Evaluating workplace-based assessment of interns in a Queensland hospital: does the current instrument fit the purpose?

TO THE EDITOR: An evaluation of 3390 prevocational progress assessment forms in New South Wales suggested that the assessment instrument appears unable to detect underperforming doctors, and may not aid their professional development.1 Similarly, Queensland interns undertake five terms per year and, in each of these, they are assessed against 11 items in three domains: clinical competence, communication, and personal and professional skills. The purpose of the assessment is to identify underperformers and subsequently to help improve their performance. However, there is limited literature on the psychometric properties of the instruments used. Our study aimed to evaluate the construct validity and reliability of the instrument as a measure of interns’ performance.

We performed principal component analysis (PCA) of the assessment items from all 72 interns who completed the five assessment forms during their intern year of the University of Queensland (2005 to 2007), working in one teaching hospital. A reliability test was performed using generalisability theory, which measured the variation in scores due to differences in performance between interns. It is generally expressed as a G coefficient ranging from 0 to 1 (no reliability to perfect reliability).

Supervisors assess interns using the rating scale: “requires substantial assistance”; “requires further development”; “consistent with level of appointment”; “performance better than expected”; “performance exceptional”; or “not applicable or not observed”. Supervisors’ ratings were coded from 1 to 5 for the data analysis.

Our examination of the assessment forms found that the mean of the skill items ranged from 3.69 to 4.16 (SD, 0.31–0.41). PCA identified only one factor from the 11 assessed items, and it explained 71% of variance in scores. The loadings of the coefficients ranged from 0.80 to 0.93 (Box) and the G coefficient was 0.16.

These data show that the instrument in this setting only measures a single global work performance factor, has low reliability, does not discriminate well between interns, and does not differentiate well between the different domains and assessment items, raising the question as to whether it is fit for purpose. The assessment program should be able to assess all these required attributes of junior doctors as outlined in the Australian Curriculum Framework for Junior Doctors in order to work safely in Australian hospitals and other health care settings.2,3 Other assessment methods involving direct observation, such as the mini clinical evaluation exercise4 and direct observation of procedural skills,5 could be integrated into the assessment to enhance its validity, but would require trained assessors and substantial time commitment for development and delivery.

Jianzhen Jenny Zhang Research Fellow
David Wilkinson Dean of Medicine and Head
Malcolm H Parker Associate Professor of Medical Ethics
Andrew Leggett Senior Lecturer
Jill Thistlithwaite Professor of Medical Education
School of Medicine, University of Queensland,
Brisbane, QLD.
jenny.zhang@uq.edu.au

Acknowledgements: We thank staff from the Medical Education Unit at Princess Alexandra Hospital for their support and advice.

Competing interests: No relevant disclosures.
doi: 10.5694/mja11.11369