Evaluation of recruitment and selection for specialty training in public health: Interim results of a prospective cohort study to measure the predictive validity of the selection process.

Pashayan N, Gray S, Duff C, Parkes J, Williams D, Patterson F, Koczwara A, Fisher G, Mason BW.

Dr Nora Pashayan, Senior Clinical Lecturer in Applied Health Research, Cancer Research UK Clinician Scientist Fellow, University College London, Department of Applied Health Research, 1-19 Torrington Place, London, WC1E 7HB.

Professor Selena Gray, Professor of Public Health, University of the West of England, Bristol, Frenchay Campus, Coldharbour Lane, Bristol, BS16 1DD.

Dr Celia Duff, Training Programme Director, Public Health East of England, 2-4 Victoria House, Capital Park, Fulbourn, Cambridge CB21 5XB.

Dr Julie Parkes, Associate Professor in Public Health Medicine, Public Health Sciences & Medical Statistics, University of Southampton, C floor, South Academic Block, Southampton General Hospital, Southampton, SO16 6YD.

Dr David Williams, Faculty of Public Health, 4 St Andrews Place, London, NW1 4LB.

Professor Fiona Patterson, Professor of Organisational Psychology, Department of Psychology, University of Cambridge, Downing Street, Cambridge, CB2 3EB.

Dr Anna Koczwara, Associate Director, Work Psychology Group, 27 Brunel Parkway, Pride Park, Derby, DE24 8HR.

Grant Fisher, Examinations and Projects Manager, Faculty of Public Health, 4 St Andrews Place, London, NW1 4LB.

Dr Brendan W Mason, Consultant Epidemiologist, Public Health Wales, Communicable Disease Surveillance Centre (CDSC), The Temple of Peace and Health, Cathays Park, Cardiff, CF10 3NW.

Correspondence to: Dr Brendan Mason. Brendan.mason@wales.nhs.uk.

Abstract

Background

The recruitment process for public health specialty training includes an Assessment Centre (AC) with three components, Rust Advanced Numerical Reasoning Appraisal (RANRA), Watson-Glaser Critical Thinking Appraisal (WGCT), and a Situation Judgement Test (SJT), which determines invitation to a Selection Centre (SC). The scores are combined into a total recruitment (TR) score which determines the offers of appointment.

Methods

A prospective cohort study using anonymous record linkage to investigate the association between applicant's scores in the recruitment process and registrar's progress through training measured by results of Membership Faculty Public Health (MFPH) examinations and outcomes of the Annual Review of Competence Progression (ARCP).

Results

Higher scores in RANRA, WGCT, AC, SC, and TR were all significantly associated with higher adjusted odds of passing Part A MFPH exam at the first attempt. Higher scores in AC, SC, and TR were significantly associated with passing Part B exam at the first attempt. Higher scores in SJT, AC, and SC were significantly associated with satisfactory ARCP outcomes.

Conclusions

The current UK national recruitment and selection process for public health specialty training has good predictive validity. The individual components of the process are testing different skills and abilities and together they are providing additive value.

200 Words

Introduction

General Medical Council (GMC) standards for speciality training require that the processes for recruitment, selection and appointment must be open, fair, and effective [1]. Evaluation is required to demonstrate the effectiveness of the recruitment and selection process. One measure of effectiveness is the predictive validity of the selection process, that is, the extent to which the process predicts applicants' future performance on criterion of interest. The intention of selection and recruitment is to identify applicants who will successfully complete training and excel in subsequent practice.

The UK public health specialty training scheme is a five year training scheme leading to registration as a public health specialist with the GMC or UK Public Health Register (UKPHR), and meet the requirements to work in consultant level senior public health posts within the UK [2]. Historically recruitment was undertaken at local Deanery level, but in 2009 a national recruitment and selection process was introduced. This consists of a two stage competency-based process, explicitly linked to a detailed person specification. Detailed development work was undertaken mapping the key competences and attributes required in the person specification for the role as a consultant in public health, and a recruitment process designed to ensure that these were all systematically tested during the recruitment process.

The first stage, assessment centre (AC), is comprised of two cognitive ability tests, which measure numerical (Rust Advanced Numerical Reasoning Appraisal test) and verbal reasoning (Watson Glaser Critical Thinking test). A situational judgment test developed specifically for use in the public health context was added in 2011. Progression to the second stage, selection centre (SC), requires applicants to pass the threshold score for each of the three tests and those with the highest combined scores are invited to the selection centre. The SC has three components; a group exercise, a written test, and a series of short interview panels. Figure 1 shows a summary of the recruitment and selection process.

We report the interim results of a prospective cohort study set up to measure the predictive validity of the recruitment process as a whole, and its individual components, on public health specialty registrars progress through speciality training.

Methods

The evaluation used a prospective cohort study design, all specialty registrars who took up an appointment to a training scheme in England and Wales following recruitment in one of the four annual recruitment rounds between 2009 and 2012 were included in the study. Applicants were given information about the planned long- term evaluation at the time of recruitment. Follow up for this interim analysis was until December 2014.

Follow up was undertaken by anonymous record linkage. Health Education East Midlands (HEEM), the lead organisation for public health national recruitment, allocated each applicant a unique anonymous identifier. The applicant's basic demographic details and scores for each part of the recruitment process were provided to the evaluation team with this unique anonymous identifier. HEEM provided the Faculty of Public Health (FPH) with details of applicants who were offered a post and their unique anonymous identifier. The FPH provided the evaluation team with outcome data with the unique anonymous identifier.

The "exposure" or predictor variables for the cohort available from the recruitment process were: T-scores standardised to the public health norm group for Rust Advanced Numerical Reasoning Appraisal (RANRA) [3]; Watson-Glaser Critical Thinking Appraisal (WGCT) [4]; and the T-score for a bespoke Situation Judgement Test (SJT) for Public Health for 2011 and 2012; and, the total overall score for the selection centre (SC score). A combined score from the assessment centre tests (AC score) and a total combined score for the whole recruitment process (TR score) was calculated.

The criterion measures or "outcomes" used to indicate progress through speciality training were: full pass at the first attempt after starting speciality training of the Membership of the Faculty of Public Health (MFPH) Part A examination; pass at the first attempt after starting speciality training of the MFPH Part B examination; and, a satisfactory outcome of the Annual Review of Competence Progression (ARCP) [5] defined as no recorded ARCP outcomes 2, 3, 4 or 5 (Supplementary Table S1).

Data on potential confounding factors, age, sex, ethnicity, and professional background (medical or other background) were obtained as part of the application process. The study size was determined by the number of registrars appointed to the training programmes and an a priori power calculation was performed. For a power of 80% at 5% significance level and the observed difference in proportions of candidates with higher AC score passing membership exam, a minimum sample size of 206 applicants are required.

Predictive validity

The predictor variables were categorised into below and above 50th percentile (median) groups. In bivariate analysis the association between demographic variables and the predictor and outcome variables was investigated.

The predictive validity of the standardised scores of each of the three AC tests (RANRA, WGCT, SJT), the AC score, SC score and TR score were examined. Logistic regression analyses were used to estimate the odds of passing Part A and Part B exam on the first attempt and having satisfactory ARCP outcomes. The predictor variables were analysed as both categorical and continuous with adjustment for potential confounders. The logistic regression analysis was not corrected for range restriction.

Receiver Operator Characteristic (ROC) analyses were used to study the ability of each of the components of the recruitment process separately and combined to discriminate between registrars likely to perform well in training, as demonstrated by passing membership examinations from the first attempt and having satisfactory outcomes of ARCPs. The discriminatory accuracy can be measured by the area under the receiver operator characteristic curve (AUC_{-ROC}). AUC_{-ROC} is the probability that a test correctly identifies an individual who will perform well in training from a pair of whom one will perform well and one will not. AUC_{-ROC} values range from 0.5 (total lack of discrimination) to 1.0 (perfect discrimination).

Individuals with missing data were excluded from any analysis which required the missing data. $P<0.05$ was considered statistically significant. Data analysis was done using STATA 13.

Results

The cohort was comprised of 274 registrars who took up an appointment after applying between 2009 and 2012. The number recruited in each consecutive year was 74, 76, 76 and 48. Demographic data were available for almost all applicants. Of these, the mean age (SD) was 33 (6) years; range: 25-50 years, 73% (197/271) were women, 86% (223/259) described themselves as White or White British and 37% (101/271) applied as registered medical practitioners.

Nineteen registrars had passed Part A MFPH and seven had passed Part B MFPH prior to taking up appointment. Eighty six percent (236/274) of registrars had sat Part A MFPH since joining a training scheme, of whom 58% (137/236) achieved a full pass at the first attempt. Sixty seven percent (188/274) of registrars had sat Part B MFPH since joining a training scheme, of whom 90% (169/188) passed at the first attempt. Overall 212 registrars had at least one ARCP record, of whom 84% (179/212) had satisfactory ARCP outcomes recorded.

The cohort's scores in the component and combined components of the recruitment process, the predictor variables, are summarised in supplementary Table S2. The association between demographic characteristics and the predictor and outcome variables are summarised in supplementary Table S3.

Higher scores in Rust Advanced Numerical Reasoning Appraisal (RANRA), Watson-Glaser Critical Thinking Appraisal (WGCT), Assessment Center (AC), Selection Center (SC), and Total Recruitment (TR) were all significantly associated with higher odds of passing Part A exam at the first attempt. There was 12% increase in odds of passing Part A exam for every one point increase in the TR score (OR_{-trend} =1.12, 95% CI 1.05-1.19) (Table 1).

There was almost four fold increase in the odds of passing Part B exam at the first attempt with higher AC score (OR-adjusted= 4.12, 95% CI 1.27-13.37), SC score

(OR-adjusted =4.28, 95% CI 1.40=13.09), and TR score (OR-adjusted=3.87, 95% CI 1.27-11.78) (Table 2).

Higher Situational Judgement Test (SJT) score, SC score and AC score were associated with statistically significant higher odds of having satisfactory ARCP outcome (Table 3). However, given the small sample size of the recruited applicants with SJT and ARCP outcome, the confidence intervals are wide.

The Area Under the Curve (AUC) from Receiver Operating Characteristic (ROC) analyses for the independent predictor variables for the progression through training outcomes are presented in Table 4. For Part A exam, RANRA, WGCT, AC, SC, and TR scores have fair discrimination, with AUC for AC was 0.62 (95% CI 0.55-0.70) and of AC and SC combined was 0.66 (95% CI 0.59-0.73). Although AC and SC did not show statistically significant discrimination for Part B exam, AC and SC combined showed statistically significant fair discrimination (AUC=0.64, 95% CI 0.56-0.74). For ARCP, SC score but not AC showed statistically significant discrimination (SC-AUC =0.65 95% CI 0.55-0.75). Despite the relatively small sample size, SJT had statistically significantly better discrimination of candidates with satisfactory ARCP outcome (AUC 0.73 95%CI 0.52-0.94).

Discussion

Main finding of this study

This is the first study that has explored the predictive validity of the UK recruitment and selection process for specialty training in public health. The recruitment and selection process for public health specialty training demonstrates good predictive value with higher scores in the process clearly associated with the likelihood of registrars passing key professional exams in a timely manner. The overall weighted Assessment Center (AC) score is a better predictor than individual AC tests, and the overall weighted Total Recruitment (TR) score is better predictor than the AC score or Selection Center (SC) score separately. While the cognitive ability tests, Rust Advanced Numerical Reasoning Appraisal (RANRA) and Watson-Glaser Critical Thinking Appraisal (WGCT), predict exam performance, Situation Judgement Test (SJT) and SC predict overall performance in training. This suggests that the individual components of the recruitment process are testing different skills and abilities or “constructs” and that together they are providing additive value. Although each component of the recruitment process adds to the cost of recruitment this can be justified by the increase in predictive validity added by each component.

What is already known on this topic

The use of standardised test and approaches is helpful in ensuring quality during recruitment processes with respect to fairness and reliability. There is empirical

evidence from other postgraduate medical recruitment processes that the use of standardised tests and approaches in recruitment is also valid in terms of predicting performance. Recruitment for UK general practice training combines short listing applicants, using clinical problem solving skill and situational judgment tests, with a selection centre which uses job-relevant simulations (patient consultation, group and written simulation exercises), to measure both clinical and non-clinical attributes [6]. Evaluation has demonstrated that each of these selection methods is a significant independent predictor of trainee performance 1 year into training assessed by educational supervisors and performance at the end of training assessed by final Royal College of General Practitioners membership examination [6 7 8]. In anaesthetic recruitment a structured selection centre predicted performance during the first year of training assessed by multiple work place based assessment [9].

Registrars less than 31 years of age were significantly more likely to pass Part A MFPH at the first attempt, and registrars who described themselves as White/British were more likely to past Part B MFPH at the first attempt (Table S3). The higher Part A pass rate might be explained by the higher scores in most components of the selection process achieved by younger applicants. Black and minority ethnic doctors are less likely to pass postgraduate examination than white doctors [10]. Further work is required to determine the possible causes of those differences.

What this study adds

This study adds to the body of evidence and shows that recruitment and selection processes within postgraduate medical training that involves standardized tests and approaches, are valid in terms of the longer term outcomes and performance in the workplace during training.

Limitations of this study

Whilst there may be debate as to whether passing professional examinations is a good proxy for performance in the workplace, there is no doubt that failure to pass relevant examinations in a timely fashion is a key problem in terms of progression and performance for registrars on training programmes. There is currently no overall metric of performance other than examination and ARCP outcomes that is routinely collected.

Although the numbers in this study are modest, they represent the entire cohort of specialty registrars appointed in this four-year period. The study was able to find significant associations with the predictor variables and outcomes despite its limited size. The small sample size was a particular issue for investigating the SJT, for which only 2 years data was available, and outcomes later in training such as the Part B MFPH examination. Follow up of this cohort over a longer period of time, and extension of the cohort to include individuals appointed in later years would be desirable and is planned.

Conclusion

The interim results indicate that the current UK national recruitment and selection process for public health specialty training has good predictive validity of satisfactory progress through training.

References

1. General Medical Council. The Trainee Doctor. London: General Medical Council; 2011.

2. Faculty Public Health. Public Health Specialty Training Curriculum 2010. London: Faculty Public Health; 2010.

3. Pearson TalentLens. *Rust Advanced Numerical Reasoning Appraisal (RANRA).* http://www.talentlens.co.uk/select/rust-advanced-numerical-reasoning-appraisal (accessed 8 March 2015).

4. Pearson TalentLens. *Watson-Glaser Critical Thinking Appraisal.* http://www.talentlens.co.uk/select/watson-glaser-unsupervised-and-supervised-uk-edition (accessed 8 March 2015).

5. Conference of Postgraduate Medical Deans (UK). *A Reference Guide for Postgraduate Specialty Training in the UK. The Gold Guide. Fifth Edition Version: GG5. MAY2014.* http://specialtytraining.hee.nhs.uk/files/2013/10/A-Reference-Guide-for-Postgraduate-Specialty-Training-in-the-UK.pdf (accessed 8 March 2015).

6. Patterson F, Lievens F, Kerrin M, Munro N, Irish B. The predictive validity of selection for entry into postgraduate training in general practice: evidence from three longitudinal studies. *Br J Gen Pract.* 2013;**63(616):**e734-41.

7. Patterson F, Ferguson E, Norfolk T, Lane P. A new selection system to recruit general practice registrars: preliminary findings from a validation study. *BMJ.* 2005;**330(7493):**711-4.

8. 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:**473-78.

9. Gale TC, Roberts MJ, Sice PJ, Langton JA, Patterson FC, Carr AS, Anderson IR, Lam WH, Davies PR. Predictive validity of a selection centre testing non-technical skills for recruitment to training in anaesthesia. *Br J Anaesth.* 2010;**105(5):**603-9.

10. General Medical Council. Outcome reports: Pass rates in postgraduate examinations. London: General Medical Council; 2015. http://www.gmc-uk.org/User_guide___Exams_data_outcome_reports.pdf_60056497.pdf (accessed 1st June 2015).

Conflict of interest:

NP, SG, CD, JP, DW, FP, AK, and BM were involved in the development of the national recruitment process. NP, SG, CD, JP, DW, and BM have been involved with or led delivery of the national recruitment process. FP and AK provided advice to Health Education England and the Public Health National Recruitment Office on development and implementation of selection methodologies through the Work Psychology Group Ltd.

Acknowledgements:

We would like to thank: Marcia Reed and staff at Health Education East Midlands for providing data; Russell Amprofo and staff at the Faculty of Public Health for providing data; and, Training Programme Directors, Head of Schools, selector, registrars, and all other staff involved in the selection and recruitment process.

Figure legend

Figure 1. Schematic diagram of the stages of recruitment process into specialty training in public health

Table 1. Odds ratios and adjusted odds ratios for passing membership examination Part A for cohort of registrars recruited between 2009 and 2012.

Predictor variable	**Pass**	**Fail**	**OR-trend***	**95% CI**	**OR**	**95% CI**	**Adjusted OR #**	**95%CI**
Part A examination N=236	**N=137**	**N=99**					**N=228**	
RANRA			**1.05**	**1.01-1.09**				
RANRA < 56	53	61			1.00		1.00	
RANRA ≥ 56	84	38			**2.54**	**1.50-4.33**	**2.63**	**1.50-4.61**
WGCT			**1.06**	**1.02-1.11**				
WGCT < 59	52	56			1.00		1.00	
WGCT ≥ 59	85	43			**2.13**	**1.26-3.60**	**2.14**	**1.21-3.79**
SJT N=103			1.01	0.91-1.12				
SJT < 58	26	20			1.00		1.00	
SJT ≥ 58	33	24			1.06	0.48-2.32	1.05	0.44-2.49
AC score			**1.08**	**1.03-1.13**				
AC <58	58	63			1.00		1.00	
AC ≥ 58	79	36			**2.38**	**1.40-4.06**	**2.43**	**1.37-4.30**
SC score			**1.07**	**1.02-1.13**				
SC < 58	59	60			1.00		1.00	
SC ≥ 58	78	39			**2.03**	**1.20-3.44**	**1.88**	**1.09-3.24**
Total Recruitment			**1.12**	**1.05-1.19**				
TR< 58	57	63			1.00		1.00	
TR ≥ 58	80	36			**2.46**	**1.44-4.18**	**2.27**	**1.31-3.93**

*OR-trend – the independent variables are continuous
Adjusted for age, sex, ethnicity and professional background

Table 2. Odds ratios and adjusted odds ratios for passing membership examination Part B for cohort of registrars recruited between 2009 and 2012.

Predictor variable	**Pass**	**Fail**	**OR-trend***	**95% CI**	**OR**	**95% CI**	**Adjusted OR #**	**95%CI**
Part B examination N=188	**N=169**	**N=19**					**N=181**	
RANRA			1.05	0.98-1.12				
RANRA < 56	74	12			1.00		1.00	
RANRA ≥ 56	95	7			2.20	0.83-5.87	2.25	0.79-6.43
WGCT			1.04	0.97-1.12				
WGCT < 59	74	10			1.00		1.00	
WGCT ≥ 59	95	9			1.43	0.55-3.69	1.25	0.44-3.54
SJT N=77			1.11	0.91-1.34				
SJT < 58	30	5			1.00		1.00	
SJT ≥ 58	40	2			3.33	0.60-18.37	2.66	0.43-16.57
AC score			1.07	0.99-1.16				
AC <58	74	14			1.00		1.00	
AC ≥ 58	95	5			**3.59**	**1.24-10.43**	**4.12**	**1.27-13.37**
SC score			1.08	0.98-1.18				
SC < 58	73	13			1.00		1.00	
SC ≥ 58	96	6			**2.85**	**1.03-7.86**	**4.28**	**1.40-13.09**
TR score			1.12	1.00-1.25				
TR< 58	73	13			1.00		1.00	
TR ≥ 58	96	6			**2.85**	**1.03-7.86**	**3.87**	**1.27-11.78**

*OR-trend – the independent variables are continuous
Adjusted for age, sex, ethnicity and professional background

Table 3. Odds ratios and adjusted odds ratios for having satisfactory ARCP outcome for cohort of registrars recruited between 2009 and 2012.

Predictor variable	**Pass**	**Fail**	**OR-trend***	**95% CI**	**OR**	**95% CI**	**Adjusted OR #**	**95%CI**
ARCP N=212	**N=179**	**N=33**						
RANRA			1.01	0.96-1.07				
RANRA < 56	90	20			1.00		1.00	
RANRA ≥ 56	89	13			1.84	0.83-4.06	2.22	0.99-5.20
WGCT			1.05	1.00-1.11				
WGCT < 59	80	16			1.00		1.00	
WGCT ≥ 59	99	17			1.16	0.54-2.48	1.43	0.62-3.27
SJT N=83			1.23	0.98-1.54				
SJT < 58	30	6			1.00		1.00	
SJT ≥ 58	45	2			3.63	0.66-19.91	**9.88**	**1.20-81.51**
AC score			1.05	0.98-1.12				
AC <58	86	22			1.00		1.00	
AC ≥ 58	93	11			**2.27**	**1.01-5.09**	**3.45**	**1.41-8.46**
SC score			**1.09**	**1.01-1.17**				
SC < 58	83	22			1.00		1.00	
SC ≥ 58	96	11			2.05	0.93-4.53	**2.41**	**1.05-5.52**
TR score			**1.12**	**1.02-1.23**				
TR< 58	85	21			1.00		1.00	
TR ≥ 58	94	12			1.71	0.78-3.73	1.99	0.89-4.49

*OR-trend – the independent variables are continuous
Adjusted for age, sex, ethnicity and professional background

Table 4: Receiver Operating Characteristic analysis to show how well the individual and combined components of the recruitment discriminate between candidates likely to pass or not membership examinations and have ARCP outcome as satisfactory or not.

	Training outcomes					
Predictor variable	**Part A Membership exam N=236**		**Part B Membership exam N=188**		**Satisfactory ARCP Outcome N=212**	
	AUC	**95% CI**	**AUC**	**95% CI**	**AUC**	**95% CI**
RANRA	**0.60**	**0.53-0.68**	0.60	0.46-0.74	0.53	0.42-0.64
WGCT	**0.60**	**0.52-0.67**	0.57	0.43-0.70	0.57	0.47-0.68
SJT score *	0.53	0.41-0.64	0.64	0.47-0.80	**0.73**	**0.52-0.94**
AC score	**0.62**	**0.55-0.70**	0.62	0.49-0.75	0.57	0.46-0.69
SC score	**0.61**	**0.54-0.68**	0.62	0.49-0.76	**0.65**	**0.55-0.75**
TR score	**0.64**	**0.57-0.71**	0.65	0.52-0.78	**0.66**	**0.56-0.76**
AC and SC combined	**0.66**	**0.59-0.73**	0.64	0.56-0.73	**0.62**	**0.55-0.71**
AC, SC, and demographics	**0.68**	**0.61-0.75**	0.58	0.50-0.67	**0.62**	**0.54-0.70**

* N=103 registrars with Part A exam; 77 with OSPHE; 83 with ARCP record