The mini clinical evaluation exercise (mini-CEX) for assessing clinical performance of international medical graduates

Balakrishnan R Nair, Heather G Alexander, Barry P McGrath, Mulavana S Parvathy, Eve C Kilsby, Johannes Wenzel, Ian B Frank, George S Pachev and Gordon G Page

Assessing the performance of junior doctors in the workplace is important but challenging. The optimum assessment is by direct observation of doctors' interactions with patients and comprises multiple assessments by multiple examiners on a variety of patient problems. Clinical supervisors are best suited to observe and certify trainees, but often do not observe them directly.[1] Performance assessment is not done well in most instances, as it requires multiple sampling over time.[2] In-training assessments done at the end of a term introduce a "halo effect".[3]

Most of these problems can be overcome by the mini clinical evaluation exercise (mini-CEX), developed by the American Board of Internal Medicine.[4] The mini-CEX involves direct observation of a trainee in a focused clinical encounter, followed by immediate feedback. The assessment is recorded on a rating form that has been shown to have high internal consistency and reliability among internal medicine trainees, giving scores comparable with a high-stake clinical examination.[5,6] The mini-CEX has higher fidelity than other formats.[7]

International medical graduates (IMGs) comprise about 25% of the medical workforce in developed countries.[8] Their certification for registration is a major task of the medical boards and registration authorities in Australia and other countries.[9] The Australian Medical Council (AMC) has conducted clinical examinations to assess IMGs since 1978.[10] Successful candidates undertake 12 months of supervised practice before obtaining full registration. Despite having passed the current AMC clinical examination, IMGs' competence and performance in the workplace have been criticised.[11]

We evaluated the feasibility, reliability and acceptability of the mini-CEX as a performance assessment tool for IMGs in the Australian workplace.

METHODS

The study was conducted in three large metropolitan teaching hospitals in Australia, one each in New South Wales, Queensland and Victoria, as part of a larger international

ABSTRACT

Objective: To evaluate the feasibility, reliability and acceptability of the mini clinical evaluation exercise (mini-CEX) for performance assessment among international medical graduates (IMGs).

Design, setting and participants: Observational study of 209 patient encounters involving 28 IMGs and 35 examiners at three metropolitan teaching hospitals in New South Wales, Victoria and Queensland, September–December 2006.

Main outcome measures: The reliability of the mini-CEX was estimated using generalisability (G) analysis, and its acceptability was evaluated by a written survey of the examiners and IMGs.

Results: The G coefficient for eight encounters was 0.88, suggesting that the reliability of the mini-CEX was 0.90 for 10 encounters. Almost half of the IMGs (7/16) and most examiners (14/18) were satisfied with the mini-CEX as a learning tool. Most of the IMGs and examiners enjoyed the immediate feedback, which is a strong component of the tool.

Conclusion: The mini-CEX is a reliable tool for performance assessment of IMGs, and is acceptable to and well received by both learners and supervisors.

MJA 2008; 189: 159–161

collaborative study with the Medical Council of Canada. The ethics committee in each centre approved the study.

Participants

All IMGs at the participating hospitals who had passed the AMC clinical examination in the previous 12 months and 50 potential examiners were asked to volunteer for the study. All IMGs gave written, informed consent.

In each centre, potential examiners were requested to attend a training session. A coordinator assisted the study team at each site. Patients were inpatients or outpatients of the participating hospitals.

Mini-CEX rating form

The following skills were rated: medical interviewing, physical examination, professionalism/humanistic qualities, counselling, clinical judgement, organisation/efficiency and overall clinical competence. Ratings were on a nine-point scale, where 1–3 signified unsatisfactory performance; 4–6, satisfactory performance; and 7–9, superior performance at a mid-postgraduate year 1 level. Examiners were also asked to grade the encounters as "met expectations", "borderline" or "did not meet expectations".

A short evaluation survey investigated any difficulties with the process and perceptions about the validity of the mini-CEX using Likert-scale questions and narrative comments.

Procedure

The examination comprised four assessments in emergency medicine (two examination, one history taking and one counselling), three in medicine (one history taking, one management and one counselling), and three in surgery (two examination and one management). These three specialties were selected because these terms are compulsory for internship in most Australian states and territories.

The IMGs and examiners were asked to schedule a mutually convenient time, with at least 30 minutes of protected time for assessment and immediate feedback. At the end of the study period, the IMGs and examiners were asked to evaluate the process.

Statistical analysis

Reliability was assessed using generalisability theory analysis. Generalisability analyses (G studies) allow estimation of the variance components associated with the different examination conditions (eg, types of tasks,

1 Variance components estimates for the generalisability study

Facets	Variance component estimate	Proportion of total variance	df
IMGs (differences between IMGs across encounters)	0.88	47%	27
Encounters nested within IMGs (differences in encounter sets between participants)	1.00	53%	157

IMG = international medical graduate. ◆

number of tasks, and number of markers). The ratio of the variance component for the object of measurement (in this case, differences between IMGs) to the total error variance yields an estimate of reliability: the generalisability (G) coefficient, with values ranging from 0 to 1. The effect of changes in the examination conditions (eg, increase in the number of tasks to be performed by the IMG, or change in the number of markers for each task) can be modelled to inform decisions on optimising the measurement (decision [D] studies).

We used a G study followed by a D study to evaluate the reliability of the measurement, and sought to determine the number of observed clinical encounters needed to attain a G coefficient of 0.90. This value was set by taking into account the high stakes of the assessment for the IMGs. The average rating for the different assessments within each encounter (eg, history taking, examination) was the outcome measure for these analyses. As each IMG interacted with different patients and a different number of patients, the "patient-encounters" factor was treated as nested within IMGs in the G-study design.

We used urGENOVA statistical software (Robert Brennan, Center for Advanced Studies in Measurement and Assessment, University of Iowa, Iowa, USA) for the analyses.

RESULTS

All 28 IMGs who had passed the AMC examinations within the previous 12 months and 35 examiners volunteered to participate in the study. Twenty-two examiners were trained in assessing the mini-CEX; the remaining 13 examiners participated without training. The examiners included specialists and specialist trainees in internal medicine, surgery and emergency medicine.

Mini-CEX ratings

The 28 IMGs were assessed by the 35 examiners on 209 clinical encounters: 122 assessments were done in wards, 70 in emergency departments, eight in intensive care units, six in outpatient clinics and two in offices; location was not recorded for one encounter. The mean number of mini-CEXs completed by IMGs and examiners was 7.2 (range, 2–13) and 6.0 (range, 1–20), respectively. Assessments were scored as "met expectations" for 150 encounters, "borderline" for 40, and 19 (9% across 12 IMGs) "did not meet expectations". The average mini-CEX duration was 20 minutes (range, 6–45 minutes). The average time for feedback was 12 minutes (range, 3–20 minutes). Complexity of encounters was rated by examiners as low for 19, moderate for 150 and high for 31 encounters; data were missing for nine encounters.

Because of differences in the number of encounters per participant, we included a maximum of eight encounters in the generalisability study. The results of the variance components estimation are shown in Box 1. The G coefficient for eight encounters was 0.88. As a measure of discrimination, the standard error of measurement for the measurement design with eight encounters was estimated at 0.35 (that is, 19/20 times, the "true" score of an IMG will fall within ±0.69 of an observed score). The results of the D study indicated that 10 encounters were necessary to achieve a reliability of 0.90 (Box 2).

Evaluation survey

In the evaluation survey, 16/28 IMGs (57%) and 18/35 examiners (51%) responded. Most respondents (10 IMGs; 15 examiners) never or only occasionally experienced difficulty arranging the mini-CEX encounters. When problems did occur, they were often due to rostering issues and patients being away from the wards.

Twelve of the 16 IMG respondents received feedback after each mini-CEX encounter, and nine of these were satisfied with this feedback. There was no response from examiners for the remaining four. Seven of 15 examiner respondents indicated they would like further training in giving feedback. The comments indicated that the feedback is a strong feature of the mini-CEX.
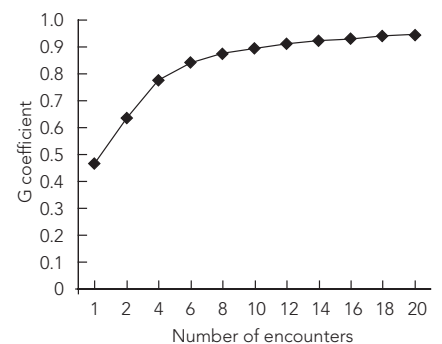
Almost half of the IMGs (7/16) and most examiners (14/18) were satisfied or very satisfied with the mini-CEX as a tool for learning. Ten of the examiner respondents were satisfied or very satisfied with the mini-CEX as an assessment tool and eight were neutral. Most respondents (8/16 IMGs; 15/18 examiners) were positive about the mini-CEX continuing as an assessment tool.

DISCUSSION

Under the conditions and settings used, the mini-CEX reliably assessed the clinical performance of IMGs with eight to 10 encounters. This is consistent with the results of other studies.[7] As the mini-CEX is conducted within the workplace with real patients, it has high fidelity and it is acceptable to both IMGs and examiners. A fail rate of 9% (19/209 encounters) across 12 IMGs is concerning, given these IMGs had passed the AMC clinical examination.

As we have demonstrated, the mini-CEX appears to be a reliable and acceptable assessment tool of clinicians in the workplace, and is a valuable method of identifying which candidates may have problems in a clinical situation. Most IMGs were satisfied with their feedback. Examiners reported that this was the most important part of the mini-CEX. However, they would have preferred more training in this task. Overall, examiners felt this tool assesses clinical performance better than the conventional methods.

2 Generalisability (G) coefficient as a function of the number of encounters*

* The G coefficient was determined for a maximum of eight actual encounters per participant. The effect of further encounters was then modelled in a D study. ◆

There are limitations to this study. Only patients from medical, surgery and emergency departments were included, and the results cannot automatically be extended to other specialties. The IMGs were a self-selected group, introducing a possible bias, and only a small group of experienced clinicians took part, limiting the generalisability of the study.

Further analyses will need to consider the effect of examiner training, and the relationship between performance on the mini-CEX with that in other assessments.

The mini-CEX is a feasible, reliable and high-fidelity instrument for workplace-based assessment of IMGs with the strong advantage of providing ongoing observation and feedback. It has the potential to be used for summative assessment of IMGs and other medical trainees. The AMC, in collaboration with some licensing bodies, has already introduced the mini-CEX as a workplace assessment tool for some IMGs.

COMPETING INTERESTS

None identified.

AUTHOR DETAILS

Balakrishnan R Nair, FRACP, FRCP, Conjoint Professor of Medicine,[1] and Director, Centre for Medical Professional Development[2]
Heather G Alexander, PhD, Associate Director[3]
Barry P McGrath, MD, FRACP, Professor of Vascular Medicine[4]
Mulavana S Parvathy, FRACGP, International Medical Graduate Supervisor[2]
Eve C Kilsby, BA(Psych), Project Officer[3]
Johannes Wenzel, MD, FACEM, Director of Clinical Teaching[5]
Ian B Frank, BA(Hons), Chief Executive Officer[6]
George S Pachev, PhD, Head, Assessment Unit, Faculty of Medicine[7]
Gordon G Page, PhD, Professor Emeritus[7]

1 University of Newcastle, Newcastle, NSW.
2 John Hunter Hospital, Newcastle, NSW.
3 Griffith Institute for Higher Education, Griffith University, Brisbane, QLD.
4 Monash University, Melbourne, VIC.
5 Southern Health, Melbourne, VIC.
6 Australian Medical Council, Canberra, ACT.
7 University of British Columbia, Vancouver, British Columbia, Canada.

Correspondence: kichu.nair@newcastle.edu.au

REFERENCES

1 Holmboe ES. Faculty and the observation of trainees' clinical skills: problems and opportunities. *Acad Med* 2004; 79: 16-22.
2 Williams RG, Dunnington GL, Klamen DL. Forecasting residents' performance — partly cloudy. *Acad Med* 2005; 80: 415-422.
3 Wilkinson TJ, Wade WB. Problems with using a supervisor's report as a form of summative assessment. *Postgrad Med J* 2007; 83: 504-506.
4 Norcini JJ, Blank LL, Arnold GK, Kimball HR. The mini-CEX (clinical evaluation exercise): a preliminary investigation. *Ann Intern Med* 1995; 123: 795-799.
5 Durning SJ, Cation LJ, Markert RJ, Pahgaro LN. Assessing the reliability and validity of the mini-clinical evaluation exercise for internal medicine residency training. *Acad Med* 2002; 77: 900-904.
6 Hatala R, Ainslie M, Kassen BO, et al. Assessing the mini-Clinical Evaluation Exercise in comparison to a national specialty examination. *Med Educ* 2006; 40: 950-956.
7 Norcini JJ, Blank LL, Duffy FD, Fortna GS. The mini-CEX: a method for assessing clinical skills. *Ann Intern Med* 2003; 138: 476-481.
8 Mullan F. The metrics of the physician brain drain. *N Engl J Med* 2005; 353: 1810-1818.
9 McGrath BP. Integration of overseas medical graduates into the Australian medical workforce. *Med J Aust* 2004; 181: 640-642.
10 Breen K, Frank I, Walters T. Australian Medical Council: the view from inside. *Intern Med J* 2001; 31: 243-248.
11 Van Der Weyden MB, Chew M. Arriving in Australia: overseas trained doctors [editorial]. *Med J Aust* 2004; 181: 633-634.