

The selection interview as a predictor of later performance

Anne Eastwood IGPE

Abstract

Background:

General Practice Education and Training (GPET) encourages Regional Training Providers (RTPs) to assess the performance of all registrars at the commencement of their training to determine their particular needs. It would be useful to RTPs, if an existing assessment tool could be used as part of the evaluation of the learning needs of new registrars.

Aim: To determine whether there is any relationship between selection interview scores and assessments of performance made during the IGPE program.

Method: A retrospective cohort study was undertaken as a quality assurance project, using existing data from 98 registrars from the 2004 to 2007 cohorts. Selection interview scores, both individual and global were compared with a number of outcomes from clinical teaching visits and written assessment activities. The data were analysed using SPSS version 14.

Results: Although there were correlations between some interview and first clinical teaching visit (CTV) scores, it was not possible to identify the lowest CTV achievers from the interview scores. Registrars with low total interview scores were more likely to achieve low assessment activity results.

Conclusions: Interview scores do not reliably predict which registrars will perform poorly in the basic term. There may be some value in adding them to current performance measures when deciding which registrars need extra assistance.

Introduction

In 2008, for the first time, The Institute of General Practice Education (IGPE) did not fill its intake quota. Over recent years, there has been a fall in the ratio of applicant numbers to available first year places¹ in general practice (GP) training in Australia. The ACIL Tasman Evaluation of Regionalisation of GP Vocational Training¹ raised the possibility that, despite minimum eligibility standards, a higher number of less able candidates are being selected and that this may lead to reduced registrar retention, an increased need for remediation activity and may ultimately influence RACGP exam pass rates.

General Practice Education and Training (GPET) encourages the assessment of performance of all registrars at the commencement of their training to determine their particular needs². If the concerns expressed in the ACIL Tasman report are well founded, the number of registrars for whom the early identification of problems is vital, may be increasing. There are, however, many demands on medical educator time at IGPE and at other regional training providers (RTPs)). It would be helpful, therefore, if an existing assessment (the selection interview) could be used to help identify particular problems and learning needs.

This study aims to determine whether there is any relationship between selection interview scores and assessments of performance made during the early part of the IGPE program.

Background

The interview as a selection tool

There is on-going debate about the usefulness of selection interviews in many areas, including medical education. There are few studies of what works in medical school selection³. A Medline search found only two (UK) studies relating selection scores to outcomes in vocational general practice programs. Concerns have been raised about the subjectivity of the interview process, the validity and reliability interview as a tool for measuring particular attributes⁴ and the observation that some important attributes are not measured. Looi⁵, for example, raised the question of including a measure of empathy in the selection of medical students. Empathy is an important attribute for a GP. Wilkinson and Zhang⁶ found that undergraduate grade point average was a better predictor of academic performance in a Queensland graduate entry medical course than interview scores. However the undergraduate grade effect waned over the duration of the program and academic ability is only one of the qualities required of a good GP.

In the UK selection is based on a written application, two interviews and a role-play. End of training assessment is based on a multiple choice exam, audit and videotaped consultations. Davison et al⁷ studied data from 64 West Midland GP registrars and found that a small correlation between interview scores and the multiple choice exam and moderate correlation with video assessment scores. Patterson et al⁸ targeted 6 competencies (empathy, communication skills, clinical expertise, problem solving, professionalism and coping with pressure). They used simulated consultations, a group exercise, a written exercise and interviews instead of the usual interviews and role play. This selection method predicted supervisor ratings after 3 months better than the standard UK method.

The General Practice Training Selection Interview in Australia

A standardised interview format is used throughout Australia. The areas targeted are; communication skills; clinical competence; analytical and problem solving skills; organisational skills, motivation and sense of vocation; personal attributes and professional and ethical attributes. One question or scenario is posed for each area and responses are graded from 1 (unsuitable and does not meet criterion) to 7 (meets criterion to a superior extent). The Behaviourally Anchored Rating Scale (BARS) provides interviewers with descriptions of the type of skills, attributes, abilities and knowledge which are indicative of a performance at a particular level.

At the 2007 GPET conference, Emblen⁹ pointed out the potential strengths and shortcomings of the Australian general practice training selection interview. While a question from each area of interest can give insight into a registrar's abilities and qualities, interviews require different skills from general practice and knowing the right answer does not necessarily result in its application in practice.

Identifying IGPE registrars who may need extra help

Currently the selection interview is not used as a source of information. The decision to offer registrars extra help is based on an overall picture gained from the areas shown in Table 1.

Table 1

Sources of information used to identify registrars needing extra help

- 1. Hospital program GP consultations**
- 2. Clinical Teaching Visit (CTV) reports**
- 3. Assessment activity results**
- 4. Direct observation teaching sessions (DOTS)**
- 5. Formal and informal feedback from supervisors**
- 6. ME feedback from individual contacts and observation at training days.**

Assessment activities are quizzes on the topic of the previous teaching session and a percentage mark is given. Direct observation teaching sessions (DOTS) are conducted by supervisors on a number of single consultations and are similar in structure to CTVs. CTVs are assessed on communication skills, history taking, physical examination, problem definition, management and investigations, organisational skill, professional and ethical behaviour and overall impression. They are graded from poor to excellent.

Method

A retrospective cohort study was undertaken as a quality assurance project. Data from the 2004 to 2007 cohorts was used. Before 2004 a different selection process was used. There were some differences between the 2004 interviews and subsequent years but these could be corrected for in the data set. Data from registrars who had transferred to IGPE from another RTP was excluded. There was insufficient data from registrars who had not completed their basic term at the start of the first term of 2008 and they were also excluded. Data was not available for all registrars for all outcomes. Registrars who did their basic term in a rural practice did not do IGPE Assessment Activities and CTVs were reported differently

IGPE maintains its own Access database (ROD), which includes interview results and the results of various assessments. ROD was not designed with program evaluation in mind, and over time, data entry has been inconsistent. Consequently much of the data had to be manually collected and entered into an Excel spreadsheet, which was then converted to SPSSv14. All IGPE registrars who have attempted the exam have passed, all but two on the first attempt. Results are recorded as pass or fail, so it was not possible to use exam results as an outcome. There was insufficient data available from supervisor assessments to allow meaningful analysis. The DOTs program has only started recently and to date the hospital program has not identified any registrars of serious concern. This left withdrawal from training, assessment activity scores and CTV scores as the outcomes available for comparison with interview scores. Since the aim of the study is to allow early intervention, only basic CTV scores were considered.

Given the ordinal nature of most of the data, Kendall's Tau b was used to examine trends in the relationship between interview and

CTV scores. Low interview scores were defined as 4 or under (there was insufficient data to use 3 as the cut off point) and low CTV score were defined as borderline or less. Bar graphs were drawn where there were significant relationships. Chi-square testing was used to examine relationships between low scores and also between overall and motivation interview scores and withdrawal from the program. Total interview scores are in practice regarded as continuous data and assessment activity and mean referee scores are continuous, so linear regressions were performed on these factors.

Results

Data from 96 registrars met the inclusion criteria. Assessment Activity data was available for 44 registrars and CTV data for 58.

Referee vs interview scores

Average referee scores had a small correlation with total interview scores on linear regression ($R=0.276$, $p=0.007$).

Withdrawal

14 registrars who accepted a place on the IGPE program withdrew from training, in all cases before undertaking a basic term. There was no relationship between withdrawal and average referee score, age at entry, total interview score or interview motivation score on Chi-Square testing.

Assessment Activities

Mean Assessment Activity scores were moderately correlated with total interview scores on linear regression ($R=0.391$, $p=0.009$). On scatter graphs there appeared to be no relationship between assessment activity and interview clinical competence or problem solving scores. Less than 10% of registrars with total interview scores over 30 had mean assessment scores of <65%, compared with 35% of those with scores >30 ($X^2=7.2$, $df=1$, $p=.007$). The relationship between low assessment activity and low interview clinical competence scores was not significant ($X^2=1.9$, $df=1$, $p=.164$)

Clinical teaching visits

The scores for the two CTV visits were compared to give an indication of test reliability. Overall 15% of first (CTV1) basic CTV scores were higher than the second (CTV2), 44% were lower and 41% were the same. The results for comparisons between interview and CTV scores where there was at least one significant relationship are shown in Table 2

Table 2
Comparing interview and CTV scores

	Kendalls Tau b correlations (all scores)		Chi square or Fisher's exact test (comparing low interview and low ctv scores) (df=1)	
Comparing	CTV1	CTV2	CTV1	CTV2
I(communication) and CTV (communication)	T _b =.231 (p=.053)	T _b =.376 (p=.002)*	(p=.271)	(p=1.00)
I(clinical comp) and CTV (physical exam)	T _b =.238 (p=.042)*	T _b = -.025 (p=.842)	(p=.706)	(p=.414)
I (clinical comp) and CTV management	T _b =.295 (p=.013)*	T _b = -.045 (p=.713)	X ² = .464 (p=.496)	(p=.070)
I (problem solving) and CTV management	T _b =.236 (p=.046)*	T _b = 0.178 (p=0.143)	X ² = .464 (p=.496)	(p=.640)
I (org) and CTV (org)	T _b =.388 (p=.001)*	T _b =.011 (p=.931)	(p=.098)	(p=.103)
I (attributes) and CTV management	T _b =.199 (p=.095)	T _b =.254 (p=0.034)*	(p=.515)	(p=1.00)
I (attributes) vs CTV overall	T _b =0.319 (p=.007)	T _b =0.293 (p=.016)	(p=.683)	(p=1.00)
I (professional) vs ctv (professional)	T _b =.351 (p=.004)*	T _b = .064 (p=.600)	(p=.229)	(p=.229)

I=Interview.

Where no test is shown the probability is for Fisher's exact test.

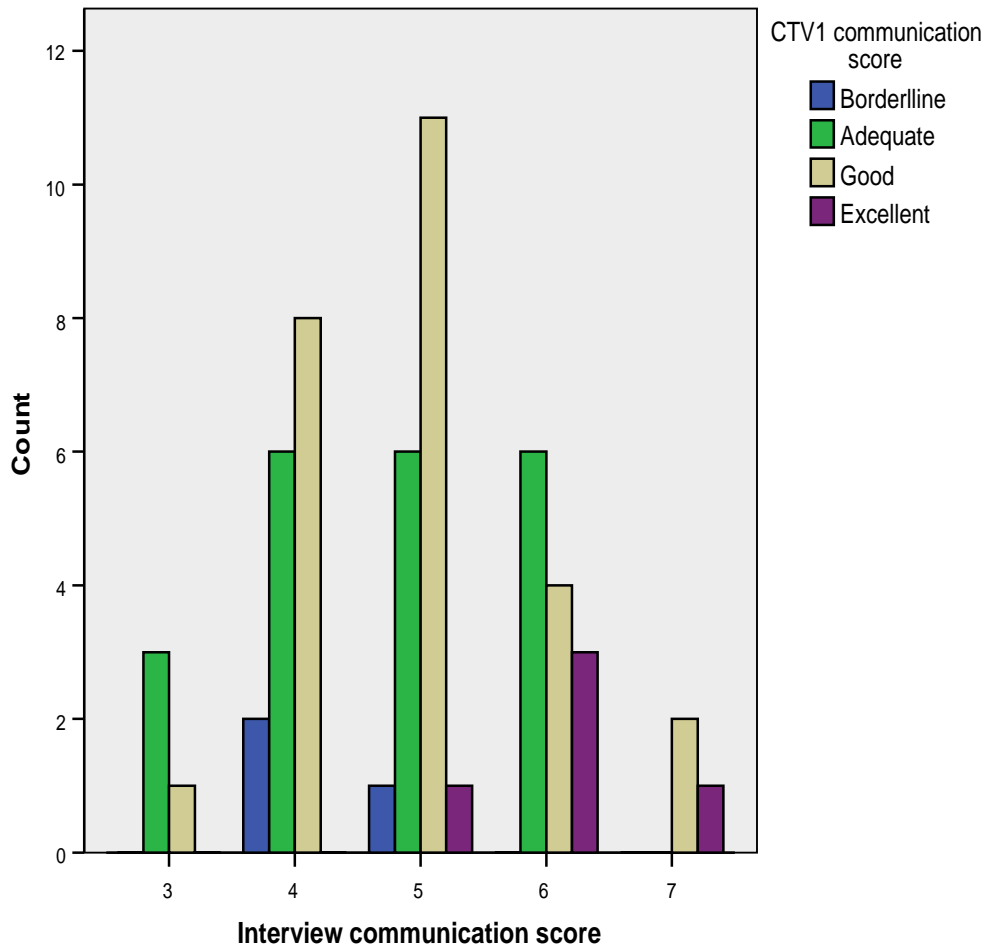


Figure 1

CTV communication score grouped by interview communication score

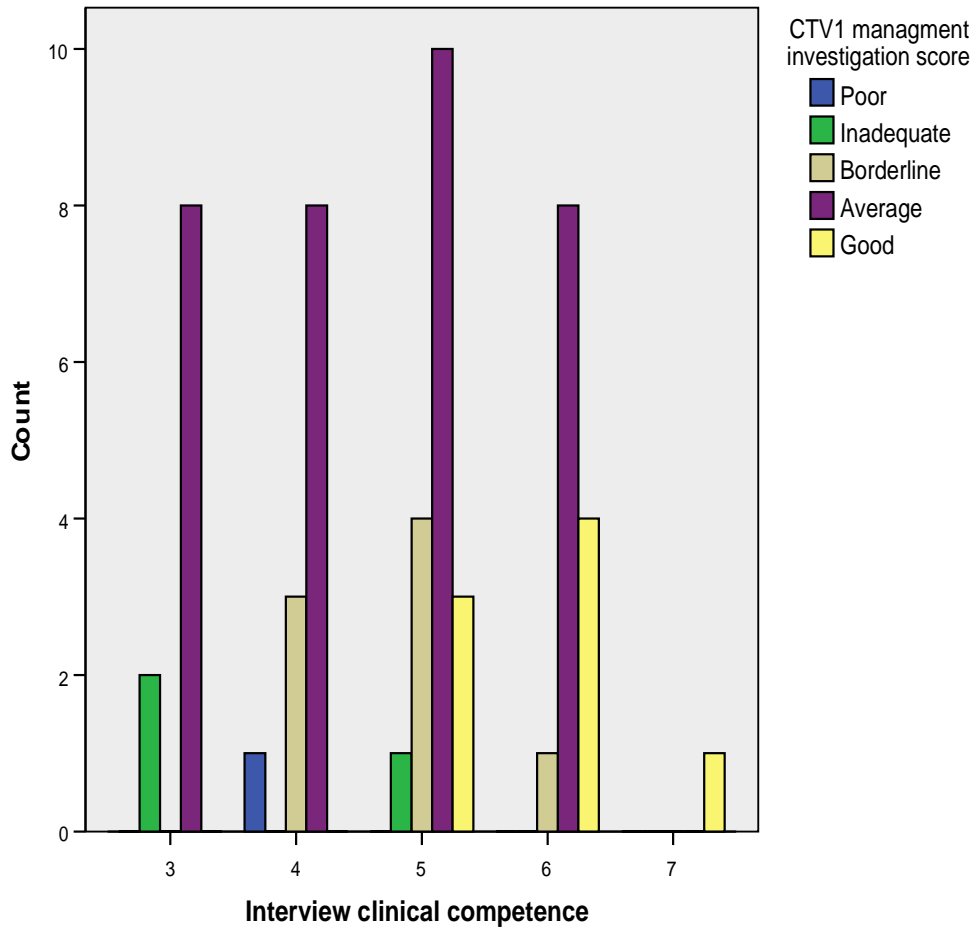


Figure 2

CTV management score grouped by interview clinical competence score

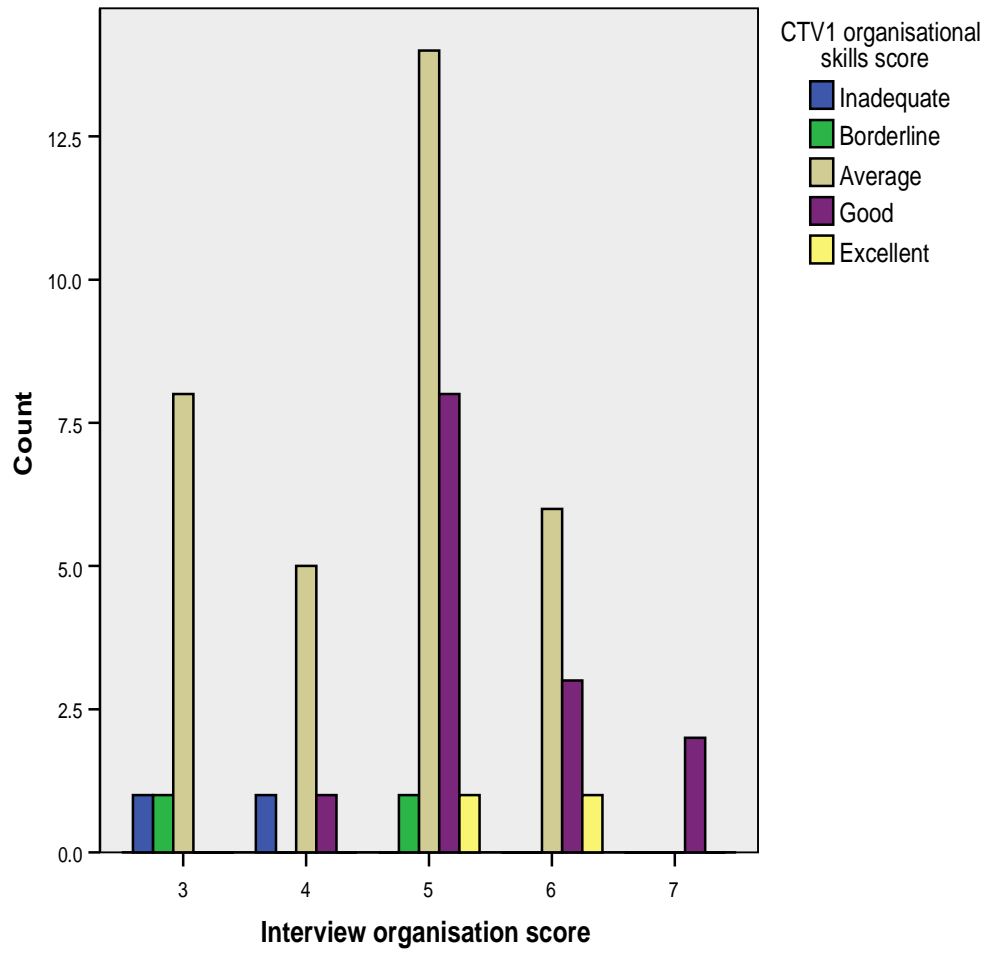


Figure 3

CTV organisational skills score grouped by interview organisational skills score

Discussion

Although 96 of a total 117 registrars from the 4 cohorts met the inclusion criteria, complete data was available for just over half of these. It is possible that more significant results would have resulted from a larger sample. The size of available samples is a problem for most evaluations within RTPs and the variability of RTP programs restricts the topics for which data can be aggregated between RTPs.

The small correlation between interview and referee scores is consistent with the findings of Salvatori's review⁴. She also found that referee reports have poor reliability and predicative validity.

There is a wide range of reasons for registrar withdrawal, from changes in person circumstances to better incomes elsewhere, and this probably explains the inability to predict withdrawal.

There was a moderate level of correlation ($r = .391$, $p = .009$) between total interview score and mean activity assessment score. There was also a significant relationship between low total interview scores and assessment activity results, although 65% of those with low interview scores did not have low assessment activity results. Assessment activities test knowledge and not the application of knowledge, skills and attitudes, which are also required for clinical competence. Nevertheless, in a small sample, the significance level of the test raises the possibility that there is a real relationship between low interview clinical competence and assessment scores.

There are significant small to moderate correlations between interview and CTV1 scores of communication skills, organisational skills and professionalism scores, as well as between interview clinical competence and CTV1 physical examination and

management scores, interview problem solving and CTV management scores and interview personal attributes and CTV overall scores. These relationships, however, mostly reduce or disappear by CTV2. This may be due to the effects of training. The aim of the study was to determine whether low interview scores could help identify registrars in need of extra assistance. With the possible exception of organisational skills, it does not appear that this is the case. Given the fairly small sample size, the relationship between interview and CTV organisational skills may be practically significant. The bar graphs (Figs 1-3) indicate that the interview is better at detecting high rather than low performing registrars. This has also been observed in other studies^{4,9}. It could reasonably be argued that statistical significance does not need to be achieved for a relationship between two factors to be useful in practice. Given the amount of time and effort, however, which goes in to helping under-performing registrars, one would want better predictive values than those achieved in this study.

Conclusions

There is no evidence that low interview scores reliably predict which registrars will later be identified as needing extra intervention at IGPE. It may be useful to add interview scores (particularly clinical competence and organisational scores) to the list assessments, which are combined to identify registrars needing extra help from medical educators.

References.

1. ACIL Tasman. *Evaluation of Regionalisation of GP Vocational Training* Chapter Four- Participation in the General Practice Training program pp 43-63 2005. Accessed online 14/5/08 at <http://www.health.gov.au/internet/wcms/publishing.nsf/Content/pcd-regionalisation-evaluation>
2. AGPT. *AGPT Policies 2006* Section 4.2.2.1. Accessed online 14/4/08 at <http://www.agpt.com.au/PoliciesPublications/RecentPolicyChanges/>
3. McManus C and Powis D. *Testing medical school selection tests*. MJA 2007;186(3);118-9.
4. Salvatori, P *Reliability and Validity of Admissions Tools Used to Select Students for the Health Professions*. Advances in Health Sciences Education 2001;6:159-175.
5. Looi J. *Empathy and competence* MJA 2008;188(7):414-416
6. Wilkinson D, Zhang J, Byrne G et al *Medical School Selection criteria and the prediction of academic performance* MJA 2008;188:349-54
7. Davison I, Burke S, Bedward J, Kelly S. *Do selection scores for general practice registrars correlate with end of training assessments?* Education for Primary Care 2006;17:473-8.
8. Patterson F, Fergusson E, Norfolk T and Lane P. *A new selections system to recruit general practice registrars: preliminary findings from a validation study*. BMJ 2005;330:711-3.

9. Powis D. *Selecting medical students*. MJA 2008;188(6):323-4
10. Emblen G, McKain P, Garside,S. *Your registrar selection interviews:- An extra tool in detecting the registrar is going to have problems? What to do with the information?* GPET Conference Proceedings 2007- Presentations (On CD).