Differential attainment in the MRCPsych according to ethnicity and place of qualification between 2013 and 2018: a UK cohort study

Paul A Tiffin (0000-0003-1770-5034), Lewis W Paton (0000-0002-3328-5634)

Author details

PAT (corresponding author): Department of Health Sciences, University of York & Health Professions Education Unit, Hull York Medical School, Y0105DD, UK, paul.tiffin@york.ac.uk

LWP: Department of Health Sciences, University of York, YO10 5DD, UK

Word count: 8360

ABSTRACT

Purpose of the study

To explore if differential pass rates exist in the clinical component of the UK postgraduate clinical psychiatry exam, the Clinical Assessment of Skills and Competencies (CASC), according to ethnicity and place of qualification (UK vs EEA vs overseas graduates).

Study design

Observational study using data from the UK Medical Education Database for 2,140 doctors, sitting the CASC for the first time between 2013 and 2018.

Results

After controlling for age, sex, time of sitting and performance in the written components of the MRCPsych, differences in CASC pass rates persisted between UK graduates self-identifying as Black and Minority Ethnicity (BME) and non-BME (odds ratio (OR) for passing 0.36, 95% confidence interval 0.23 to 0.56, $p<0.001$). Both EEA (OR 0.25, 0.15 to 0.40, $p<0.001$) and overseas graduates (OR 0.07, 0.05 to 0.11, $p<0.001$) were less likely to pass the CASC at first attempt, even after controlling for the influence of educational and background variables. These groups, on average, had lower scores on written exams with substantial content relating to procedural skills (e.g. critical appraisal) rather than pure recall of factual knowledge.

Conclusions

Substantial differences exist in clinical examination performance between UK BME and non-BME candidates, as well as between UK and non-UK graduates. These differences are not explained by differing levels of clinical knowledge. In the interests of equality this situation requires further investigation and remediation. Future research should focus on understanding how potential bias may be acting within different stages of recruitment, training and assessment within psychiatry.

KEYWORDS

psychiatry, medical education

MAIN MESSAGES

- Substantial differences in performance exist for the clinical examination of the MRCPsych (the CASC) between UK medical graduates who identify as Black and Minority Ethnic (BME) and those who identify as non-BME.
- Similar differences, though of larger magnitude, also exist for CASC performance between UK medical graduates and those who obtained their primary medical qualification outside of the UK.
- These differences persist even after controlling for the potential influence of educational and background variables, including performance on the written components of the MRCPsych examination.

RESEARCH QUESTIONS

- Are there any sources of potential bias in the MRCPsych clinical examination, especially in relation to possible interactions between and rater characteristics?
- Are there any differences in the postgraduate training experiences between differing groups of psychiatrists in training that could help explain the differential performance in the MRCPsych, and in particular, the CASC?
- Do any inter-group differences in academic performance translate into any meaningful differential variations in actual clinical behaviour and/or patient outcomes?

INTRODUCTION

In order to practise as a Consultant Psychiatrist in the United Kingdom (UK) doctors generally must pass the MRCPsych (Membership of the Royal College of Psychiatrists) examination, which is set by the Royal College of Psychiatrists. The only exception to this is when doctors may be placed on the General Medical Council (GMC) specialist register for psychiatry via a Certificate of Eligibility for Specialist Registration (CESR). This alternate process involves providing evidence to the Royal Colleges of Psychiatrists that previous training and experience would be equivalent to that obtained by the normal membership route.[1] Thus, postgraduate examinations, such as the MRCPsych, serve as the main gateways, and potential barriers, to advanced specialty practise. Consequently, it is vital that they are both rigorous, though fair, and not unduly disadvantaging any specific groups or individuals. For this latter reason there has been considerable interest in any differences in pass rates between groups of doctors for such examinations.

Such a differential pass rate at the practical component of the Royal College of General Practitioners membership examination, between British black and minority ethnic (BME) candidates and those self-identifying as white, has previously been reported. This difference persisted even after controlling for previous performance on the written, knowledge-based component of the examination.[2] These findings led to (ultimately unsuccessful) legal action taken by the British Association of Physicians of Indian Origin (BAPIO), against the Royal College of General Practitioners and the GMC, who felt that they constituted evidence of institutional racial bias.[3] However, the underlying reasons for such inter-group differences are still not fully understood, though they are likely to be related, at least partly, to subtle cultural issues. For example, a linguistic analysis of candidates sitting the practical component of the general practice examination, the 'clinical skills assessment', concluded that black and minority ethnic (BME) UK graduates may show some of the subtle differences in communication style also observed in non-UK doctors.[4]

In addition, differential pass rates according to place of primary medical qualification have also been reported across a number of postgraduate clinical examinations.[5] Indeed, differences in pass rates, between UK and overseas medical graduates, have previously been reported for the MRCPsych in 1999.[6] Indeed, relatively reduced postgraduate academic performance across a number of markers has been observed for doctors graduating from countries other than that of their practise.[7-10] For example, at the Annual Review of Competence Progression (ARCP) panels, which review progress in training for doctors, non-UK graduates were more likely to receive poor ratings. Moreover, the specialties with the largest disparities between groups were psychiatry and general practice.[11] It was also notable that once the effects of postgraduate (membership) exam failures were removed, the degree of difference between graduate groups, in terms of ARCP outcomes, diminished substantially, though remained statistically significant. This suggests that such clinical exams could pose a specific hurdle to overseas doctors' career progression. Moreover, at the time it was hypothesised that these two specialties may have been the most sensitive to issues relating to culture and communication, especially in relation to the clinical components of the postgraduate examinations.

Understanding, and addressing, such differential attainment is a matter of priority for a number of key reasons. Firstly, there is clearly a matter of social justice at stake, and if there is ethnic bias at work, even if it was unconscious bias, then it should be remediated immediately. Secondly, in Western countries, such as the UK, psychiatry is a less popular specialty than most. For example, during the 2019 recruitment round only 86% of 'core' (basic) psychiatric training places were filled across England, though this is a marked increase over previous years.[12] This has led to a heavy reliance on doctors who qualified overseas; over 40% of UK psychiatrists graduated from outside of the country.[13] Thus, if clinical examinations are a barrier to career progression for overseas and ethnic minority

doctors this presents a major workforce issue. The Royal College of Psychiatrists is aware of these issues and has already taken steps to help ensure that the membership examination is as fair as possible.[14] The availability and linkage of data relating to place of qualification, ethnicity and postgraduate academic achievement, via the UK Medical Education Database (UKMED) now makes a detailed study of differential attainment in psychiatry possible.[15] Thus, we are in a position to evaluate the extent to which any differences in performance between groups of candidates persist, and consider what further investigations and other measures may be required to address these.

Psychiatric speciality training in the UK

Recruitment to UK-wide speciality training (excluding Northern Ireland) is currently organised by Health Education England (HEE) via the National Psychiatry Recruitment Office. Once a doctor completes the initial two year 'foundation programme' after qualifying, they may apply for a place on a specialty training scheme. In psychiatry, 'core' training last three years and is usually made up of six-month approved, supervised training placements, rotating through a variety of specialties (e.g. learning disability, older people's psychiatry etc.). During this period a trainee will aim to pass all parts of the MRCPsych, which are required for progression to higher psychiatry training in one or two psychiatry specialties that the trainee wishes to work in as a consultant, for example general psychiatry.[16] Successful completion of the training placements, as evaluated by the doctor's e-portfolio, which includes workplace based assessments and feedback, and passing the MRCPsych permits an application to be made to 'higher specialist training'. This usually comprises of three years of further supervised experience in 9-12 month blocks within a psychiatric specialty. Successful completion of this stage of training results in an award of a 'certificate of completion of training'. This entitles the doctor's name to be placed on the list of specialists held by the GMC (the UK medical regulator) and to apply for Consultant grade roles with healthcare providers.

The route for non-UK graduates often differs from this, although depending on prior experience, overseas doctors may join this pipeline at different stages. Currently doctors graduating from a recognised institution within the European Economic Area (EEA) do not have to provide evidence of language fluency in English or have to pass the Professional and Linguistic Assessment Board (PLAB) examinations. In contrast, this requirement must be met by most overseas (non-EEA) doctors seeking UK-based medical training and registration. To take the PLAB examinations, overseas medical graduates must provide evidence of language competency in English. This usually means having obtained sufficient scores on the International English Language Test System (IELTS). The IELTS is made up of four parts—*listening*, *speaking*, *reading*, and *writing*—and can be taken as many times as desired.[17] Each part is rated between band one (non-user) and band nine (expert user). In order to be eligible to sit the PLAB examination an IELTS score of at least 7.0 on each part is required with an overall rating of 7.5 achieved, recently raised from the need for an overall score of 7.0. The PLAB examination is designed to ensure that overseas doctors demonstrate the clinical competencies equivalent to those that a UK graduate would be expected to have obtained by the end of the first foundation year when they are eligible for full registration with the GMC. Thus, successful PLAB graduates are able to apply for foundation year two posts in competition with both UK and EEA medical graduates. Some overseas graduates may be able to provide evidence of previous experience which allows them to apply straight for specialty training without completing the second year of the foundation programme.[18]

The MRCPsych examination

This postgraduate examination, which must be passed in order to become a full member of the Royal College of Psychiatrists and progress to higher specialist training, is made up of several components. Over the last decade or so some changes have been made to the structure of the MRCPsych. From March 2008, it was composed of three written exams and

a practical exam, the 'Clinical Assessment of Skills & Competencies' (CASC). The three written exams ('papers 1 to 3') contained selected response questions, with both single best answer and 'extended matching' formats. These written tests evaluated knowledge of relevant basic science, research evidence, clinical psychiatry and medical statistics. Changes were made in 2015 so that only two written papers were taken, papers A and B. As a transitional arrangement, in order to accommodate candidates who had previously passed papers 1 or 2, paper A was split into two halves (AI and AII), covering different aspects of the MRCPsych syllabus.[19] Paper A tests knowledge of the science and theory underpinning psychiatric practice (behavioural and social sciences, human development, neuroscience, psychopharmacology and psychiatric classification). Paper B tests knowledge of critical appraisal of research evidence and clinical topics relevant to all the psychiatric specialties. The CASC uses the format of an Objective Structured Clinical Examination (OSCE), made up of two (morning and afternoon) circuits of stations designed to evaluate a candidate's clinical skills. In the morning circuits, candidates have 4 minutes to read any instructions for the station and 7 minutes to complete the task. In the afternoon, this becomes 90 seconds of reading and, again, 7 minutes to complete the task. In total, there are 16 CASC stations, which test 'history taking', 'examination' (physical and mental) and 'management'. The scoring system and pass standard for the CASC employs the borderline regression method.[20] In the case of the CASC, this means that each station is marked by a trained examiner who provides two sets of scores. The first is a five point 'analytic' global domain score, ranging from 1 ('poor') to 5 ('excellent') for between three and five domains. The second is a 6-point overall global judgement- 'Excellent Pass', 'Pass', 'Borderline Pass', 'Borderline Fail', 'Fail', or 'Severe Fail', which have associated grade descriptors to help anchor them. The total weighted domain scores are then regressed onto their global scores to produce a linear (regression) equation for each station for all candidates. The total domain score for borderline candidates, determined through the 'line of best fit' is used to set the pass mark for that station. The pass mark for the whole exam then is calculated as the average of the station pass marks for that day, with the addition of the standard error of

measurement. In order to pass the CASC a candidate must achieve a passing score in at least 12 of the 16 stations and meet or exceed the overall total borderline regression score set. Thus, high scores on some stations will not necessarily compensate for low scores in others. The minimum of 12 stations as a pass standard was set on the basis that the five history taking and five examination stations cover basic clinical skills and that a 'borderline pass' candidate should be expected to pass eight out of these ten stations. Six of the stations relate to clinical management and a 'borderline pass' candidate was expected to pass four out of these six stations.[21] The CASC is thus administered according to accepted standards of practice in the area of clinical educational assessment, three years after its introduction. A previous study reported that, three years after its introduction, candidates and examiners were somewhat divided regarding whether the examination evaluated all the advanced skills required to practise effectively as a psychiatrist.[22]

The aim of this study was to examine two distinct, though overlapping issues, related to differential attainment in the MRCPsych. Firstly, to examine whether self-reported ethnicity was associated with lower pass rates at the MRCPsych in UK medical graduates, and in particular at the CASC, and whether any differences persisted after controlling for potential confounding factors. These latter factors would also include performance in the written, knowledge-based parts of the examination. Secondly, whether world region of qualification (UK, EEA or non-EEA ['Overseas']) was associated with differential pass rates at the MRCPsych, and, again, to what extent these were independent of potentially confounding factors. Our findings would have implications for both the education and support of psychiatrists in training, as well as medical regulatory policy relating to doctors qualifying outwith the UK.

METHODS

As stated above, our primary aim was to describe the pass rates at the CASC for UK medical graduates according to their self-reported ethnicity (BME vs non-BME). A secondary aim was to explore the impact of world region of qualification (UK vs EEA vs Overseas) on CASC pass rates. In terms of implications for policy and practice, it is the raw (unadjusted) results from univariable analyses of pass rates that matter most- that is, those apparent before adjustment is made for any potential confounding factors. However, we also sought to evaluate the degree to which the main effects of interest (ethnicity and place of qualification) were independent of potential confounding factors. For this reason the effects of several key demographic and educational factors, known to be associated with postgraduate educational performance (such as candidate gender and age), were also controlled for.

Data sources and preparation

All data were held and linked within the UKMED which then provides research extracts in a Safe Haven.[15] With a Safe Haven data are managed and analysed within a remote secure server from which only reports on aggregated data can be extracted. Note also that UKMED statistical disclosure controls stipulate that all numbers included in public documents are rounded to the nearest five.[23] This blunting has been applied to all numbers cited in this report. The flow of data through the study is shown in Figure 1.

As the study used routinely collected, deidentified data, ethical approval was not required. This was confirmed in writing by the Chair of the University of York Health Sciences Ethics Committee. Moreover, individual informed consent for the use of the data was not required for this study. This is because all the data used were held within the UKMED. The use of personal data in UKMED is not reliant on individual consent from data subjects, as it is not a necessary condition for processing under the General Data Protection Regulation (GDPR), which allows personal data to be used without consent where it is necessary for statutory functions. The Medical Act 1983 gives the General Medical Council (GMC) a legal responsibility to promote high standards of medical education and co-ordinate all stages of

medical education. This enables the creation of the database and the disclosure of data by other providers in compliance with the GDPR. For more information, see www.ukmed.ac.uk/faq.

The primary outcome for this study was the odds of passing the CASC at first attempt. The UKMED contained data on 2140 registered doctors with such an outcome reported for a first attempt at the CASC. The UKMED holds Royal College Examination data from those who sat such tests from 1st August 2013. Doctors are identified as being in psychiatric training according to the specialty recorded by the deaneries in their returns to the GMC that allow the GMC to administer the National Training Survey,[24] which is conducted annually by the regulator.[25] Only the first attempts were analysed, for two reasons. Firstly, the first attempt at an examination is considered to be the best marker of true underlying ability, with the subsequent probability of passing increasingly capitalising on chance/practise effects.[26] Secondly, UKMED holds data on all UK registered doctors who entered medical school from 2002 or joined postgraduate training schemes from 2012. Thus, given this timescale, the majority of doctors represented in UKMED who had attempted the CASC had done so only once. For example, whilst there was an outcome recorded for first attempt for 2,140 doctors, there was only one for second attempt in 590 cases.

The primary outcome for this study was performance on the CASC part of the MRCPsych. This was because previous concerns relating to possible ethnic bias in the Membership of the Royal College of General Practitioners (MRCGP) examination were related to the Clinical Skills Assessment, as the clinical component, rather than the written knowledge test.[2] As these clinical examinations are judged by human raters there is more scope for such bias to occur compared to written tests, which are machine-marked. The individual marks on the CASC were not used as a continuous outcome for modelling purposes because, due to the usual standard setting processes, raw scores did not equate directly to the pass/fail status. Thus, for modelling purposes, the odds of passing the CASC at first

attempt, rather than overall mark achieved, was used as the (binary) outcome of interest. Moreover, the granularity of the data on CASC performance that were received by the UKMED from the Royal College of Psychiatrists varied across years.

As highlighted earlier, the structure of the written component of the MRCPsych has changed slightly during the study period. For this reason, in order to obtain an estimate of the overall academic performance of the doctors the scores obtained on each written paper, relative to the pass mark for that sitting, were averaged for each candidate across each paper taken to provide an overall metric of performance at the written component of the MRCPsych. A quantile ('Q-Q plot') demonstrated that these scores for performance on the written component approximated a normal distribution, except for relatively extreme values (15 points below, or 20 points above the pass mark).

The nationality and the name and country of the medical schools where the primary medical qualifications were obtained were derived from the GMC List of Registered Medical Practitioners (LRMP). Where dual nationality was recorded, in a very small number of cases, only the first nationality provided to the GMC was used. Graduate group status was categorised via the world region where the primary medical qualification was obtained, in order to approximately align with UK regulatory policy (UK, EEA and outside of the EEA). Ethnicity, as reported to the GMC at registration, was only mainly available for UK graduates. The GMC also fills in missing data on ethnicity in registered doctors in some cases via the NHS electronic staff record or when recorded when a doctor completes the National Training Survey.[25] Thus, this variable was only used when modelling CASC pass rates in this graduate group. For analytic purposes ethnicity was dichotomised into 'white' and BME. Sex and years of birth and medical registration were also obtained from the LRMP. The latter two variables were used to estimate age and years in UK practise at the time of their first CASC sitting. This would also include, in some cases, clinical experience in specialties other than psychiatry. The identification of the 'deanery' (UK administrative region

for medical postgraduate training) was undertaken, but was only available for only 700 of the doctors, mainly UK graduates, in the dataset. However, previous research indicated that the impact of deanery was relatively trivial [2] with an intraclass correlation of only 0.009 or lower for this effect. Likewise, a variance components model, using the present data, evaluating the odds of passing by deanery (where region was reported) was only 0.035 for all doctors and only 0.008 for UK graduates. This indicates a negligible effect overall. For this reason, the effect of deanery was not adjusted for in our models.

Scores for the IELTS and PLAB examination were also available from the GMC for the vast majority of doctors qualifying from outside of the EEA, referred to hereafter as 'overseas graduates' (see Figure 1). Thus, the relationship between these variables and the odds of passing the CASC at first attempt were modelled.

Statistical analyses

A series of binary logistic regressions were performed where the odds of passing the CASC at first attempt was estimated, according to each predictor variable. Initially the raw (unadjusted) relationship between the predictors and outcome were explored using a series of univariable logistic regressions. Following this a series of multivariable models were built, which included the potential confounding variables. As the main focuses of the study was estimating the impact of ethnicity and place of qualification on MRCPsych performance, all relevant confounders were placed in the multivariable models. However, it should be noted that when modelling the independent effects of the predictors specific to overseas graduates, namely the IELTS scores and PLAB performance, only one component of these two assessments were placed in the multivariable model at a time, along with the other demographic and educational variables. This is because the IELTS and PLAB components are dependent on each other. For example, PLAB part 2 can only be taken once PLAB part 1 is passed. Also, some of the subtest scores of the IELTS tend to correlate highly with each other, risking substantial multicollinearity when entering them both in the same model.

Trends over time were also explored by introducing a variable which represents the particular sitting of the CASC taken. As there are two sittings of the MRCPsych per year (spring and autumn) there were ten sittings during the study period (spring 2013 to autumn 2018). As the structure of the written component changed considerably during the study period, and this portion of the MRCPsych was not the main focus of the study, time trends were only explored in relation to the CASC.

In order to estimate the potential impact of a change in regulatory policy we conducted "simulation" studies of a change in PLAB pass mark and IELTS requirements. This was performed by dividing overseas graduates into quartiles according to their performance at the most recent attempt at the PLAB part 2 examination and comparing their odds of passing CASC at first attempt with UK graduates. Similarly, the CASC pass rates of overseas graduates who had achieved at least of eight in each language domain on IELTS were compared to UK graduates. The numbers of overseas graduates in the sample precluded more fine-grained analyses in this respect.

Missing data were relatively few, with the exception of PLAB part 2 and written papers (see Figure 1). PLAB part 2 scores were only present for 30% (n=195) of the overseas graduates. Likewise, written test scores were missing for 37% of the overseas graduates, though only 19% of UK graduates. Therefore, values for the PLAB part 2 and written paper scores were imputed, where missing for overseas graduates, using chained equations. One hundred imputed datasets were created this way, with the imputed values informed by observed values for sex, age, overall, PLAB part 1 performance at first attempt and IELTS overall speaking scores (which were relatively complete). Thus, analyses which involved either PLAB part 2 scores or written tests scores for the overseas graduates were repeated using the imputed datasets. These additional analyses were performed as a form of sensitivity analyses to assess the potential impact of the missing data on the results. Specifically the analyses of the imputed data was used to indicate the extent to which any missing data were

'missing at random'- that is the missing values were related to the observed variable values. Otherwise, listwise deletion used to deal with the variables with more sparse missing data.

All data management and analyses were conducted in Stata v14. The code is available from the lead author on request.

RESULTS

Table 1 depicts the demographic characteristics and educational performances for the different graduate groups, as well as a breakdown of the CASC pass rate, at first attempt, by sitting. As can be seen, the CASC pass rate, at first attempt, was highest in UK graduates identifying as 'white' and lowest in overseas graduates (see also later). UK graduates identifying as BME tended to be slightly younger at the time of attempt at the CASC and were modestly more likely to be male, compared to those UK qualifying doctors identifying as 'white'. In contrast, compared to UK graduates, overseas graduates were older and had been registered with the GMC for relatively longer at the time of the first sitting of the CASC. Compared to UK graduates they also had lower average scores on the written papers.

For overseas graduates IELTS scores tended to be distributed around the minimum scores required for eligibility to sit the PLAB examination. Indeed, the mean overall IELTS score for this group was only 7.4- slightly below the recently raised requirement for an overall mark of 7.5. We also noted that the mean scores (relative to pass) at first attempt for both part 1 (0.43, SD 19.59) and part 2 (3.56, SD 4.18) of the PLAB exam for this cohort was considerably lower than those reported for a more general cohort of overseas graduates registering with the GMC.[27] In this latter case, the mean scores for PLAB part 1 were 7.47 (SD 19.29) and 6.00 (SD 4.58) for PLAB part 2 at first attempt. As can be seen, overall there is a general trend for the overall CASC pass rate, at first attempt, to have increased over the study timeframe. A logistic regression of this CASC pass rate against time of sitting

(coded '1' for autumn 2013 through to '10' for spring 2018) yielded an odds ratio (OR) of 1.14 (95% confidence interval 1.11 to 1.18, p<0.001). When analysed by group, modest time trends, of borderline statistical significance, were observed in this respect for both UK (OR 1.07, 1.00 to 1.13. p=0.037) and non-UK graduates (OR 1.05, 1.00 to 1.10, p=0.043). The flow of data through the study is shown in figure 1.

[Insert Figure 1 here]

World region of qualification and self-reported ethnicity (UKGs only)				
	White UKG (n=940)	**BME UKG (n=350)**	**EEA Graduate (n=155)**	**Overseas Graduates (n=650)**
Age in years (SD) at first attempt at the Clinical Assessment of Skills and Competencies (n=2140)	31.74 (4.75)	30.69 (2.99)	34.85 (5.83)	39.02 (5.42)
Years since UK registration at first attempt at the Clinical Assessment of Skills and Competencies (n=2140)	5.52 (1.56)	5.57 (1.58)	9.29 (5.06)	14.07 (5.12)
Male (proportion (%)) (n=2140)	360/940 (38.1%)	155/350 (44.3%)	65/155 (42.0%)	305/650 (46.7%)
CASC pass rate at first attempt†	870/940 (92.9%)	280/350 (79.4%)	90/155 (57.3%)	185/650 (28.6%)
Paper 1 pass rate at first attempt† (n=545)	215/265 (81.6%)	70/95 (74.0%)	30/40 (70.0%)	95/135 (70.8%)
Paper 2 pass rate at first attempt† (n=625)	230/305 (75.6%)	80/105 (76.7%)	25/40 (65.9%)	115/160 (71.9%)
Paper 3 pass rate at first attempt† (n=750)	260/305 (84.0%)	80/115 (71.3%)	20/55 (38.9%)	105/255 (40.5%)
Paper A pass rate at first attempt† (n=595)	270/325 (84.0%)	95/120 (80.67%)	20/35 (66.7%)	70/110 (63.1%)
Paper B pass rate at first attempt† (n=950)	375/455 (82.1%)	130/170 (75.6%)	45/70(59.7%)	100/230 (42.2%)
Average score, relative to the pass mark, on written papers sat† (N=1600)	6.54 (7.57)	5.09 (8.19)	1.77 (8.33)	0.03 (8.43)
Proportion of candidates passing at first attempt for each sitting				
Sitting	**CASC pass rate at 1st attempt- all candidates†**	**CASC pass rate at 1st attempt- UK graduates†**	**CASC pass rate at 1st attempt- non-UK graduates†**	
Autumn 2013	210/425 (49.3%)	145/170 (84.1%)	65/255(26.0%)	
Spring 2014	110/175 (61.0%)	85/105 (80.0%)	25/70 (33.3%)	
Autumn 2014	170/240 (71.8%)	130/145 (89.6%)	40/95 (44.7%)	
Spring 2015	110/155 (70.1%)	80/85 (96.5%)	25/70(37.7%)	
Autumn 2015	160/205 (78.4%)	140/155 (92.2%)	20/50 (37.3%)	
Spring 2016	120/180 (66.5%)	95/110 (89.0%)	20/70 (31.4%)	
Autumn 2016	185/235 (78.8%)	160/175 (91.5%)	25/60 (40.7%)	
Spring 2017	110/150 (73.8%)	90/105 (87.5%)	20/45 (42.2%)	
Autumn 2017	180/220 (82.6%)	165/180 (91.2%)	15/35 (38.9%)	
Spring 2018	105/155 (68.9%)	90/100 (88.1%)	19/55 (32.7%)	
IELTS average scores (SD) (overseas graduates only)				
Reading (n=440)	-	-	-	7.63 (0.62)

Speaking (n=615)	-	-	-	7.53 (0.60)
Listening (n=510)	-	-	-	7.68 (0.61)
Writing (n=480)	-	-	-	7.33 (0.51)
Overall score (n=615)				7.40 (0.46)
Professional and Linguistic Board (PLAB) examination performance (overseas graduates only)				
PLAB part 1 average score at 1st attempt, relative to pass mark (n=605)	-	-	-	0.43 (19.59)
PLAB part 1 average score at pass, relative to pass mark (n=600)	-	-	-	13.35 (9.57)
Attempts at PLAB part 1 (SD) (n=610)	-	-	-	1.74 (0.95)
PLAB part 2 average score at 1st attempt (n=190)	-	-	-	3.56 (4.18)
PLAB part 2 average score at pass (n=195)	-	-	-	5.27 (2.95)
Attempts at PLAB part 2 (SD) (n=205)	-	-	-	1.39 (0.63)

Table 1. Sociodemographic characteristics and educational performance for the different graduate groups in the study sample (N=2140).

†Note that numbers have been rounded in line with UKMED statistical disclosure controls, so that percentage values will not correspond precisely to the proportions shown.[23]

Ethnicity and MRCPsych performance in UK medical graduates

As can be seen in Table 1, UK medical graduates who identified as of 'white' ethnicity had an average CASC pass rate, at first attempt, of around 93%. This contrasts with those UK graduates identifying as BME, where the average CASC pass rate at first attempt was approximately 79%. Although not the main focus of the study, it is informative to examine the pass rates, at first attempt, for the written parts of the MRCPsych for UK graduates, as this will reflect levels of clinical knowledge, as well as skills, such as examination preparation. As can be seen from Table 2, those UK graduates identifying as BME had a lower odds of passing Paper 3 (OR 0.47, 0.28 to 0.78, p=0.004) compared to non-BME UK graduates. A number of univariable trends, relating to written examination performance and demographic factors, can also be

Paper and predictor variable	Odds Ratio	lower 95% confidence interval	upper 95% confidence interval	p
Results from univariable (unadjusted) analyses				
BME vs non-BME				
Paper 1: Basic and social science and adult psychiatry (n=360)	0.64	0.37	1.11	0.11
Paper 2: Theory, pharmacology, epidemiology and research (n=410)	1.06	0.63	1.80	0.82
Paper 3: Clinical topics and critical appraisal (n=420)	0.47	0.28	0.78	0.004
Paper A: Science and theory (n=445)	0.80	0.46	1.37	0.42
Paper B: Critical appraisal and clinical topics (n=630)	0.68	0.44	1.03	0.07
Male sex				
Paper 1: Basic and social science and adult psychiatry (n=365)	0.90	0.54	1.51	0.70
Paper 2: Theory, pharmacology, epidemiology and research (n=425)	0.66	0.42	1.04	0.08
Paper 3: Clinical topics and critical appraisal (n=440)	0.90	0.55	1.45	0.66
Paper A: Science and theory (n=455)	0.53	0.32	0.87	0.01
Paper B: Critical appraisal and clinical topics (n=645)	0.76	0.52	1.13	0.17
Age at time of first sitting (years)				
Paper 1: Basic and social science and adult psychiatry (n=365)	0.98	0.93	1.03	0.40
Paper 2: Theory, pharmacology, epidemiology and research (n=425)	0.96	0.92	1.00	0.07
Paper 3: Clinical topics and critical appraisal (n=435)	0.95	0.91	1.00	0.03
Paper A: Science and theory (n=455)	1.01	0.96	1.08	0.66
Paper B: Critical appraisal and clinical topics (n=645)	0.94	0.91	0.98	0.001
Clinical experience at time of first sitting (years)				
Paper 1: Basic and social science and adult psychiatry (n=365)	1.05	0.89	1.23	0.57
Paper 2: Theory, pharmacology, epidemiology and research (n=425)	1.01	0.92	1.11	0.81
Paper 3: Clinical topics and critical appraisal (n=435)	0.99	0.90	1.09	0.81
Paper A: Science and theory (n=455)	1.07	0.93	1.23	0.32
Paper B: Critical appraisal and clinical topics (n=645)	0.94	0.88	1.01	0.10
Results from multivariable (adjusted) analyses				
BME vs non-BME				
Paper 1: Basic and social science and adult psychiatry (n=360)	0.58	0.33	1.03	0.06
Paper 2: Theory, pharmacology, epidemiology and research (n=410)	0.96	0.56	1.65	0.89
Paper 3: Clinical topics and critical appraisal (n=420)	0.44	0.26	0.74	0.002
Paper A: Science and theory (n=445)	0.82	0.47	1.42	0.48
Paper B: Critical appraisal and clinical topics (n=630)	0.61	0.40	0.94	0.03
Male sex				
Paper 1: Basic and social science and adult psychiatry (n=360)	0.94	0.56	1.59	0.83
Paper 2: Theory, pharmacology, epidemiology and research (n=410)	0.65	0.41	1.03	0.07
Paper 3: Clinical topics and critical appraisal (n=420)	0.89	0.54	1.47	0.66
Paper A: Science and theory (n=445)	0.57	0.35	0.94	0.03
Paper B: Critical appraisal and clinical topics (n=630)	0.80	0.54	1.20	0.28
Age at time of first sitting (years)				
Paper 1: Basic and social science and adult psychiatry (n=360)	0.96	0.91	1.01	0.15
Paper 2: Theory, pharmacology, epidemiology and research (n=410)	0.95	0.90	1.00	0.03
Paper 3: Clinical topics and critical appraisal (n=420)	0.94	0.89	0.99	0.02
Paper A: Science and theory (n=445)	0.99	0.93	1.07	0.88
Paper B: Critical appraisal and clinical topics (n=630)	0.94	0.90	0.98	0.007
Clinical experience at time of first sitting (years)				
Paper 1: Basic and social science and adult psychiatry (n=360)	1.07	0.87	1.32	0.50

Paper 2: Theory, pharmacology, epidemiology and research (n=410)	1.01	0.85	1.20	0.89
Paper 3: Clinical topics and critical appraisal (n=420)	1.06	0.89	1.27	0.53
Paper A: Science and theory (n=445)	1.00	0.82	1.21	1.00
Paper B: Critical appraisal and clinical topics (n=630)	0.98	0.85	1.12	0.77

Table 2. Results of a series of logistic regressions predicting passing the written papers of the MRCPsych at first sitting for UK medical graduates. Both univariable (raw) and multivariable (adjusted) results are provided. Note that numbers have been rounded in line with UKMED statistical disclosure controls.[23]

observed. When the influence of these were controlled for in the logistic regression analysis the independent effect of self-reported ethnicity remained, if anything, increasing in magnitude (OR 0.44, 0.26 to 0.74, p=0.002). Other, independent influences of demographic factors on the written exam pass rates can also be observed, though are not the focus of the present study.

The results of logistic regression models for predicting CASC performance in UK medical graduates are shown in Table 3. As can be seen, BME status was a both a univariable (OR 0.30, 0.21 to 0.42, p<0.001) and independent predictor of reduced odds of passing the CASC at first attempt (OR 0.36, 0.23 to 0.55, p<0.001). The other educational and demographic factors were statistically significant univariable (p<0.05) predictors. However, in the multivariable model, only average, standardised performance at the written components of the examination, male sex and younger age at sitting were independent predictors of passing the CASC at first attempt.

Predictor	**Unadjusted (raw) ORs (95% confidence interval)**	**P**	**Adjusted ORs (95% confidence interval)**	**P**
Ethnicity (BME vs non-BME) (n=1290)	0.30 (0.21 to 0.42)	<0.001	0.36 (0.23 to 0.56)	<0.001
Performance at written parts (z-score) (n=1075)	1.05 (1.02 to 1.07)	<0.001	1.03 (1.01 to 1.06)	0.02
Male sex (n=1330)	0.55 (0.39 to 0.78)	0.001	0.49 (0.32 to 0.74)	0.001
Age at first sitting (n=1330)	0.95 (0.92 to 0.98)	<0.001	0.95 (0.91 to 0.99)	0.02
Clinical experience at first sitting (n=1330)	0.94 (0.89 to 0.99)	0.03	0.96 (0.84 to 1.10)	0.56
Time of sitting (more recent vs earlier) (n=1330)	1.07 (1.00 to 1.13)	0.04	1.04 (0.95 to 1.13)	0.38

Table 3. Results of a series of logistic regressions with the odds of passing CASC at first attempt in UK medical graduates only. Note, the number for the univariable analyses are shown (n) and varied according to data completeness, with n=1045 for the multivariable analysis (note, that numbers have been rounded in line with UKMED statistical disclosure controls [23]). Both unadjusted (raw) and adjusted results are shown.

World region of medical qualification and MRCPsych performance

The overall pass rates for non-UK graduates, especially for those from outside the EEA ('overseas') , were generally considerably lower for first attempts at both the written and clinical components of the MRCPsych (Table 1). The results of the logistic regression models, for predicting success at first sitting for the written components of the examination, are shown in Table 4. The results from univariable (unadjusted) analyses, with the main effect of interest being world region of qualification (UK vs EEA vs overseas graduates) are shown alongside those from a multivariable model, where the effects were adjusted for sex, age and clinical experience (years) at time of sitting. It can be seen that markedly lower pass rates at first attempts at written Papers '3' and 'B' are present for non-UK compared to UK medical graduates. The coefficients for these effects change little, once the potential influence for sex, age and experience are accounted for. Thus, it appears, that the most marked differences between non-UK and UK medical graduates are observed for those written examination components with a procedural skills component, that may involve, for example, some knowledge of statistical methods and critical appraisal of relevant research.

Few differences between EEA and overseas graduates were observed, in terms of performance on the written components of the MRCPsych. Indeed, as can be seen in Table 4, the only independent, statistically significant effect was observed for Paper B, with overseas doctors less likely to pass at first attempt than their EEA graduate counterparts.

Paper and graduate group comparison	N	Unadjusted ORs (95% confidence intervals)	p	Adjusted ORs (95% confidence intervals)	p
Paper 1 Basic and social science and adult psychiatry					
EEA vs UK graduates	405	0.59 (0.29 to 1.22)	0.15	0.60 (0.27 to 1.30)	0.20
Overseas vs UK graduates	505	0.61 (0.39 to 0.96)	0.03	0.41 (0.21 to 0.80)	0.009
Overseas vs EEA graduates	175	1.04 (0.48 to 2.24)	0.92	0.86 (0.36 to 2.07)	0.74
Paper 2 Theory, pharmacology, epidemiology and research					
EEA vs UK graduates	465	0.60 (0.30 to 1.19)	0.14	0.48 (0.23 to 1.01)	0.05
Overseas vs UK graduates	585	0.79 (0.53 to 1.20)	0.27	0.77 (0.43 to 1.40)	0.39
Overseas vs EEA graduates	200	1.33 (0.64 to 2.75)	0.45	1.15 (0.52 to 2.56)	0.72
Paper 3 Clinical topics and critical appraisal					
EEA vs UK graduates	490	0.15 (0.08 to 0.28)	<0.001	0.14 (0.07 to 0.28)	<0.001
Overseas vs UK graduates	695	0.16 (0.12 to 0.23)	<0.001	0.16 (0.10 to 0.27)	<0.001
Overseas vs EEA graduates	310	1.07 (0.59 to 1.95)	0.83	0.96 (0.51 to 1.84)	0.91
Paper A Science and theory					
EEA vs UK graduates	485	0.41 (0.19 to 0.88)	0.02	0.38 (0.17 to 0.85)	0.02
Overseas vs UK graduates	565	0.35 (0.22 to 0.55)	<0.001	0.33 (0.16 to 0.70)	0.004
Overseas vs EEA graduates	145	0.85 (0.38 to 1.94)	0.71	0.80 (0.29 to 2.19)	0.67
Paper B Critical appraisal and clinical topics					
EEA vs UK graduates	715	0.37 (0.22 to 0.61)	<0.001	0.39 (0.23 to 0.67)	0.001
Overseas vs UK graduates	875	0.18 (0.13 to 0.25)	<0.001	0.22 (0.14 to 0.36)	<0.001
Overseas vs EEA graduates	305	0.49 (0.29 to 0.84)	0.01	0.51 (0.27 to 0.93)	0.03

Table 4. Results of a logistic regression modelling odds of passing the written parts of the MRCPsych at first attempt in UK and non-UK medical graduates. Both unadjusted (raw) results, as well as adjusted results, controlling for the influence of sex, age and experience at time of sitting and average, standardised performance in the written components are

shown. Note that numbers have been rounded in line with UKMED statistical disclosure controls.[23]

The results for the models predicting the odds of passing the CASC at first attempt in trainee psychiatrists are shown in Table 5. As can be seen, overseas medical graduates have much lower odds of passing the CASC at first sitting compared to EEA graduates, who, in turn, how much lower odds of passing compared to UK graduates. These effects relating to world region of primary medical qualification change little after conditioning on the potential confounding variables (performance in the written components, sex, age, experience and time of sitting). It can also be seen that, in general, males, and those registered with the GMC for longer also have relatively reduced odds of passing at first attempt, even after controlling for the influence of potential confounding variables. Taking the CASC more recently is also associated with increased odds of success at first attempt. Increasing age is associated with reduced odds of passing the CASC at first attempt on univariable analysis. However, this effect is reduced to virtually zero after conditioning on the other demographic variables, time of sitting and performance in the written component of the MRCPsych.

As can be seen, Table 5 also depicts the results from univariable and multivariable models exploring the predictors of CASC outcome at first attempt specific to overseas medical graduates; namely performance on the PLAB examination and IELTS English fluency test. For overseas graduates, higher ratings on the reading, listening and overall IELTS components of the test were significantly associated, on univariable analysis, with relatively higher odds of passing the CASC at first attempt. However, once the IELTS scores were conditioned on the demographic variables, performance in the written components of the MRCPsych and time of the sitting in the multivariable analysis this pattern changed somewhat. That is, in the multivariable model only the IELTS speaking, listening and overall ratings were statistically significantly related with increased odds of passing the CASC at first attempt.

Performance at part 1 of the PLAB examination was not associated with subsequent performance on the CASC, either on univariable or multivariable analyses. However, better performance at part 2 (the practical component) of the PLAB examination was associated with an increased odds of passing the CASC at first attempt. This was true for both the initial attempt and the most recent score (relative to the pass mark that sitting) for the examination. Multiple attempts at PLAB part 2 were also associated with an increased odds of passing the CASC at first attempt, though the effect for sitting the PLAB part 2 three or more times (versus passed first time) became statistically non-significant once the influence of potential confounding variables were controlled for in the multivariable model.

Predictor	**N**	**Raw ORs (95% CIs)**	**p**	**N**	**Adjusted‡ ORs (95% CIs)**	**p**
EEA vs UK graduates	1490	0.17 (0.12 to 0.24)	<0.001	1185	0.25 (0.15 to 0.40)	<0.001
Overseas vs UK graduates	1980	0.05 (0.04 to 0.06)	<0.001	1485	0.07 (0.05 to 0.11)	<0.001
Overseas vs EEA graduates	810	0.30 (0.21 to 0.43)	<0.001	525	0.36 (0.22 to 0.58)	<0.001
Male sex	2140	0.69 (0.57 to 0.83)	<0.001	1600	0.65 (0.50 to 0.85)	0.001
Age (years) at first sitting	2140	0.85 (0.83 to 0.86)	<0.001	1600	1.00 (0.97 to 1.04)	0.94
Experience (years)	2100	0.79 (0.77 to 0.81)	<0.001	1600	0.81 (0.78 to 0.85)	<0.001
Average performance in written exams (score relative to pass mark)	1600	1.09 (1.07 to 1.11)	<0.001	1600	1.06 (1.05 to 1.08)	<0.001
Time of sitting (more recent vs earlier)	2135	1.14 (1.11 to 1.18)	<0.001	1600	1.07 (1.02 to 1.13)	0.007
IELTS performance (overseas graduates only)						
Reading	440	1.49 (1.08 to 2.05)	0.02	285	1.20 (0.80 to 1.81)	0.38
Speaking	615	1.32 (0.99 to 1.76)	0.06	385	1.72 (1.16 to 2.54)	0.007
Listening	510	1.86 (1.36 to 2.54)	<0.001	325	1.59 (1.05 to 2.39)	0.03
Writing	480	0.99 (0.67 to 1.47)	0.97	300	1.00 (0.57 to 1.74)	1.00
Overall score	615	2.22 (1.52 to 3.23)	<0.001	385	1.97 (1.19 to 3.27)	0.009
PLAB examination performance (overseas graduates only)						
PLAB part 1, relative to pass mark, first attempt	605	1.01 (1.00 to 1.01)	0.28	380	1.00 (0.99 to 1.01)	0.69
PLAB part 1, relative to pass mark, at pass	600	1.02 (1.00 to 1.04)	0.09	380	1.01 (0.98 to 1.03)	0.52
PLAB part 1 sat twice vs passed first time	610	0.90 (0.59 to 1.37)	0.62	385	1.17 (0.69 to 2.00)	0.56
PLAB part 1 sat thrice vs passed first time	610	0.96 (0.54 to 1.72)	0.89	385	1.12 (0.52 to 2.40)	0.77
PLAB part 1 sat four or more times vs passed first time	610	0.46 (0.21 to 1.02)	0.06	385	0.63 (0.24 to 1.62)	0.34
PLAB part 2, relative to pass mark, first attempt	190	1.26 (1.15 to 1.39)	<0.001	170	1.22 (1.10 to 1.36)	<0.001
PLAB part 2, relative to pass mark, at pass	195	1.32 (1.17 to 1.48)	<0.001	170	1.27 (1.12 to 1.45)	<0.001
PLAB part 2 sat twice vs passed first time	205	0.29 (0.12 to 0.68)	0.005	180	0.33 (0.13 to 0.83)	0.02

PLAB part 2 sat three or more times vs passed first time	205	0.11 (0.01 to 0.85)	0.03	180	0.13 (0.02 to 1.09)	0.06

Table 5. Results of a series of logistic regressions modelling the odds of passing the CASC at first sitting in UK and non-UK medical graduates. Note that numbers used in the analyses have been rounded in line with UKMED statistical disclosure controls.[23]

‡ Adjusted for average (standardised performance in written parts of exam, sex, age, years of experience and the time of the exam sitting (later vs earlier).

As a form of sensitivity analysis for the potential impact of missing data, t[23]he multivariable (adjusted) model for overseas vs UK graduates was re-run with missing written paper scores multiply imputed. The results from imputed and non-imputed datasets showed little difference. Likewise, the final multivariable model for the overseas graduates (see Table 5) was re-run in a dataset where the vast majority of the missing average written examination and PLAB part 2 scores were imputed (in all but 60 cases). The effect of PLAB part 2 score, at first sitting, was somewhat attenuated in the imputed data, though remained statistically significant (OR 1.07, 1.01 to 1.14, p=0.02). The multivariable (adjusted) model for overseas vs UK graduates was re-run with missing written paper scores imputed. The results from imputed and non-imputed datasets showed little difference.

As in previous research, we also simulated raising the standards for passing the PLAB part 2 examination.[27] For those overseas graduates with these scores reported we compared to overseas graduates who had passed the examination with a score, relative to the pass mark, placing them in the top quartile in this sample. The difference in odds of passing the CASC, compared with UK graduates diminished somewhat for these high performing PLAB candidates but remained substantial (OR 0.16, 0.09 to 0.29, p<0.001) with the respective pass rates being 56.3% (top PLAB scorers) vs 89.0% (UK graduates). Likewise when we compared to those who scored 8 or more on their overall IELTS test we noted some very modest reduction in difference with UK graduates CASC pass rates (OR 0.10, 0.07 to 0.15, p<0.001).

DISCUSSION

Our results showed marked differences in pass rates at the first sitting at the CASC between UK medical graduates, according to their self-reported ethnicity(BME versus non-BME). These differences were not explained by performance at the written papers, which evaluated clinically relevant knowledge, or other demographic factors. A number of statistically significant ($p<0.05$) trends, relating to educational demographic factors and CASC pass rates at first attempt, were also noted, though not all of these were independent of each other. We also noted some differences in performance on the written components of the MRCPsych, according to self-reported ethnicity of the UK medical graduates. Specifically, substantial inter-group pass rate differences existed specifically for Paper 3 and, its replacement, Paper B. It is worth noting that both these written papers test procedural (skills) as well as semantic (fact-based) knowledge. For example, skills relating to the critique of scientific research papers.

Differential performance in the MRCPsych was also observed across world region of qualification. That is, considerably lower pass rates for the CASC at first attempt were observed for non-UK medical graduates, especially those who qualified from outside of the EEA, compared to those holding UK-based medical degrees. Again, a number of educational and demographic predictors were observed, and most of these had effects that were independent of each other. For those trainee psychiatrists who obtained their medical degree outside of the EEA aspects of English language fluency, as indicated by IELTS scores, and performance in part 2 (though not part 1) of the PLAB examination were also independent predictors of success at the CASC. Again, in relation to the written components of the MRCPsych, a similar pattern to that observed for ethnicity was seen, with the most marked, differences observed for the papers with procedural knowledge components (Papers '3' and 'B').

Our results almost exactly mirror the findings reported by Esmail and Roberts for the MRCGP, in terms of the findings in relation to both ethnicity and place of qualification. This is unsurprising given that both general practice and psychiatry are specialties that place a strong emphasis on inter-personal ability and communication skills, as well as other procedural skills. This may especially disadvantage those from overseas, or even those who study undergraduate medicine in the UK, but who speak English as a second language. Moreover, both general practice and psychiatry are both relatively unpopular medical career choices with a historical reliance on overseas graduates. Moreover, in terms of ARCP outcomes these were the two specialties noted to have the most marked performance differences between UK and overseas graduates. Moreover, the disparities substantially reduced when ARCP outcomes associated with postgraduate examination failure were excluded.[11] This suggested that for many overseas graduates the relevant Royal College membership exams were acting as barriers to career progression. It may also be that this relative unpopularity of psychiatry means that trainees entering such specialities are starting from a lower baseline of knowledge and skills. Indeed, UK medical graduates who choose psychiatry, on average, have lower levels of undergraduate academic achievement, compared to their peers.[28] Moreover, for general practice and medicine, McManus has previously suggested that the PLAB test may be too easy to ensure equivalence between UK and non-UK medical graduates in terms of postgraduate exam performance.[5] Indeed, in the present study we noted the relatively low IELTS and PLAB test scores achieved by the MRCPsych candidates, in relation to a more general cohort of overseas graduates.[27] It was also interesting to note that PLAB part 2 (the practical component of the PLAB) but not PLAB part 1 scores at passing were predictive of CASC performance. These findings echo those previously shown for the MRCP and MRCGP, where PLAB part 2 scores, rather than those for part 1, predicted performance in the clinical components.[5]

There are also likely to be more subtle and complex reasons underlying the differences observed in the present study. For overseas graduates there may be issues relating to culture- both of the UK in general and within the NHS in particular. Our findings also highlight a degree of disjoint between semantic (fact-based) and procedural knowledge in medicine. It is possible that some doctors in training place more importance on the acquisition of factual knowledge. However, it appears to be procedural skills that are more crucial in determining relative performance on clinical practical exams, especially where candidates have already had to pass knowledge tests. Indeed, in one study performance at the MRCGP clinical skills assessment in overseas doctors was more strongly predicted by scores on a situational judgment test, evaluating interpersonal skills, than by achievement on a knowledge based test [29]. However, we do not agree with Esmail and Roberts' earlier assertion that "*Previous training experience and cultural factors… could help explain these differences between UK candidates and international medical graduates. However, these cultural factors cannot explain differences between white candidates and black and minority ethnic candidates who have trained in the UK, and who would have had similar training experiences and language proficiency*".[2] Indeed, we would highlight the observation reported from a subsequent linguistic analysis of MRCGP clinical skills assessment candidates that there were distinctive communication styles in UK graduates identifying as BME that were also observed in non-UK doctors.[4] Relative average lower academic performance in medical students and doctors from minority ethnic groups is well evidenced [30] and is probably not adequately explained by personality or obvious social factors.[31] However, it should be stressed that, even if subtle cultural influence exist, these should be understood, and not be used to perpetuate any further marginalisation of certain groups of doctors.

If bias, conscious or unconscious, is a driver behind these disparities in pass rates it could operate at a number of points in training and assessment. One previous study investigated examiner behaviour in the clinical part of the Membership of the Royal Colleges of

Physicians of the United Kingdom (MRCP(UK)). The authors' reported that examiners reporting to be from BME groups were, on average, more stringent, than those identifying as white, though did not identify any bias as such towards BME candidates.[32] Thus, it is quite possible that there is similarly no bias evident in the CASC examination scoring. However, the CASC will inevitably place a greater degree of emphasis on communication skills, compared to the PACES exam. Therefore it is important that further detailed analysis, of the sort conducted by McManus et al. in the MRCP(UK) examination, should be performed. This will help rule this potential source of bias out. It is also possible that there may be bias at work during speciality recruitment and training. Esmail and Roberts also noted that the proportion of ethnic minority trainees varied across deaneries and that selection and training processes may work against the interests of weaker recruits, encouraging a cycle of 'educational deprivation'.[2] In the present study the nature of the data precluded an exploration of these factors. This may be a focus of future research into psychiatric training. Moreover, it is likely that qualitative, rather than quantitative findings will shed further light on the underlying reasons for differential attainment in the specialty. In this regard, previous explorations of the experiences of overseas doctors in UK training highlighted the multiple challenges to career progression. These included feelings that seniors lacked trust in them, cultural differences, separation from the usual sources of social support, as well as the perceived difficulty of professional examinations.[33] A report commissioned by the GMC, as part of a review into the fairness of career progression, included interviews with 262 UK doctors in training, around half identifying as being from BME groups. The trainees were drawn from a number of specialties, including psychiatry.[34] Although it is not clear how representative the sample was, some of the quotes raise the possibilities that trainees from BME backgrounds may be subject to regular 'microaggressions' in the workplace. These can be defined as 'brief, but frequent, verbal, behavioural, or environmental indignities that communicate prejudices or insults toward any group'. They may be conscious or unconscious and subtle, though relentless in their persistence.[35] For example, one UK BME psychiatric trainee was quoted in the report as follows: "*I'm expecting to get a lower mark*

because I'm- I know it's a stupid way of thinking but actually it got to the point where I was thinking "What is it? Am I…?" I wasn't sure if it was my knowledge anymore, I wasn't sure if it was my confidence, I wasn't sure if it was my skin colour. So you start-I think it creates almost like a nasty way of thinking and how you perceive yourself to be. And if that someone's expectation of you is low, subconsciously your performance will be low." It may be that trainees from certain socio-cultural backgrounds are, on average, less able or willing to seek help with exam preparation. It is also possible that trainees from certain backgrounds may receive fewer or less intense helping behaviours from their trainers, which may additionally depend on, and interact with, their trainer characteristics. This trainer/trainee relationship could be assumed to be especially critical in the months leading up to a CASC attempt. Moreover, the relative lack of practical support is likely to be evident by underperformance in tests of procedural, rather than semantic knowledge, the latter being more easily obtained merely by private study. The report also highlighted that those involved with medical Royal College exams, especially those linked to psychiatry, medicine and general practice, felt under pressure to address these differential attainment issues since the BAPIO Court case.

Strengths and limitations

Although this was a large, national dataset a number of limitations should be borne in mind. Firstly, we did not have data relating to ethnicity for non-UK graduates. However, our previous analyses indicate that, at least for overseas graduates, the vast majority identify as BME. Therefore, in this case, recorded ethnicity tends to co-vary almost perfectly with graduation region, and does not add any additional information. Moreover, we noted that in a previous study of the clinical skills assessment for general practitioners in training very small numbers of overseas graduates identified as of 'white' ethnicity. Unlike the study by Esmail and Roberts we were not in a position to examine 2nd and 3rd sittings at the clinical examination, though the authors pointed out that few conclusions could be drawn from any

observations in this regard, as subsequent outcomes were dependent on the preceding ones.[2] Our data ruled out the possibility that BME or overseas candidates were choosing to take the CASC earlier than non-BME UK graduates in their training, and thus were more likely to fail at first attempt. In this regard overseas graduates were substantially older and more experienced at the time of the first attempt. Whilst BME UK graduates were sitting the examination with roughly the same post-registration experience as non-BME UK graduates, they were, on average, slightly younger. This could have reflected a modestly decreased tendency for BME doctors to take time out of training between the foundation years and specialty recruitment.

With the exception of self-reported ethnicity for non-UK graduates, generally, there were relatively few missing data. However, where information was less complete, i.e. for written paper scores and PLAB examination performance for overseas graduates, our sensitivity analyses using multiply imputed datasets suggested that such data were 'missing at random'. That is, their absent values were, to some extent, related to those that were observed. As such, the missing data are very unlikely to have affected any of our key findings.

Although detailed information on socio-economic status was unavailable, our findings in relation to CASC performance are very unlikely to be explained by this factor. This is because academic performance is associated with sociodemographic characteristics, such as the type of school previously attended.[36 37] Thus, once our models were adjusted for scores on the written papers the effects of such factors would almost certainly have been minimal. The numbers of doctors graduating from the EEA were relatively few in the sample. Therefore more detailed analyses relating to this group in comparison with the other two graduate groups were not performed.

We also acknowledge that the number of analyses performed may have increased the risk of a type 1 error- that is apparently statistically significant finding due to chance alone. However, the p values associated with our key findings tended to be very small (e.g. $p<0.001$), making them highly unlikely to be due to chance alone. Moreover, the results were consistent with findings from other groups of doctors in training. Nevertheless, it cannot be ruled out that some of the secondary findings, with p values nearer to 0.05, could still be the work of chance.

Implications for policy and directions for future research

Organisations, such as the medical Royal Colleges, already have training in relation to unconscious bias in place. However, the evidence for the effectiveness of such training, especially if it consists of courses of one day (or less) duration, is weak.[38] Moreover, it is likely that the reasons underlying the findings in this study are complex, and are therefore unlikely to be mediated via a simple single strategy, even the institutional level. Our simulations, albeit crude ones, relating to the PLAB and IELTS, suggest that substantial differential attainment in the CASC are likely to continue, even if the requirements for performance in these assessments are made substantially more stringent. This finding is in keeping with previous studies in relation to membership exams and ARCP that indicated that only PLAB candidates above the upper duodecile (12^{th}) for performance in the examination exhibited postgraduate academic achievement roughly equivalent to UK medical graduates.[5] [11] This situation is likely to remain unchanged even if a medical licensing assessment is introduced as a requirement for all UK registered practitioners, as planned.[39] Certainly, substantial differential attainment between home and overseas graduates in countries that have such licensing arrangements in place remain. Moreover, there is little evidence that quality of care patient safety is improved by the introduction of such medical licensing examinations.[40]

What is also unclear is whether differences in academic performance between graduate groups of doctors translate into differential clinical care and patient outcomes. In this regard there is scant evidence, though previous research reported that US citizens who obtained their medical degrees from outside of North America may have poorer patient outcomes when compared with either home graduates or non-US citizens who graduated abroad.[41] Similar UK-based research has not been performed. Nevertheless one study evaluating ARCP outcomes in 'UK overseas graduates' (UK citizens who graduated from outside of Britain) reported that such doctors were less likely to receive favourable outcomes at progression assessment than other graduate groups. It was also noted that such UK overseas graduates were more likely than overseas nationals graduating abroad to be successful at obtaining a place on a specialty training scheme.[42]Currently specialty pass rates at postgraduate exams are publically available via the GMC website.[43] However, breakdowns of pass rates by place of medical qualification, region and ethnicity for each Royal College are not currently possible for this site. In the interests of transparency this should be addressed.

Consequently our findings suggest that further, more detailed, research is performed with some urgency in order to help identify and understand any sources of bias present. In particular, analysis, such as that previously performed by McManus et al. for the practical component of the Royal College of Physicians exam could be repeated for data relating to the CASC. Such analyses could help understand CASC rater behaviour, and identify whether there are any interactions in relation to candidates depending on their ethnicity or medical graduation group. If, as in the case of the MRCP, obvious examiner bias was ruled out, then other putative sources of bias or discrimination during training could be investigated. This could focus particularly on selection and recruitment into core training for psychiatry, as well as more qualitative research focusing on trainer/trainee interactions with a focus particularly on preparation and support for the CASC. Locally, in-house training programmes exist, which include simulation-based training with feedback, which could

encourage confidence in communication and other procedural skills in both UK medical graduates and overseas doctors.[44] These can often take the form of formative objective structure clinical examinations (OSCEs).[45] There are some indications that, where such programmes have been implemented the differences in CASC pass rates between UK and non-UK graduates may diminish.[46] It was also noted that, compared to the more semantic knowledge focussed multiple-choice tests, the two written papers ('3' and 'B') which involved evaluation of critical appraisal and statistical skills showed marked performance differences between UK and non-UK medical graduates, as well as some inter-ethnic differences in UK doctors. Thus, some graduates groups working in psychiatry may benefit from specific additional support in these areas.

Conclusions

'Brexit' and a continued shortage of UK medical graduates choosing careers in psychiatry, is likely to continue our reliance on overseas doctors for NHS mental health services is likely to continue for the foreseeable future. Thus, the marked differences in CASC performance must be urgently investigated and remediated from a medical workforce perspective. Moreover, the backdrop of the high fatality rate in BME health workers during the Covid-19 pandemic and the 'Black Lives Matter' movement is likely to add impetus to efforts to address any potential unfair treatment of NHS employees. Moreover, particular in the case of UK graduates who identify as from BME groups, our findings need to be acted on from an equality perspective.

LIST OF ABBREVIATIONS

- ARCP: Annual Review of Competence Progression
- BAPIO: British Association of Physicians of Indian Origin
- BME: Black and Minority Ethnicity
- CASC: Clinical Assessment of Skills and Competencies

- CESR: Certificate of Eligibility for Specialist Registration
- EEA: European Economic Area
- GMC: General Medical Council
- HEE: Health Education England
- IELTS: International English Language Test System
- LRMP: List of Registered Medical Practitioners
- MRCGP: Membership of the Royal College of General Practitioners
- MRCPsych: Membership of the Royal College of Psychiatrists
- MRCP(UK): Membership of the Royal Colleges of Physicians of the United Kingdom
- NHS: National Health Service
- OSCE: Objective Structured Clinical Examination
- PLAB: Professional and Linguistic Assessment Board
- UKMED: UK Medical Education Database

DECLARATIONS

Ethics approval and consent to participate: As the study used routinely collected, deidentified data, ethical approval was not required. This was confirmed in writing by the Chair of the University of York Health Sciences Ethics Committee. Moreover, individual informed consent for the use of the data was not required for this study. This is because all the data used were held within the UKMED. The use of personal data in UKMED is not reliant on individual consent from data subjects, as it is not a necessary condition for processing under the General Data Protection Regulation (GDPR), which allows personal data to be used without consent where it is necessary for statutory functions. The Medical Act 1983 gives the GMC a legal responsibility to promote high standards of medical education and co-ordinate all stages of medical education. This enables the creation of the

database and the disclosure of data by other providers in compliance with the GDPR. For more information see www.ukmed.ac.uk/faq.

Availability of data and materials: This study involved the analyses of secondary data. Access to the data may be obtained from the UK Medical Education Database (www.ukmed.ac.uk) following approval of an application. Code for data management and analysis is available from the lead author on request.

Competing interests: The authors declare that they have no competing interests

Funding: PAT is supported in his research by a National Institute for Health Research Career Development Fellowship. This paper presents independent research funded by the NIHR. The views expressed are those of the authors and not necessarily those of the University of York, NHS, the NIHR or the Department of Health and Social Care.

Author contributions: PAT conceptualised the study and led on data analysis, interpretation and drafting the manuscript. LWP contributed to data analysis, interpretation and critically appraising the manuscript. Both authors approve the final version of the manuscript for submission.

Acknowledgements: The authors wish to thank Mr Daniel Smith at the General Medical Council for his help with the data extraction. The authors also wish to thank Ian Hall, Kate Lovett, Elemawusi Rockson and Kiran Sanghara from the Royal College of Psychiatrists, and Dr Indranil Chakarabarti, for their comments on an earlier version of this manuscript.

FIGURE LEGENDS

Figure 1. Flow of data and relative completeness through the study for each medical graduate group (note; in accordance with UKMED policy all numbers have been rounded to nearest five [23]).

References

1. The Royal College of Psychiatrists. Curricula Review Project & CESR. London: The Royal College of Psychiatrists, 2019.
2. Esmail A, Roberts C. Academic performance of ethnic minority candidates and discrimination in the MRCGP examinations between 2010 and 2012: analysis of data. *BMJ* 2013;347(347):f5662.
3. The High Court of Justice. Neutral Citation Number: [2014] EWHC 1416 (Admin) London: The High Court of Justice; 2014 [Available from: http://www.rcgp.org.uk/news/2014/may/~/media/Files/News/Judicial-Review-Judgment-14-April-2014.ashx accessed 16th April 2020].
4. Roberts C, Atkins S, Hawthorne K. Performance features in clinical skills assessment: Linguistic and cultural factors in the Membership of the Royal College of General Practitioners examination. London: King's College London, 2014.
5. McManus IC, Wakeford R. PLAB and UK graduates' performance on MRCP(UK) and MRCGP examinations: data linkage study. *BMJ* 2014;348 doi: 10.1136/bmj.g2621
6. Tyrer SP, Leung WC, Smalls J, et al. The relationship between medical school of training, age,gender and success in the MRCPsych examinations. *Psychiatr Bull* 2002;26:257-63.
7. Spike NA, Hays RB. Analysis by training status of performance in the certification examination for Australian family doctors. *Med Educ* 1999;33(8):612-5.
8. Boulet JR, Swanson DB, Cooper RA, et al. A comparison of the characteristics and examination performances of U.S. and non-U.S. citizen international medical graduates who sought Educational Commission for Foreign Medical Graduates certification: 1995-2004. *Acad Med* 2006;81(10 Suppl):S116-9.
9. Bessant R, Bessant D, Chesser A, et al. Analysis of predictors of success in the MRCP (UK) PACES examination in candidates attending a revision course. *Postgrad Med J* 2006;82(964):145-49.

10. General Medical Council. Progression Reports London: The General Medical Council; 2019 [Available from: https://www.gmc-uk.org/education/reports-and-reviews/progression-reports accessed 2nd December 2019].
11. Tiffin PA, Illing J, Kasim AS, et al. Annual Review of Competence Progression (ARCP) performance of doctors who passed Professional and Linguistic Assessments Board (PLAB) tests compared with UK medical graduates: national data linkage study. *BMJ* 2014;348 doi: 10.1136/bmj.g2622
12. Health Education England. Specialty recruitment: round 1 - acceptance and fill rate: Health Education England; 2019 [Available from: https://www.hee.nhs.uk/our-work/medical-recruitment/specialty-recruitment-round-1-acceptance-fill-rate accessed 2nd December 2019].
13. General Medical Council. The state of medical education and practice in the UK 2017 [Available from: https://www.gmc-uk.org/static/documents/content/SoMEP-2017-final-full.pdf accessed 16th April 2020].
14. Royal College of Psychiatrists. A fair exam, 2019 [Available from: https://www.rcpsych.ac.uk/training/exams/afairexam accessed 4th December 2019].
15. Dowell J, Cleland J, Fitzpatrick S, et al. The UK medical education database (UKMED) what is it? Why and how might you use it? *BMC Med Educ* 2018;18(1):6.
16. General Medical Council. General psychiatry curriculum London: General Medical Council; 2019 [Available from: https://www.gmc-uk.org/education/standards-guidance-and-curricula/curricula/general-psychiatry-curriculum accessed 8th October 2019].
17. British Council. What is IELTS? : British Council; 2017 [Available from: http://takeielts.britishcouncil.org/choose-ielts/what-ielts accessed 16th April 2020].
18. IMT Recruitment. Foundation competences London: IMT Recruitment; 2019 [Available from: https://www.imtrecruitment.org.uk/recruitment-process/am-i-eligible/foundation-competences accessed 16th April 2020].

19. Royal College of Psychiatrists. Preparing for exams London, 2019 [Available from: https://www.rcpsych.ac.uk/training/exams/preparing-for-exams accessed 20th September 2019].
20. Kramer A, Muijtjens A, Jansen K, et al. Comparison of a rational and an empirical standard setting procedure for an OSCE. Objective structured clinical examinations. *Med Educ* 2003;37(2):132-9.
21. Royal College of Psychiatrists. CASC guide for candidates, 2019 [Available from: https://www.rcpsych.ac.uk/training/exams/preparing-for-exams/casc-guide-for-trainees accessed 20th September 2019].
22. Kamatchi R, Bandyopadhyay S, Jainer A, et al. 3 years on: Examiners' and candidates' views on the CASC (Clinical Assessment of Skills and Competencies). *Br J Med Practitioners* 2012;5(4):a537.
23. General Medical Council. Process for completing UKMED research, 2019 [Available from: https://www.ukmed.ac.uk/documents/UKMED_research_process.pdf accessed 8th October 2019].
24. General Medical Council. National training survey briefing note 2 2018 [Available from: https://www.gmc-uk.org/-/media/documents/2018-briefing-note-2_pdf-72918521.pdf accessed 8th October 2019].
25. General Medical Council. National training surveys – Doctors in training. London: The General Medical Council, 2019.
26. McManus I, Ludka K. Resitting a high-stakes postgraduate medical examination on multiple occasions: nonlinear multilevel modelling of performance in the MRCP(UK) examinations. *BMC Med* 2012;10:60.
27. Tiffin PA, Paton LW, Mwandigha LM, et al. Predicting fitness to practise events in international medical graduates who registered as UK doctors via the Professional and Linguistic Assessments Board (PLAB) system: a national cohort study. *BMC Med* 2017;15(1):66.

28. Lambe PJ, Gale TCE, Price T, et al. Sociodemographic and educational characteristics of doctors applying for psychiatry training in the UK: secondary analysis of data from the UK Medical Education Database project. *BJPsych Bulletin* 2019;43(6):264-70.

29. Patterson F, Tiffin PA, Lopes S, et al. Unpacking the dark variance of differential attainment on examinations in overseas graduates. *Medical Educ* 2018;52(7):736-46.

30. Woolf K, Potts HWW, McManus IC. Ethnicity and academic performance in UK trained doctors and medical students: systematic review and meta-analysis. *BMJ* 2011;342

31. Woolf K, McManus I, Potts H, et al. The mediators of minority ethnic underperformance in final medical school examinations. *Br J Educ Psychol* 2013;83:135 - 59.

32. McManus I, Thompson M, Mollon J. Assessment of examiner leniency and stringency ('hawk-dove effect') in the MRCP(UK) clinical examination (PACES) using multi-facet Rasch modelling. *BMC Med Educ* 2006;6:4.

33. Woolf K, Rich A, Viney R, et al. Perceived causes of differential attainment in UK postgraduate medical training: a national qualitative study. *BMJ Open* 2016;6(11):e013429.

34. Woolf K, Rich A, Viney R, et al. Fair Training Pathways for All: Understanding Experiences of Progression. London: University College London, 2016.

35. Sue DW. Microaggressions in Everyday Life: Race, Gender, and Sexual Orientation. London: Wiley 2010:xvi.

36. Kumwenda B, Cleland JA, Walker K, et al. The relationship between school type and academic performance at medical school: a national, multi-cohort study. *BMJ Open* 2017;7(8):e016291-e91.

37. Mwandigha LM, Tiffin PA, Paton LW, et al. What is the effect of secondary (high) schooling on subsequent medical school performance? A national, UK-based, cohort study. *BMJ Open* 2018;8(5).

38. Atewologun D, Cornish T, Tresh T. Unconscious bias training: An assessment of the evidence for effectiveness- research report 113. London: Equality and Human Rights Commission, 2018.

39. Rimmer A. All medical students to sit the same exam under GMC plans. *BMJ* 2017;356:j541.

40. Archer J, Lynn N, Coombes L, et al. The impact of large scale licensing examinations in highly developed countries: a systematic review. *BMC Med Educ* 2016;16(1):212.

41. Norcini JJ, Boulet JR, Dauphinee WD, et al. Evaluating the quality of care provided by graduates of international medical schools. *Health Aff* 2010;29(8):1461-8.

42. Tiffin PA, Orr J, Paton LW, et al. UK nationals who received their medical degrees abroad: selection into, and subsequent performance in postgraduate training: a national data linkage study. *BMJ Open* 2018;8(7)

43. General Medical Council. Specialty exam pass rates London: General Medical Council; 2019 [Available from: https://webcache.gmc-uk.org/analyticsrep/saw.dll?Dashboard&PortalPath=%2fshared%2fNTS_LTD%2f_portal%2fNTS&Page=Specialty%20examinations%20(postgraduate)&Done=Dashboard%26PortalPath%3d%252fshared%252fNTS_LTD%252f_portal%252fNTS%26Page%3dMain%2520Menu%26ViewState%3d8gi1qd1nhntil3ku7n5dp7l5be accessed 8th October 2019].

44. Murray B. The use of high-fidelity simulation in psychiatric and mental health nursing clinical education. *Int J Health Sci Educ* 2014;2(1)

45. Pugh D, Desjardins I, Eva K. How do formative objective structured clinical examinations drive learning? Analysis of residents' perceptions. *Med Teach* 2018;40(1):45-52

46. Boylan J. Increasing CASC pass rates follwoing implemtation of a programme of high-fidelity simulation training for core trainees in psychiatry. Personal communication, 2019.