Hays  <http://orcid.org/0000-0002-3875-3134>
 Karen D'Souza  <http://orcid.org/0000-0003-3230-5100>
 Shannon L. Saad  <http://orcid.org/0000-0002-0423-5478>
 Helen Rienits  <http://orcid.org/0000-0002-1159-8357>
 Antonio Celenza  <http://orcid.org/0000-0001-9792-9526>
 Rinki Murphy  <http://orcid.org/0000-0002-0043-2423>

References

- Ajjawi R, Tai J, Huu Nghia TL, Boud D, Johnson L, Patrick CJ. 2020. Aligning assessment with the needs of work-integrated learning: the challenges of authentic assessment in a complex context. *Assess Eval High Educ.* 45(2):304–316.
- Biggs J. 1996. Enhancing teaching through constructive alignment. *High Educ.* 32(3):347–364.
- Biggs J. 2003. *Teaching for quality learning at university*. 2nd ed. Buckingham: Open University Press.
- Boursicot K, Kemp S, Wilkinson T, Findyartini A, Canning C, Cilliers F, Fuller R. 2021. Performance assessment: consensus statement and recommendations from the 2020 Ottawa Conference. *Med Teach.* 43(1):58–67.
- Byrne A, Tweed N, Halligan C. 2014. A pilot study of the mental workload of objective structured examination examiners. *Med Educ.* 48(3):262–267.
- Crossley J, Johnson G, Booth J, Wade W. 2011. Good questions, good answers: construct alignment improves the performance of workplace-based assessment scales. *Med Educ.* 45(6):560–569.
- Cruess RL, Cruess SR, Steinert Y. 2016. Amending Miller's pyramid to include professional identity formation. *Acad Med.* 91(2):180–185.
- de Jong LH, Bok HGJ, Schellekens LH, Kremer WDJ, Jonker FH, van der Vleuten CPM. 2022. Shaping the right conditions in programmatic assessment: how quality of narrative information affects the quality of high-stakes decision-making. *BMC Med Educ.* 22(1):409.
- de Jonge LPJWM, Timmerman AA, Govaerts MJB, Muris JWM, Muijtjens AMM, Kramer AWM, van der Vleuten CPM. 2017. Stakeholder perspectives on workplace-based performance assessment: towards a better understanding of assessor behaviour. *Adv Health Sci Educ Theory Pract.* 22(5):1213–1243.
- Govaerts M, van der Vleuten CP. 2013. Validity in work-based assessment: expanding our horizons. *Med Educ.* 47(12):1164–1174.
- Harasym PH, Woloschuk W, Cuning L. 2008. Undesired variance due to examiner stringency/leniency effect in communication skill scores assessed in OSCEs. *Adv Health Sci Educ Theory Pract.* 13(5):617–632.
- Hays RB, Davies HA, Beard JD, Caldon LJM, Farmer EA, Finucane P M, McCrorie P, Newble DI, Schuwirth LWT, Sibbald GR. 2002. Selecting performance assessment methods for experienced physicians. *Med Educ.* 36:910–917.
- Humphris GM, Kaney S. 2001. Examiner fatigue in communication skills objective structured clinical examinations. *Med Educ.* 35(5):444–449.
- Lockyer J, Carraccio C, Chan M-K, Hart D, Smee S, Touchie C, Holmboe ES, Frank JR, on behalf of the ICBME Collaborators. 2017. Core principles of assessment in competency-based medical education. *Med Teach.* 39(6):609–616.
- Malau-Aduli BS, Hays RB, D'Souza K, Smith AM, Jones K, Turner R, Shires L, Smith J, Saad S, Richmond C, et al. 2021. Examiners' decision-making processes in observation-based clinical examinations. *Med Educ.* 55(3):344–353.
- Malau-Aduli BS, Hays RB, D'Souza K, Jones K, Saad S, Celenza A, Turner R, Smith J, Ward H, Schlipalius M, et al. 2022. "Could you work in my team?": exploring how professional clinical role expectations influence decision-making of assessors during exit-level medical school OSCEs. *Front Med.* 9:844899.
- Malau-Aduli BS, Jones K, Saad S, Richmond C. 2022. Has the OSCE met its final demise? Rebalancing clinical assessment approaches in the peri-pandemic world. *Front Med.* 9:825502.
- Miller GE. 1990. The assessment of clinical skills/competence/performance. *Acad Med.* 65(9 Suppl.):S63–S67.
- Norcini J, Anderson MB, Bollela V, Burch V, Costa MJ, Duvivier R, Hays R, Palacios Mackay MF, Roberts T, Swanson D. 2018. Consensus framework for good assessment. *Med Teach.* 40(11):1102–1109.
- Norcini J, Zaidi Z. 2018. Workplace assessment. In: Swanwick T, Forrest K, O'Brien B, editors. *Understanding medical education*. 3rd ed. Oxford: Wiley Blackwell; p. 319–334.
- Sturman N, Wong WY, Turner J, Allan C. 2018. Online examiner calibration across specialties. *Clin Teach.* 15(5):377–381.
- Ten Cate O, Regehr G. 2019. The power of subjectivity in the assessment of medical trainees. *Acad Med.* 94(3):333–337.
- Ten Cate O. 2013. Nuts and bolts of entrustable professional activities. *J Grad Med Educ.* 5(1):157–158.
- Valentine N, Durning S, Shanahan E, Van der Vleuten C, Schuwirth L. 2022. The pursuit of fairness in assessment: looking beyond the objective. *Med Teach.* 44(4):353–359.
- Van der Vleuten CPM, Schuwirth LWT, Driessen EW, Dijkstra J, Tigelaar D, Baartman LKJ, van Tartwijk J. 2012. A model for programmatic assessment fit for purpose. *Med Teach.* 34(3):205–214.
- van der Vleuten CPM, Schuwirth LWT. 2019. Assessment in the context of problem-based learning. *Adv Health Sci Educ Theory Pract.* 24: 903–914.
- Wilkinson TJ, Moore M, Flynn EM. 2012. Professionalism in its time and place: some implications for medical education. *NZ Med J.* 29(1358): 64–73.
- Wang X, Su Y, Cheung S, Wong E, Kwong T. 2013. An exploration of Biggs' constructive alignment in course design and its impact on students' learning approaches. *Assess Eval High Educ.* 38(4): 477–491.
- Yates J. 2011. Development of a 'toolkit' to identify medical students at risk of failure to thrive on the course: an exploratory retrospective case study. *BMC Med Educ.* 11(1):95.
- Yepes-Rios M, Dudek N, Duboyce R, Curtis J, Allard RJ, Varpio L. 2016. The failure to fail underperforming trainees in health professions education: A BEME systematic review: BEME Guide No. 42. *Med Teach.* 38(11):1–8.