Title

The predictive validity of selection for entry into postgraduate training in general practice and psychiatry: what value do face-to-face selection processes add? A retrospective cohort study

Summary

Medical specialty training selection processes have needed to change substantially in response to the COVID-19 pandemic, where face-to-face selection procedures were not feasible. The Multi-Specialty Recruitment Assessment (MSRA) is a computer-based assessment used by several specialties, including General Practice and Psychiatry. The MSRA has now been used for recruitment into post-graduate medical training for these specialties for a number of years. Given the MSRA has been shown to be a reliable form of standardised assessment, and that it can be delivered remotely it became an important part of the pandemic contingency planning. Previous research has already demonstrated that performance on the MSRA validity predicts future performance in relation to GP training (Patterson et al., 2013; Patterson et al., 2015). While there is no published research into the predictive validity of the MSRA in psychiatric training, unpublished research conducted using Royal College data demonstrated that the MSRA scores are predictive of selection outcomes, postgraduate exam performance and ARCP ratings. In addition, provisional, as yet unpublished findings from analyses of data drawn from the UKMED have also supported the validity of the MSRA in these specialties (application ref UKMEDP154). Bearing in mind the previous evidence supporting the use of the MSRA in these two specialties, the respective selection teams made the decision to use the MSRA as the sole decision maker for selection into General Practice and Psychiatry postgraduate training in 2020 and 2021. The relevant data held in the UKMED were accessed to explore the potential risks around using a single assessment method for selection and to inform immediate decisions. Previously, the value (or otherwise) of face-to-face selection processes were raised in a review of GP selection (Davison et al., 2016). Therefore, the aim of this study is to further explore the incremental value of face-to-face selection processes for recruitment into GP and psychiatry training. Such components of selection generate a relatively high cost to the NHS. The data will be used to explore, in more depth, the relationships between selection assessment performance and subsequent in-training outcomes. This will provide a better understanding of the potential added value of face-to-face selection in this context. This would inform future discussions around use of the MSRA in specialty selection. For example, whether only certain applicants should be prioritised for face-to-face assessment.

References

1. An evaluation of GP specialty selection which questions the value of face to face selection processes
2. Research exploring the relationship between general practice selection scores and MRCGP examination performance - it is proposed this application will build upon this research
3. Research exploring the predictive validity of selection into GP training using longitudinal studies - it is proposed this application will build upon this research

Research questions

1. To what extent does face-to-face selection demonstrate incremental validity over the MSRA, in predicting in-training performance outcomes, in terms of pass at first attempt at the clinical component of the relevant Royal College Membership exams and Annual Review of Competence Progression (ARCP) ratings? 2. For poorer performers on the MSRA (i.e. those scoring below the lowest quartile in either / all components), to what extent does face-to-face selection demonstrate incremental validity over the MSRA components in predicting in-training performance outcomes (ARCP and clinical postgraduate exams)? 3. What are the levels of differential attainments for several groups of individuals with protected characteristics on the differing components of the MSRA and face-to-face selection? Analyses will be conducted separately for recruitment into General Practice and Core training in Psychiatry.

Data required

Data Item	Data Type	Description
BME_INT	#N/A	#N/A Higher level ethnicity coded as an integer.
BME	#N/A	Higher level ethnicity coding: BME or white.
ETHNICITY_SRC	#N/A	Source system of the ETHNICITY_L1 and ETHNICITY_L2 data stored for this record.
ETHNICITY_L2_INT	#N/A	Ethnicity Level 2 information coded as an integer
ETHNICITY_L2	#N/A	Ethnicity Level 2 information.
ETHNICITY_L1_INT	#N/A	Ethnicity Level 1 information coded as an integer
ETHNICITY_L1	#N/A	Ethnicity Level 1 information
PRIMARY_PMQ	#N/A	Primary medical qualification as per the GMC’s register.
PRIMARY_PMQ	#N/A	Primary medical qualification as per the GMC’s register.
GENDER_INT	#N/A	Gender coded as integer for analysis: 1 = Female and 0 = Male.
GENDER	#N/A	Gender
BIRTH_YEAR	#N/A	Birth year extracted from Date of Birth held on the GMC register. Date of birth is too identifiable for inclusion in extracts.
BIRTH_MONTH	#N/A	Birth month extracted from Date of Birth held on the GMC register. Date of birth is too identifiable for inclusion in extracts.

Data Item	Data Type	Description

Data Item	Data Type	Description

Data Item	Data Type	Description

Data Item	Data Type	Description

Data Item	Data Type	Description

Data Item	Data Type	Description

Data Item	Data Type	Description

Data Item	Data Type	Description

Data Item	Data Type	Description

Data Item	Data Type	Description

Methodology

Analyses will be conducted separately for each specialty. The analysis will be conducted using the data for those who went through the specialty selection process and sat the MSRA between 2015 to 2018, and who were offered and accepted a training post. The MSRA scores from the applicant's successful application prior to entry into the specialty will be used, where multiple applications over time have been made. Royal College exam data will only be used if the exam has been sat following entry into the specialty. A limitation of this research will be a restriction of range. Consideration will be given to how this may affect the composition of the sample analysed. The spread of scores across metrics will be examined, compared to the full MSRA sample, to explore the potential skew of the data. Given both specialties have low competition ratios, it is anticipated the restriction may not be substantial, however. Listwise deletion will be used, by default, to treat missing data for each component of the analysis. However, if missing data are substantial, we will use imputation to address this issue. In this context imputation could be applied to both predictors, but also to missing values for the categorical outcomes (e.g. Clinical exam pass at first attempt/yes/no) as both a way of addressing the restriction of range and as a form of sensitivity analysis. It is likely that any imputation, if used, would be conducted using multiple imputation using chained equations (MICE) as implemented in Stata. A final decision about the approach to missing values will be made following an exploration of the data. Initially univariable analyses will be conducted to model the relationship between the selection scores on the MRSA (predictor variables), and in-training performance data. In this respect the primary outcome variables of interest will be pass at the clinical components of the respective membership exams. These dichotomous outcomes have been selected for several reasons. Firstly, scores relative to the pass mark, do not correspond perfectly with a pass or fail. For example, if too many stations are failed in the psychiatry clinical exam- the Clinical Assessment of Skills and Competencies (CASC). Secondly, dichotomous pass/fail of clinical examinations is more likely to reflect examiner judgments of whether a candidate is, overall, deemed clinically competent. Thirdly, in the case of the CASC the structure of the examination has been changed over the study period. The dichotomous metric of pass/fail will thus be less affected by the fact that the structure has changed over the years, and also this mitigates against those trainees who fail several stations, and therefore fail with a higher mark than another candidate with a bare pass. Secondary outcomes that may be evaluated include scores on clinical knowledge exams and performance at the Annual Review of Competence Progression (ARCP). The clinical knowledge exam results will be used in two ways. The deviation from the pass score will be used for some analysis, as it allows better comparison of data across diets. In addition, the proportion of trainees passing the exam will be used for some analyses. In terms of predicting ARCP ratings: the selection scores will be compared to ARCP ratings across training. ARCP ratings will be analysed as ordinal outcomes in a multilevel ordinal logistic regression. The approach used will be as described by Tiffin et al (BMJ 2014), previously shown to be valid. Record of in-training assessment scores will be recoded to the equivalent ARCP outcome codes. Only ARCP 'competency-based' outcomes indicating training progress were included (e.g., 'out of programme' was excluded). The remaining outcomes will be recoded onto a 4-point ordinal scale: 1. satisfactory progression / training programme completed = ARCP outcomes '1' or '6', respectively); 2. additional evidence requested = ARCP outcome '5'; 3. targeted training required (no extended time) = ARCP outcome '2'; 4. extended training time required/left programme = ARCP outcomes '3' or '4', respectively. A limitation is that the ARCP ratings lack granularity and are unlikely to discriminate between average and good performance. Nonetheless, this metric is helpful in exploring relationships between selection methods and poor performance, and when considered alongside other measures of performance. Educational Performance Measure (EPM) data may also be used to understand to what extent this metric adds to the prediction of later performance, in UK graduates. Therefore this data has been requested, along with PLAB scores. The univariable analyses will be used identify which of the predictor variables are most predictive of outcomes, when considering multivariable model building. The univariable analyses will also be used to address the third research question; the potential impact of any proposed changes on different demographic sub-groups (by gender, place of qualification and ethnicity). Multivariable models will be used to evaluate the incremental validity of each selection measure, to inform how best the selection data can be used and combined (weighted), to select those who are most likely to go on to perform well in training.

Analysis proposed

For research questions 1 and 2, initial analysis has already been outlined (and conducted) in research application ref UKMEDP154. That is; descriptive statistics will be run, to show average MSRA scores and selection scores by ARCP rating (all satisfactory / one or more unsatisfactory rating) and exam result (pass or fail). Formal statistical tests, as appropriate (e.g. one way independent t-tests) will be used to explore whether any differences observed are unlikely to be due to chance (i.e. statistically significant). MSRA and selection scores will be banded and frequencies the proportion of trainees receiving satisfactory or less than satisfactory ARCP ratings will be presented in a histogram, to demonstrate performance compared to MSRA and selection scores. Regression analyses and correlations (as appropriate) will be carried out to explore the strength of relationships between selection scores and exam performance. This research will build on our previous analysis by using regression analysis to further examine the predictive and incremental validity of the selection methods. In analysing associations between variables, for continuous outcomes, coefficients will be corrected for multivariate restriction of range. The aim will be to define a structural model to evaluate the added value of each selection method. If skew is observed, the standard errors will be bootstrapped. Modelling pass at clinical examinations at first attempt will be modelled using single level logistic regression. As mentioned earlier, as multiple ARCPs are taken, then multi-level ordinal logistic regression will be used to model this outcome. Continuous clinical knowledge exam scores, when treated as outcomes, will be modelled using linear regression, using bootstrapped standard errors, if not normally distributed. The overall analytic strategy is to build models of increasing complexity, from univariable models, to multivariate regressions, and if indicated and supported by the data and earlier findings, to path analysis. In this sense path analysis could be useful in testing and modelling the direct and indirect effects of various predictors, for example performance on the situational judgement test components of the MSRA. Any path analysis would be implemented in the software package Mplus. In this situation one of the team's Mplus licenses would be installed within the Safe Haven as Mplus is not standard software for the HIC at present. For research question 2, this analysis will be repeated for poorer performers on the MSRA (i.e. those scoring below the lowest quartile in either / all components). This will be used to explore whether there are any trends in the data which may suggest where further intervention may be beneficial for the trainees who fall in this group. Analyses will be conducted to explore the adverse impact of selection methods, in line with research question 3. Regression analyses will be run, with effect sizes reported by demographic group.

Timeline

Proposed Date: 2022-02-01

Duration: 8 month(s)

February - April 2022 Data received and cleaned May - June 2022 Data analysis conducted July - September 2022 Write up report

Proposal for dissemination

The results will be reported to subgroups within HEE, as well as the Royal Colleges (including chief examiners and recruitment leads), to inform their ongoing approach to selection. It is intended that the findings of this research are submitted to be published in a peer-reviewed journal, with the aim of contributing to the research in this area and providing to context to stakeholders around selection into GP and Psychiatry.

Researchers and partners

Name	Org	Role	Proposed Role	Experience with managing, cleaning and organising large datasets (>5000 cases)	Experience with statistics and data modelling	What software do you have experience and expertise in using?	CV	
Professor Fiona Patterson f.patterson@workpsychologygroup.com Office 110, Cubo Derby, Victoria Street, Derby DE1 1EQ +44 (0)793 122 9344	Work Psychology Group	Lead Researcher	Lead and project supervisor; lead on design of the analysis and writing the results for publication.	Extensive experience with managing the UKCAT and MSRA selection data sets, amongst other large data sets.	Extensive expertise in defining research plans and publishing in peer reviewed journals.	SPSS, Excel	Fiona is a leading expert in the field of assessment, selection and innovation in organisations. She is the founding Director for Work Psychology Group, a research-led organisational psychology consulting practice. She has worked and published extensively in the healthcare sector, and was appointed as honorary visiting Professor at the University of Nottingham Medical School in 2010. She recently guest edited for a special issue of AMEE's MedEdPublish journal on selection and recruitment.	Included in communication: Yes Access to safe haven: No
Ms Elis Sugiyarto e.sugiyarto@workpsychologygroup.com Office 110, Cubo Derby, Victoria Street, Derby DE1 1EQ +44 (0) 7388 925 987	Work Psychology Group	Researcher	Elis' role will be to support with the management, cleaning, and analysis of data, and contribute toward the write up.	Experience with managing large datasets for selection into international medical schools (e.g. Stellenbosch University, South Africa, >8000 cases) and large UKCAT selection datasets (>35,000 cases).	BSc in Psychology, MSc in Occupational Psychology. Experience of using parametric and non-parametric tests; experience conducting exploratory and confirmatory factor analysis and structural equation modelling.	Excel, SPSS, MPlus programme (introductory level).	Elis is a Psychologist at Work Psychology Group, holding a First Class Honours BSc in Psychology from Cardiff University and an MSc with Distinction in Occupational Psychology from the University of Manchester. At a national and international level, Elis has worked as a Research Assistant at King's College London's Institute of Psychiatry, Psychology and Neuroscience (KCL IoPPN) and the University of Santo Tomas, Philippines.	Included in communication: Yes Access to safe haven: Yes

Doctor Paul Tiffin Paul.tiffin@york.ac.uk Mental Health and Addiction Research Group, Department of Health Sciences, University of York Area 4, A/RRC/103, Alcuin Research Resource Centre, University of York, Heslington, York YO10 5DD 01904 321117	University of York	Researcher	Support with data cleaning, analysis and write up.	Extensive experience of managing, cleaning and analysing large datasets using conventional statistical techniques.	I have extensive experience of conventional statistical approaches as well as machine learning. This includes the use of multilevel models, structural equation modelling, latent variable modelling and the application of item response theory, as well as the use of Bayesian approaches. I also have experience of developing and implementing machine learning algorithms.	Stata Mplus R Python	Paul is a Reader in Psychometric Epidemiology/Consultant Psychiatrist and quantitative methodologist. He is a member of the UKMED Research Subgroup and led UKMED P001- a study into FtP in undergraduates. Paul has published over 80 journal papers, mainly as first/last named author and raised over £5.5M in research funding as a lead and co-applicant.	Included in communication: Yes Access to safe haven: Yes
Mrs Emma Morley e.morley@workpsychologygroup.com Office 110, Cubo Derby, Victoria Street, Derby DE1 1EQ +44 (0) 7572 871 228	Work Psychology Group	Researcher	Emma's role will be to coordinate the project team, support with the data management and cleaning, and input into the analysis and write up.	Extensive experience with managing the UKCAT and MSRA selection data sets, amongst other large data sets.	BSc in Psychology, MSc in Occupational Psychology, CPsychol; Occupational Psychologist registered with the Health and Care Professionals Council (HCPC). Experience of using a range of parametric and non-parametric tests, for example, analysis, correlations, multiple regression, logistic regression, ANOVA, ANCOVA, MANOVA and MANCOVA.	SPSS Excel	Emma is a Chartered Member of the British Psychological Society (BPS) and an Occupational Psychologist registered with the Health and Care Professionals Council (HCPC). She holds an MSc with Distinction in Occupational Psychology from the University of Sheffield. Emma specialises in designing and evaluating bespoke and innovative assessment solutions in the healthcare professions, informed by the latest research literature and intention to disseminate key learnings.	Included in communication: Yes Access to safe haven: Yes

Funding

No additional funding is required for this work to proceed.

Support software details

Ideally both SPSS and Stata if possible. If progressing to path analysis Mplus would have to be installed. At least one of the applicants has a current, personal, Mplus license so we understand for a modest fee (£200-£300) this can be installed in the Safe Haven to be used by the team member.