Workplace-based assessments: what, why, and how to implement?

HY So, FHKCA, FHKCA (IC), **YF Choi,** FHKEM, **PT Chan,** FHKCOS, **Albert KM Chan,** FHKCA, FANZCA, **George WY Ng,** FHKCP, **George KC Wong *,** MD, FCSHK

The Hong Kong Jockey Club Innovative Learning Centre for Medicine, Hong Kong Academy of Medicine, Hong Kong SAR, China

* Corresponding author: georgewong@fellow.hkam.hk

This article was published on 3 Jun 2024 at www.hkmj.org.

Hong Kong Med J 2024;30:250–4
https://doi.org/10.12809/hkmj2311361

Introduction

Assessments in postgraduate medical education have undergone significant changes over the past few decades.[1] We are more familiar with assessment methods that assess the 'knows', 'knows how', and 'shows' levels of Miller's pyramid, also known as 'assessment of competence' (online supplementary Fig).[2,3] These methods emphasise objectivity through standardisation and by minimising the role of human judgement.

However, in the 1990s, several factors led to a shift in thinking. First, it was recognised that assessment methods prioritising objectivity (rather than professional judgement) can oversimplify complex skills, diminishing the true value of the assessment.[4] It was also understood that clinical encounters are 'context-specific', and that competency lies in doctors' abilities to adapt and respond to the various circumstances they encounter.[5] 'Assessments of competence' conducted in controlled settings have weak correlations with doctors' actual practices in real clinical settings.[2] Furthermore, the introduction of competency-based medical education has highlighted the importance of skills such as communication, collaboration, and professionalism, which are not easily quantifiable.[6] These factors indicate a need to return assessments to the clinical environment. Additionally, educators have found that excessively focusing on objectivity and quantitative results for summative purposes can cause students to prioritise succeeding in the assessments, rather than learning to become good clinicians. It is important to address the impact of assessments on learning by involving learners as active participants and providing them with meaningful feedback.[7] The current consensus is that expert judgement should be recognised and respected during the assessment process.[8]

Workplace-based assessments (WBAs) involve the assessment of day-to-day practices within the working environment.[9] They represent a form of 'assessment of performance' which evaluates doctors' actual professional practices.[2,3] These types of assessments can include direct observation of clinical procedures and patient management, or retrospective presentation of cases. Each assessment is followed by guided reflection to identify possible learning points. Action plans should be formulated and subsequently carried out. Various WBA tools have been defined, and tools currently in use by our Colleges are summarised in online supplementary Table 1.[10]

Purposes

An integrated set of WBAs can be designed primarily for learning enhancement (formative) or performance evaluation (summative). The design of the WBAs should be aligned with their intended purpose. The use of WBAs as formative assessments may have more learning benefits compared with their use as summative assessments alone, or their use as combined assessments.[11] Confusion surrounding the purposes of WBAs is a common obstacle hindering effective implementation among trainers and trainees. The Table provides a summary of the features of WBAs as formative assessments in comparison with traditional summative assessments.[12]

Confusion about the purposes of WBAs can lead to misconceptions, such as the use of psychometric criteria of validity and reliability to evaluate WBAs. The validity of WBAs is primarily supported by their authenticity.[8] Additionally, the validity of WBAs as a formative assessment relies on high-quality feedback from trainers and feedback literacy among trainees.[13,14]

Because WBAs are non-standardised assessments, factors such as case selection, context restriction, and rater cognition can influence inter-rater variability. There are three sources of variability related to rater cognition.[15] First, trainers may fail to correctly apply assessment criteria. Training for the trainers can reduce this source of variability. Second, variability can arise from limitations in human cognition, leading to various forms of bias. Efforts to understand the impacts of cognitive influences and use cognitive tools can help address this variability. Finally, competence is a complex phenomenon; different trainers may focus on unique aspects that actually are complementary. This 'meaningful

 ©2024 Hong Kong Academy of Medicine. CC BY-NC-ND 4.0

TABLE. Comparison of workplace-based assessments and conventional summative assessments[12]

	Workplace-based assessments	Conventional summative assessments
Examples of assessment tools	EPA, Mini-CEX, DOPS, CBD, chart review, MSF	Multiple choice questions, short answer questions, Viva voce, written essay
Areas of assessment	Able to assess multiple clinical competencies such as procedure skills, patient care, service improvement, communication, and professionalism	Mainly written examinations. Assess textbook knowledge. Some may partially assess clinical competency (eg, OSCE)
Timing of assessment	Throughout the training period, enabling continuous and programmatic assessment	Usually at the end of a training period
Results of assessment	Extensive information (feedback) given to trainees, with or without summative decisions	A mark, grade, or decision with limited information
Impacts of assessment on trainees' shortcomings	Facilitates timely improvement for trainees based on feedback. Ideally, enables early identification of underperformance	A summative decision is made at the end of the training period. Remedial training and discontinuation of training are consequences of failure. Underperformance is recognised at the end of the training period Trainees may learn how to improve when it is too late
Generalisation of the assessment process	Direct improvements to patient care and other clinical competencies	Because only textbook knowledge is examined, results cannot be generalised to clinical competency
Level of assessment	Intended to be low-stakes environment	High-stakes environment

Abbreviations: CBD = Case-Based Discussion; DOPS = Direct Observation of Procedural Skills; EPA = entrustable professional activity; Mini-CEX = Mini–Clinical Evaluation Exercise; MSF = Multi-Source Feedback; OSCE = objective structured clinical examination

idiosyncrasy' is not considered problematic—it represents a strength of this form of assessment. When WBAs are used for formative purposes, reliability is not a major concern; when they are used for summative purposes, reliability should be considered.[8] The main determinant of reliability in all types of assessments is sample size, rather than 'objectivity'. Therefore, it is important to ensure that each trainer conducts an adequate number of assessments.[16]

Implementation

To address the challenges of integrating assessments into the clinical environment, we used the Consolidated Framework for Implementation Research (Fig[17]) for categorisation of issues identified in the existing literature and in the results of a Younger Fellows Chapter survey conducted at the Hong Kong Academy of Medicine Medical Education Conference 2021 (online supplementary Table 2).[18-22] Based on these identified issues and recommendations from the Ottawa Conference 2020, we propose the following implementation framework.[8]

Design workplace-based assessments according to their intended purposes

Because WBAs are most beneficial as formative assessments, the focus should be on designs that maximise their impacts on learning.[8] It is crucial to involve both trainers and trainees in the design process; this ensures that their input is incorporated.[18] The WBA tools should be user-friendly and utilise simple language.[18,20] Although the application of a checklist to facilitate identification of specific feedback may be helpful, the checklist should not be overly burdensome.[20] The use of digital technology for documentation can improve accessibility to WBA tools and enable data collection for learning analytics.[12,19] Assessments should focus on narrative feedback instead of rating scales or scores. Whenever possible, the decision at the end of each WBA should be based on narrative comments

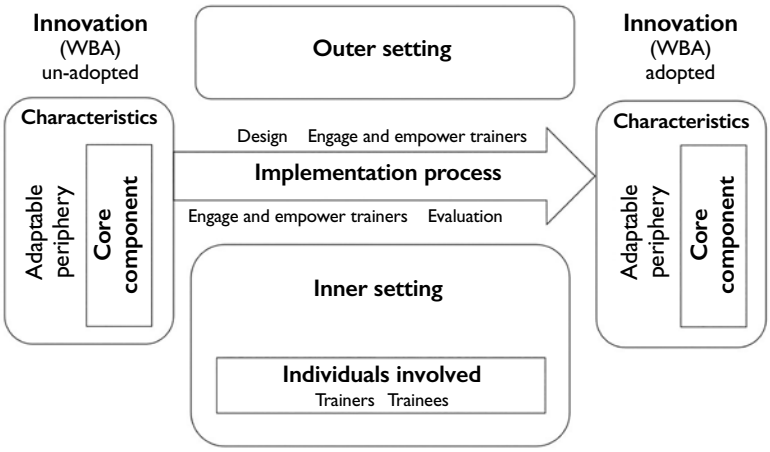

FIG. Consolidated Framework for Implementation Research[17]
Abbreviation: WBA = workplace-based assessment

that aid learning, rather than a pass/fail decision, to avoid the 'failure-to-fail' phenomenon.[8]

Work structure is also important. Trainees often rotate through multiple wards or hospitals, resulting in short and constantly changing relationships with their trainers. This can make it difficult for supervisors to assess a trainee's performance because there is a lack of familiarity. It is challenging but crucial to foster longitudinal and trusting relationships between trainees and trainers, such as by prolonging trainees' rotations or assigning them specific trainers for longer periods of time.[8,20-22]

Engage and empower trainers

The effectiveness of WBAs is greatly influenced by trainers' knowledge and understanding of how to conduct assessments and provide feedback to trainees.[19-21] Attainment of this knowledge and understanding requires trainers to familiarise themselves with relevant assessment tools and engage in medical education, which is currently not included in most Colleges' fellowship training programmes. Trainers' willingness to engage in WBAs is affected by organisational culture and the value placed on teaching and feedback. A lack of understanding regarding WBAs can also lead to a lack of engagement.[19,20] Therefore, all trainers involved in WBA should be required to receive training focused on conducting assessments and understanding the rationale behind them.[19]

The quality of trainer feedback is crucial for effective learning and for trainees to recognise the value of WBAs. Trainers must be skilled in providing feedback.[18-20] They should also ensure that the tasks selected for assessments are appropriate for each trainee's level of experience and competence.[19] To address these issues, the Hong Kong Jockey Club Innovative Learning Centre for Medicine (HKJC ILCM) has developed Train-the-Trainer WBA Courses in collaboration with various Colleges.

Engage and empower trainees

If the purposes of WBAs are not clear during implementation, the tools may be used ineffectively; trainees may cynically view the assessments as a 'reductive "tick-box"' approach to evaluating the complexities of professional behaviour. Trainees should also understand that WBAs are designed for formative purposes, not summative purposes; the perception that WBAs serve as summative assessments may encourage learners to adopt strategic and undesirable behaviours, such as avoiding discussion of challenging patient cases or seeking lenient assessors.[18,19] Therefore, it is equally important to engage trainees by explaining the purposes and uses of WBAs.[18,19] The HKJC ILCM has piloted a WBA Trainee Course to improve trainees' feedback literacy and to promote a growth mindset and self-regulated learning.[14,20]

Evaluate the implementation process

Given that WBAs are considered an ever-evolving approach, it is essential for Colleges to establish mechanisms for regular evaluation of the implementation process to ensure that the WBAs remain relevant and effective.[23]

Resolve the issue of time constraints

Numerous studies have consistently highlighted the challenge of allocating sufficient time for trainees and trainers to integrate WBAs into their daily routines.[18-21] According to information from informal communication with different Colleges, most local surveys showed that debriefing sessions ranged from 10 to 20 minutes per WBA. Recognising this challenge, the Hong Kong Academy of Medicine emphasised the importance of ongoing discussions and collaborative efforts among various parties to address the resource implications of WBA implementation in its recent position paper concerning postgraduate medical education.[23] Additionally, resource allocation is influenced by organisational culture and the value placed on teaching and feedback.[20,21]

The way forward

We have discussed how assessments in medical education evolved from a measurement role to a judgement role. Another paradigm shift, which began in around 2010, has led to the perception of assessments as systems.[1] Medical education requires multiple cognitive, psychomotor, and attitudinal/relational skills. Because no single assessment method can capture all of these skills, multiple measures are necessary. However, if these assessments are applied in an uncoordinated manner and combined to reach an overall decision based on traditional weighting, they cannot effectively reflect a trainee's competence. An assessment system should integrate and combine single assessments to meet the diverse needs of various stakeholders.[24] Therefore, each single WBA tool should be part of an integrated, coherent set of WBAs; this set of WBAs should be embedded in a broader assessment system.[8] Attention should be given to the criteria for creating effective assessment systems.[24] Programmatic assessment, a logical approach for building such systems,[8,25] is based on the principle that each assessment method or tool has limitations; compromises are needed if individual assessments alone are used for pass/fail decisions. A contrasting perspective is that each assessment should be regarded as a single data point and optimised for learning by providing meaningful feedback to the learner. Pass/fail and

high-stakes decisions should be made in a credible and transparent manner, using multiple data points in a holistic approach.[25]

There are several unresolved issues regarding WBAs that warrant further investigation.[8,18] These include inquiries into the effectiveness of individual WBA tools at various levels of training, the potential extension of WBAs into continuing professional development, and the use of WBAs to assess complex outcomes and competencies (eg, teamwork). Additionally, there is need to identify the optimal method for synthesising WBA results that can support informed decisions and promote learning. It is also worth exploring whether a programmatic approach to WBAs could enhance their learning effects. Considering the context-specific nature of educational interventions, the HKJC ILCM should collaborate with College fellows to conduct local investigations that address these questions.

Author contributions

All authors contributed to the concept or design, acquisition of data, analysis or interpretation of data, drafting of the manuscript, and critical revision of the manuscript for important intellectual content. All authors had full access to the data, contributed to the study, approved the final version for publication, and take responsibility for its accuracy and integrity.

Conflicts of interest

All authors have disclosed no conflicts of interest.

Funding/support

This commentary received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Supplementary material

The supplementary material was provided by the authors and some information may not have been peer reviewed. Accepted supplementary material will be published as submitted by the authors, without any editing or formatting. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by the Hong Kong Academy of Medicine and the Hong Kong Medical Association. The Hong Kong Academy of Medicine and the Hong Kong Medical Association disclaim all liability and responsibility arising from any reliance placed on the content. To view the file, please visit the journal online (https://doi.org/10.12809/hkmj2311361).

References

1. Schuwirth LW, van der Vleuten CP. A history of assessment in medical education. Adv Health Sci Educ Theory Pract 2020;25:1045-56.
2. Rethans JJ, Norcini JJ, Barón-Maldonado M, et al. The relationship between competence and performance: implications for assessing practice performance. Med Educ 2002;36:901-9.
3. Miller GE. The assessment of clinical skills/competence/performance. Acad Med 1990;65(9 Suppl):S63-7.
4. Norman GR, Van der Vleuten CP, De Graaff E. Pitfalls in the pursuit of objectivity: issues of validity, efficiency and acceptability. Med Educ 1991;25:119-26.
5. ten Cate O, Snell LS, Carraccio C. Medical competence: the interplay between individual ability and the health care environment. Med Teach 2010;32:669-75.
6. Frank JR, Snell LS, ten Cate O, et al. Competency-based medical education: theory to practice. Med Teach 2010;32:638-45.
7. Cilliers FJ, Schuwirth LW, Adendorff HJ, Herman N, van der Vleuten CP. The mechanism of impact of summative assessment on medical students' learning. Adv Health Sci Educ Theory Pract 2010;15:695-715.
8. Boursicot K, Kemp S, Wilkinson T, et al. Performance assessment: consensus statement and recommendations from the 2020 Ottawa Conference. Med Teach 2021;43:58-67.
9. Postgraduate Medical Education and Training Board Workplace Based Assessment Subcommittee. Workplace based assessment. 2005. Available from: https://webarchive.nationalarchives.gov.uk/ukgwa/20091211203900/http://www.pmetb.org.uk/media/pdf/3/b/PMETB_workplace_based_assemment_paper_(2005).pdf. Accessed 7 May 2024.
10. Royal College of Physicians and Surgeons of Canada, Royal Australasian College of Physicians, Royal Australasian College of Surgeons. Work-based assessment: a practical guide. 2014. Available from: https://www.surgeons.org/-/media/Project/RACS/surgeons-org/files/becoming-a-surgeon-trainees/work-based-assessment-a-practical-guide.pdf?rev=64c62242e777411eb43be8ac781dfa4a&hash=DCEE633AC11B7EE63975DF1A6948C99A. Accessed 7 May 2024.
11. Harlen W, James M. Assessment and learning: differences and relationships between formative and summative assessment. Assess Educ Principles Policy Pract 1997;4:365-79.
12. Garrison C, Ehringhaus M. Formative and summative assessments in the classroom. Available from: https://www.amle.org/wp-content/uploads/2020/05/Formative_Assessment_Article_Aug2013.pdf. Accessed 1 Jul 2023.
13. Saedon H, Salleh S, Balakrishnan A, et al. The role of feedback in improving the effectiveness of workplace based assessments: a systematic review. BMC Med Educ 2012;12:25.
14. Carless D, Boud D. The development of student feedback literacy: enabling uptake of feedback. Assess Eva Higher Educ 2018;43:1315-25.
15. Gingerich A, Kogan J, Yeates P, Govaerts M, Holmboe E. Seeing the 'black box' differently: assessor cognition from three research perspectives. Med Educ 2014;48:1055-68.
16. van der Vleuten CP, Schuwirth LW. Assessing professional competence: from methods to programmes. Med Educ 2005;39:309-17.
17. Damschroder LJ, Reardon CM, Lowery JC. The Consolidated Framework for Implementation Research (CFIR). In: Handbook on Implementation Science. Cheltenham: Edward Elgar Publishing; 2020: 88-113.
18. Anderson HL, Kurtz J, West DC. Implementation and use of workplace-based assessment in clinical learning environments: a scoping review. Acad Med 2021;96:S164-74.
19. Massie J, Ali JM. Workplace-based assessment: a review

of user perceptions and strategies to address the identified shortcomings. Adv Health Sci Educ Theory Pract 2016;21:455-73.
20. Lörwald AC, Lahner FM, Greif R, Berendonk C, Norcini J, Huwendiek S. Factors influencing the educational impact of Mini-CEX and DOPS: a qualitative synthesis. Med Teach 2018;40:414-20.
21. Lörwald AC, Lahner FM, Mooser B, et al. Influences on the implementation of Mini-CEX and DOPS for postgraduate medical trainees' learning: a grounded theory study. Med Teach 2019;41:448-56.
22. Young JQ, Sugarman R, Schwarz J, O'Sullivan PS. Faculty and resident engagement with a workplace-based assessment tool: use of implementation science to explore enablers and barriers. Acad Med 2020;95:1937-44.
23. So HY, Li PK, Lai PB, et al. Hong Kong Academy of Medicine position paper on postgraduate medical education 2023. Hong Kong Med J 2023;29:448-52.
24. Norcini J, Anderson MD, Bollela V, et al. 2018 consensus framework for good assessment. Med Teach 2018;40:1102-9.
25. Henneman S, de Jong LH, Dawson LJ, et al. Ottawa 2020 consensus statement for programmatic assessment—1. Agreement on the principles. Med Teach 2021;43:1139-48.