BMJ Open

Getting the right balance? A mixed logit analysis of the relationship between UK training doctors' characteristics and their specialties using the 2013 National Training Survey

Idaira Rodriguez Santana, Martin Chalkley

To cite: Rodriguez Santana I, Chalkley M. Getting the right balance? A mixed logit analysis of the relationship between UK training doctors' characteristics and their specialties using the 2013 National Training Survey. *BMJ Open* 2017;**7**:e015219. doi:10.1136/bmjopen-2016-015219

► Prepublication history for this paper is available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2016-015219).

Received 5 December 2016
Revised 30 May 2017
Accepted 16 June 2017

Centre for Health Economics, University of York, York, UK

Correspondence to
Idaira Rodriguez Santana; idaira.rodriguezsantana@york.ac.uk

ABSTRACT

Objective To analyse how training doctors' demographic and socioeconomic characteristics vary according to the specialty that they are training for.

Design Descriptive statistics and mixed logistic regression analysis of cross-sectional survey data to quantify evidence of systematic relationships between doctors' characteristics and their specialty.

Setting Doctors in training in the United Kingdom in 2013.

Participants 27 530 doctors in training but not in their foundation year who responded to the National Training Survey 2013.

Main outcome measures Mixed logit regression estimates and the corresponding odds ratios (calculated separately for all doctors in training and a subsample comprising those educated in the UK), relating gender, age, ethnicity, place of studies, socioeconomic background and parental education to the probability of training for a particular specialty.

Results Being female and being white British increase the chances of being in general practice with respect to any other specialty, while coming from a better-off socioeconomic background and having parents with tertiary education have the opposite effect. Mixed results are found for age and place of studies. For example, the difference between men and women is greatest for surgical specialties for which a man is 12.121 times more likely to be training to a surgical specialty (relative to general practice) than a woman (p-value<0.01). Doctors who attended an independent school which is proxy for doctor's socioeconomic background are 1.789 and 1.413 times more likely to be training for surgical or medical specialties (relative to general practice) than those who attended a state school (p-value<0.01).

Conclusions There are systematic and substantial differences between specialties in respect of training doctors' gender, ethnicity, age and socioeconomic background. The persistent underrepresentation in some specialties of women, minority ethnic groups and of those coming from disadvantaged backgrounds will impact on the representativeness of the profession into the future. Further research is needed to understand how the processes of selection and the self-selection of applicants into specialties gives rise to these observed differences.

Strengths and limitations of this study

- ► Results are derived from a comprehensive survey with a 97.7% response rate and account for the role of doctors' socioeconomic background in the specialty allocation process.
- ► Multinomial mixed logistic regression accounts for multiple covariates to be correlated and to control for doctors' unobserved heterogeneity.
- ► Access to data was limited by confidentiality requirements of the data sharing agreement.
- ► The survey data used omits some potentially important control variables such as medical school attended and academic performance of training doctors.
- ► Only outcomes of a complex specialty allocation process involving application, selection and acceptance are observed, not the process itself.

INTRODUCTION

Becoming a medical practitioner in the UK is a competitive process and represents a substantial investment of time and financial resources, much of that funded out of taxation. The outcome determines the composition of the medical profession. There is growing concern that the profession should reflect not only appropriate skills but a balance of socioeconomic background, gender and ethnicity[1] to be representative of the society the doctors serve. Achieving a greater balance could improve patient outcomes[2] and foster public health policies targeted at deprived and minority groups.[3]

Evidence has been accumulating regarding imbalance in one or other of gender, ethnicity and socioeconomic background across specialties. For example, despite the increase in the number of women entering the medical profession in the last three decades,[4] there exists a large gender difference in the distribution of doctors across specialties[4–8] and as a result, women now predominate in paediatrics, obstetrics or general practice but are a minority in surgery or radiology.

While there is no direct evidence regarding the distributions of ethnicity and socioeconomic background across specialties in the UK, studies[9 10] have shown that applicants from disadvantaged and/or from non-white ethnic backgrounds have less probability of receiving offers from some medical schools, which is an important determinant of specialty allocation.[11] There is also evidence that national and overseas educated doctors have different application patterns[12] and that overseas educated doctors have restricted access[13] to the most popular specialty training posts. This restriction may create an underclass within the NHS.[14 15]

There is thus a patchwork of evidence indicating that specialties may be unbalanced in regard to the gender, ethnicity and socioeconomic background of their constituent doctors but no overall view of how these imbalances relate to each other. A fundamental problem is that characteristics such as gender and ethnicity may be correlated so that an apparent gender imbalance can in part or in whole be accounted for by an ethnicity imbalance or vice versa.

We analyse data from the 2013 National Training Survey to examine the distribution of doctors across specialties along with their gender, ethnicity and socioeconomic background. We consider whether there are imbalances in regard to any one of the demographic or socioeconomic covariates, holding the other characteristics constant, by means of a mixed logistic regression model. Our study provides an evidence base for stimulating debate and discussion regarding the possible need to intervene in doctors' training in the UK to redress these imbalances across specialties. A better understanding of how individuals are assigned to specialties is a necessary precondition for the formulation of effective strategies to ensure greater representativeness across medical specialties. With the increasing role of women in the medical workforce, the larger dependence of the UK on overseas educated doctors and the desire to widen access to those coming from deprived backgrounds, the need for that knowledge is urgent.

DATA AND METHODS

Data and variables

The General Medical Council (GMC) National Training Survey (NTS) is a cross-sectional survey carried out each year. From 2013 it included questions about doctors' socioeconomic background.[16] The survey has a high response rate, 97.7% for 2013, which translates to a total of 52 797 responding individuals out of 54 055 who were eligible.[17] Due to the commitment to confidentiality by the GMC, our study is restricted to individuals who are not unique in respect of the combination of their characteristics and is focused on 40 889 doctors. To establish whether there is probable bias from the omission of some individuals we compared the descriptive statistics for the main demographic and socioeconomic characteristics of the complete survey[16 17] and our sample, finding that differences between our sample and population mean values are all smaller than three percentage points. The 13 359 doctors carrying out foundation training were excluded from the analysis since they had not selected their specialty, resulting in 27 530 doctors in the analysis sample. These were divided into two groups for analysis: a *general sample* containing all doctors in specialty training and a *UK sample* composed of 18 588 who attended both secondary and university education in the UK.

For each individual there is information on their demographic and socioeconomic characteristics which we encoded as categorical or binary dummy variables: variable *man* was assigned the value one if the doctor is a man; BME has the value one for black and minority ethnic doctors; age was given in four bands, <30, 30–39, 40–49 and 50+, which we merged into two groups and defined the variable *mature* to take the value one if the individual is 40 years old or older and variable *UK University* is equal to one if a doctor completed their secondary and medical undergraduate studies in the UK. For UK graduates, there is additional information concerning parental education and socioeconomic proxies. The variable *parent uni* takes the value one if at least one parent has tertiary education. The variables *state*, *grammar* and *independent* take the value one according to the type of secondary school attended. S*tate* school is the omitted category in the multivariate analysis. Following previous analyses[18] school type is used as a proxy for socioeconomic background. In the United Kingdom approximately 7% of pupils attend independent schools, of which only 1% receive means-tested scholarships. Finally, the variable *income sup* is assign the value one if the doctor's household received income support at any point in their childhood.

Each doctor could be assigned to one of 13 categories of training according to their specialisation. We reduced this categorisation to six *specialties* to group options that have the same core training or that can be regarded as close substitutes.[19] The resulting specialties analysed are:

1. acute care, emergency medicine and anaesthetics (ACEM);
2. general practice (GP);
3. surgical (SUR);
4. hospital-based specialties including medical specialties, paediatrics and childcare, medical-surgical specialties (ie, obstetrics and gynaecology and ophthalmology), and occupational medicine (HBS);
5. psychiatry (PSY);
6. others including pathology, radiology and public health (OTH).

Methodology

The specialties associated with doctors in NTS 2013 describe the outcome of the allocation of a doctor to a training post. We only observe the result of a complex selection process involving application, selection and acceptance, not the process itself, but seek to examine whether the process operates so as to sort doctors according to their demographic or socioeconomic characteristics. To establish evidence of systematic relationships between doctors' characteristics and specialty allocation in the presence of correlation between these characteristic we estimate individual-level multivariate analysis, by means of mixed logit regression.[20 21] Since specialties as defined are mutually exclusive categories a multinomial logit approach gives a natural means of establishing the effect of an individual's characteristics on the probability of observing them in one specialty, conditional on fixing their other characteristics. A mixed model permitted us to relax some of the strong distributional assumptions implied by a fixed coefficient approach and allows for the estimation of a multinomial logit model with unobserved heterogeneity.[21] We capture the heterogeneity by allowing the constant term in the model to vary across individuals following a normal distribution and allowing the estimates from the different alternatives to be correlated. Moreover, since covariates with missing observations account for 0.5% and 5.7% of our *general* and *UK sample* respectively, we proceeded as if data were missing completely at random (MCAR) and based our analysis on a sample of complete observations.

RESULTS

Descriptive statistics

Tables 1 and 2 show the distributions of individuals' characteristics by specialty for the *general sample* and the *UK sample* respectively. If the allocation process is unaffected by demographic and socioeconomic characteristics we would expect a similar distribution of characteristics in every specialty since in that case each specialty would appear as a random sample from the overall population of doctors in training. The figures show a different picture.

In terms of gender, 45.49% of the total sample consists of men but in surgical specialties men make up 78.38% of the total while in general practice they constitute 30.72%. In terms of ethnicity, the greatest deviations from the overall percentage of BME doctors (41.05%) are observed for ACEM (22.85%) and for PSY (56.21%). Similar differences emerge when comparing UK-educated and overseas educated doctors. For example, the largest number of overseas students is observed in psychiatry (54.19%) and the smallest in ACEM (18.25%) and SUR (24.7%) respectively. Table 2 concerns the *UK sample* and there is additional information on socioeconomic variables. Overall, doctors have attended an *independent* school in a larger proportion (35.31%) than the general UK population. There is again an uneven distribution across specialties; SUR being the group with the largest representation (42.19%) and GP the smallest (29.93%). We observe the opposite for *state* school with the largest representation in GP (44.57%) and the smallest from SUR doctors (32.43%). Other socioeconomic characteristics present in the data are parental education and income support. The means across all doctors are 65.83% and 10.23% respectively. For these variables, there is relatively little variability across specialties.

Regression results

In the regression tables we report the maximum likelihood coefficient estimates (MLE), its associated z-score, the implied odds ratio (OR) and the associated 95% confidence interval (95% CI). We set GP as the omitted category; it is the single largest category and recurrent problems in recruitment to GP make it a relevant object of comparison. Individual characteristics are captured by the dummy variables described in the section 'Data and variables'.

Table 3 shows the results for all doctors in training. In terms of gender, we observe a positive significant effect of the variable *man* for all of the alternative specialties to GP, confirming the relationships observed in the descriptive statistics. The greatest effect is associated with SUR for which male doctors are 12.121 times more likely to be allocated in a surgical specialty relative to the GP option. The variable BME has a negative coefficient estimate for all the categories with respect to the base outcome. In this case, the greatest effect is found in the ACEM category, with an OR of 0.289. The regression estimates for the variable *mature* also coincide with the results observed in the descriptive statistics. The greatest positive significant effect was found for PSYwith an OR of 2.918. The rest of the categories have negative coefficients and ORs less than one, implying for example that being 40 years old or older reduces the probability of being based in any of these specialties relative to GP. Finally, the variable *UK university* indicates a positive significant effect for the surgical specialties such that a UK-educated doctor is 2.072 times more likely to appear in this specialty with respect to GP. The greatest negative effect is found for PSY with an OR of 0.209.

Table 4 shows the results for the UK-educated doctors; those who completed both secondary school education and undergraduate studies in the UK. The estimated coefficients and ORs for the variables *man*, BME and *mature* are of the same signs and of similar magnitudes to those shown in table 3. In respect of schooling variables (*state* school is omitted category), which were used to proxy socioeconomic background, we observe positive and significant estimates and ORs greater than one for all specialties with respect to GP. The largest effect is found for surgical specialties for which doctors who attended an *independent* school are 1.789 times more likely to be in surgical specialties than those who attended a *state* school relative to GP. The smallest positive effect is associated to PSY with an associated OR of 1.301. Overall, having attended an *independent* or *grammar* school reduces the probability of GP with respect to any other specialty. Finally, the estimates for parental education are positive but modest compared with the schooling estimates. Here the greatest effect in magnitude is related to HBS and

Table 1 Characteristics of the doctors in the *general sample* (n=27 530)

	All sample		ACEM		GP		SUR		HBS		PSY		OTH	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Gender														
Man	12 524	45.49	1761	53.22	2134	30.72	3104	78.38	4030	40.2	914	42.81	581	50.26
Women	15 006	54.51	1548	46.78	4812	69.28	856	21.62	5994	59.8	1221	57.19	575	49.74
Ethnicity														
BME	11 301	41.05	756	22.85	2890	41.61	1628	41.11	4322	43.12	1200	56.21	505	43.69
White	16 215	58.9	2553	77.15	4056	58.39	2328	58.79	5692	56.78	935	43.79	651	56.31
Missing	14	0.05					4	0.1	10	0.1				
Age														
40 years old or more	1404	5.1	67	2.02	344	4.95	184	4.65	426	4.25	366	17.14	17	1.47
Less than 40 years old	26 126	94.9	3242	97.98	6602	95.05	3776	95.35	9598	95.75	1769	82.86	1139	98.53
Place of studies														
UK	19 425	70.56	2705	81.75	5226	75.24	2982	75.3	6767	67.51	978	45.81	767	66.35
Overseas	8105	29.44	604	18.25	1720	24.76	978	24.7	3257	32.49	1157	54.19	389	33.65

ACEM, acute care, emergency medicine and anaesthetics; BME, black and minority ethnic; GP, general practice; HBS, hospital-based specialties; OTH, pathology, radiology and public health; PSY, psychiatric specialties; SUR, surgical specialties.

Table 2 Characteristics of the doctors in the *UK sample* (n=19 425)

	All sample		ACEM		GP		SUR		HBS		PSY		OTH	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Gender														
Man	8081	41.6	1369	50.61	1393	26.66	2262	75.86	2334	34.49	361	36.91	362	47.2
Women	11 344	58.4	1336	49.39	3833	73.34	720	24.14	4433	65.51	617	63.09	405	52.8
Ethnicity														
BME	4909	25.27	339	12.53	1432	27.4	926	31.05	1709	25.25	264	26.99	239	31.16
White	14 504	74.67	2366	87.47	3794	72.6	2052	68.81	5050	74.63	714	73.01	528	68.84
Missing	12	0.06					4	0.13	8	0.12				
Age														
40 years old or older	261	1.34	8	0.3	71	1.36	56	1.88	63	0.93	53	5.42	10	1.3
Less than 40 years old	19 164	98.66	2697	99.7	5155	98.64	2926	98.12	6704	99.07	925	94.58	757	98.7
School type														
State	7490	38.56	1030	38.08	2329	44.57	967	32.43	2522	37.27	370	37.83	272	35.46
Grammar	4414	22.72	610	22.55	1166	22.31	610	20.46	1627	24.04	236	24.13	165	21.5
Independent	6859	35.31	962	35.56	1564	29.93	1258	42.19	2430	35.91	340	34.76	305	39.77
Missing	662	3.41	103	3.81	167	3.19	147	4.93	188	2.78	32	3.27	25	3.26
Parental university														
Yes	12 788	65.83	1747	64.58	3237	61.94	1978	66.33	4649	68.7	669	68.4	508	66.23
No	6000	30.89	855	31.61	1827	34.96	859	28.81	1943	28.71	282	28.83	234	30.51
Missing	637	3.28	103	3.8	162	3.1	145	4.87	175	2.49	27	2.76	25	3.26
Income support														
Yes	1988	10.23	282	10.43	581	11.12	297	9.96	627	9.27	101	10.33	100	13.04
No	15 609	80.36	2157	79.74	4146	79.33	2377	79.71	5541	81.88	795	81.29	593	77.31
Missing	1828	9.41	266	9.83	499	9.55	308	10.33	599	8.85	82	8.38	74	9.65

ACEM, acute care, emergency medicine and anaesthetics; BME, black and minority ethnic; GP, general practice; HBS, hospital-based specialties; OTH, pathology, radiology and public health; PSY, psychiatric specialties; SUR, surgical specialties.

Table 3 Mixed logit regression estimates for the *general sample*

	ACEM			SUR			HBS			PSY			OTH		
	MLE	OR	95% CI	MLE	OR	95% CI	MLE	OR	95% CI	MLE	OR	95% CI	MLE	OR	95% CI
Man	1.471 (5.63)***	4.354.	2.610 to 7.264	2.495 (15.2)***	12.121	8.786 to 16.722	0.507 (4.43)***	1.660	1.327 to 2.077	0.061 (−0.68)	1.063	0.891 to 1.268	0.626 (6.18)***	1.869	1.533 to 2.280
BME	−1.240 (−9.18)***	0.289	0.222 to 0.377	−0.271 (−2.44)**	0.763	0.614 to 0.948	−0.131 (−1.73)*	0.878	0.757 to 1.018	−0.071 (−0.86)	0.932	0.793 to 1.095	−0.050 (−0.56)	0.951	0.799 to 1.133
Mature	−1.021 (−4.97)***	0.360	0.222 to 0.377	−0.380 (−2.58)***	0.684	0.512 to 0.913	−0.496 (−3.89)***	0.609	0.474 to 0.782	1.071 (5.23)***	2.918	1.954 to 4.357	−1.309 (−4.18)***	0.270	0.146 to 0.499
UK university	0.221 (1.67)*	1.247	0.963 to 1.615	0.728 (7.02)***	2.072	1.690 to 2.539	−0.235(−2.87)***	0.790	0.673 to 0.928	−1.567 (−7.31)***	0.209	0.137 to 0.318	−0.655 (−2.98)***	0.520	0.338 to 0.799
Constant	−1.684 (−4.66)***	0.186	0.091 to 0.377	−2.600 (−12.14)***	0.074	0.049 to 0.113	0.446 (5.03)***	1.562	1.313 to 1.859	−1.384 (−3.26)***	0.251	0.109 to 0.576	−2.962 (−6.83)***	0.052	0.022 to 0.121
N	27 516														
Log likelihood	−41153.377														

z scores in parentheses: ***p<0.01; **p<0.05 and *p<0.1; MLE (mixed logit estimate);OR (odds ratio); CI (confidence intervals).
Man vs. woman; BME vs. non-BME; mature vs. non-mature (>40 vs.<40 years old); UK university vs. overseas educated student.
ACEM, acute care, emergency medicine and anaesthetics; BME, black and minority ethnicity; GP, general practice; HBS, hospital-based specialties; OTH, pathology, radiology and public health; PSY, psychiatric specialties; SUR, surgical specialties.

PSY, with associated ORs of 1.311 and 1.321 respectively. To achieve a parsimonious model, we excluded the variable *income sup* as we did not find statistically significant estimates of its effect in any of the specifications we tested.

DISCUSSION

Our analysis shows that in respect of socioeconomic and demographic characteristics there are substantial differences across specialties. All of the characteristics we considered constitute potentially important indicators of the representativeness of the medical profession and specialties perform very differently. Our analysis confirmed that the well known[5–8] gender gap in some specialties is also present in the cohort of training doctors in 2013. The gap is greatest between GP and SUR . Overall, surgical specialties are more male, white British and socioeconomically privileged than GP. Surgical specialties are typically highly competitive and the mean income for surgical specialties is above the 75% percentile of the distribution.[22]

In regard to schooling variables our results are new and statistically significant. They imply that doctors from better-off socioeconomic backgrounds are less likely to be based in GP than in any other specialty. In contrast to previous work in this area, we have been able to examine imbalances in both demographic and socioeconomic characteristics simultaneously using a survey that contains rich information on individuals and that has a very high response rate. These data are amenable to robust statistical methods to control for correlation between variables and allow for unobserved heterogeneity. Our results are both novel and more comprehensive than has previously been possible, however they are necessarily specific to the particular selection of doctors studied. Nevertheless, the observed distribution of doctors corresponds to historical trends and the characteristics associated with the different specialties.

As with any survey there are missing data and our study is further limited by confidentiality requirements that reduced the sample we were able to analyse. Since covariates with missing observations only account for 0.5% and 5.7% of our general and UK samples respectively, we proceeded as if data were missing completely at random (MCAR) and based our analysis on a sample of complete observations. Comparing our sample to the full survey population did not reveal any differences more than three percentage points in the means of the variables of interest. Nevertheless, there is always a possibility that our sample is biased in ways that cannot be detected by simple comparison of means. For future work on this survey it may be possible to work within the context of a 'safe' environment that would permit sharing of the complete dataset. A further limitation is that there are characteristics of individuals that are relevant to understanding their specialty allocation that are not reported in this survey. In particular, the educational background of doctors in training in respect of the medical school they

Table 4 Mixed logit regression estimates for the *UK sample*

	ACEM			SUR			HBS			PSY			OTH		
	MLE	OR	95% CI	MLE	OR	95% CI	MLE	OR	95% CI	MLE	OR	95% CI	MLE	OR	95% CI
Man	1.140 (21.73)***	3.127	2.822 to 3.466	2.165 (34.57)***	8.714	7.707 to 9.852	0.405 (6.78)***	1.499	1.334 to 1.685	0.491 (5.94)***	1.634	1.390 to 1.922	0.924 (10.43)***	2.520	2.118 to 2.998
BME	−1.145 (−16.27)***	0.318	0.277 to 0.365	−0.120 (−2.1)**	0.887	0.793 to 0.992	−0.181 (−3.9)***	0.834	0.762 to 0.914	−0.092 (−1.08)	0.912	0.771 to 1.078	0.061 (0.68)	1.063	0.892 to 1.268
Mature	−1.648 (−4.35)***	0.193	0.092 to 0.405	0.005 (0.03)	1.005	0.675 to 1.496	−0.545 (−2.75)***	0.580	0.393 to 0.856	1.488 (6.75)***	4.428	2.875 to 6.819	−0.133 (−0.38)	0.876	0.438 to 1.749
Parent uni	0.139 (2.6)***	1.149	1.035 to 1.275	0.235 (4.23)***	1.265	1.134 to 1.410	0.271 (6.1)***	1.311	1.202 to 1.430	0.278 (3.38)***	1.321	1.124 to 1.552	0.149 (1.69)*	1.161	0.976 to 1.380
Independent school	0.373 (6.42)***	1.452	1.296 to 1.627	0.582 (9.75)***	1.789	1.592 to 2.011	0.346 (7.14)***	1.413	1.285 to 1.554	0.263 (2.99)***	1.301	1.095 to 1.547	0.448 (4.76)***	1.566	1.302 to 1.883
Grammar school	0.179 (2.78)***	1.196	1.054 to 1.356	0.221 (3.27)***	1.247	1.093 to 1.424	0.254 (4.85)***	1.289	1.163 to 1.428	0.241 (2.52)**	1.272	1.055 to 1.534	0.170 (1.58)	1.186	0.960 to 1.465
Constant	−1.141 (−16.19)***	0.320	0.278 to 0.367	−2.062 (−25.15)***	0.127	0.108 to 0.149	−0.212 (−3.29)***	0.809	0.712 to 0.918	−2.514 (−12.99)***	0.081	0.055 to 0.118	−2.824 (−16.04)***	0.059	0.042 to 0.084
N	18588														
Log likelihood	−27595.305														

z scores in parentheses: ***p<0.01;**p<0.05 and *p<0.1; MLE (mixed logit estimate); OR (odds ratio); CI (confidence intervals).
Man vs. woman; BME vs. non-BME; mature vs. non-mature (>40 vs.<40 years old); parent uni vs. parent with no university education; income support vs. no income support; independent and grammar school vs. state school.
ACEM, acute care, emergency medicine and anaesthetics; BME, black and minority etnicity; GP, general practice; HBS, hospital-based specialties; OTH, pathology, radiology and public health; PSY, psychiatric specialties; SUR, surgical specialties.

attended and their academic records could potentially be confounders of the estimates of demographic and socioeconomic characteristics. Future work to address this limitation might include linking individual survey responses to administratively recorded data.

Our results can be viewed describing the outcome of a complex process of specialty allocation in which doctors apply, are selected and subsequently choose from the offers that are made to them. We have established that the outcome of this process is highly unbalanced and that some specialties exhibit a dearth of doctors with some characteristics. There are many potential causes for such an imbalance. It might be the result of differences in preferences between socioeconomic groups in terms of characteristics of the specialties, potential earnings and other non-pecuniary benefits of the alternatives and those differences might also have foundations in the secondary school, or as previous literature suggests[11] in the medical school attended. Alternatively, another determinant of the observed differences between socioeconomic and different gender groups might be due to the existence of some form of statistical discrimination.[23] The latter is a theory of inequality between groups that is based on stereotypes that do not arise from prejudice or racial and gender bias. For example, doctors might self-select themselves into the less competitive training posts, by actively not applying[6 24] or by not investing in the necessary skills to be an admissible candidate because they perceive that the chances of success are low. This then becomes a self-fulfilling prophecy.

Identifying the causes of the imbalances we have documented has considerable importance for policymakers who are concerned to redress them. Medical education is costly and in the UK relies on a substantial injection of public funds. It is therefore natural that policymakers will be concerned that the outcome of medical education reflects societal values and a key task for future research is to find the means of discriminating between the competing explanations. Therein lies the means to intervene successfully. Our research also highlights another potential cause for concern. Approximately 30% of doctors in the training scheme are educated overseas and they are very unequally distributed over specialties. From an international perspective importing doctors from low-income countries might be seen as a brain drain and in conflict with what has been termed 'ethical recruitment".[25] Additionally, previous literature[14 15] has suggested that overseas educated doctors who are allocated to less desirable training positions might be creating an 'underclass' within the NHS. This is a concern for both policymakers and the profession. Future research might usefully focus on the quality of training experience and satisfaction of those overseas doctors.

Acknowledgements We are grateful for comments and suggestions from Karen Bloor, Nigel Rice and Beatriz González López-Valcárcel and those received during presentations at the Health, Econometrics and Data Group Seminar (University of York), the 35th Spanish Health Economics Conference (Granada) and the second Conference on the Economics of the Health Workforce (Milan).

Contributors IRS managed the data and did the statistical analysis. All authors interpreted data and drafted the manuscript. All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent Not Applicable.

Ethics approval Dr Stephen Holland, chair of University of York's research governance committee approved this study by chair's action because the data are (i) extant (ii) fully anonymised and (iii) non-identifiable. The study raises no other substantive ethical issues.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement This research has been undertaken using data from the National Training Survey data set available on application from the General Medical Council. We are not permitted to share the raw data with other parties, without the advance written permission of the GMC.

REFERENCES

1. General Medical Council. *Generic standards for specialty including GP training*. London: General Medical Council, 2010.
2. Tsugawa Y, Jena AB, Figueroa JF, *et al*. Comparison of hospital mortality and readmission rates for medicare patients Treated by Male vs female physicians. *JAMA Intern Med* 2017;177:206–13.
3. Cohen JJ, Gabriel BA, Terrell C. The case for diversity in the health care workforce. *Health Aff* 2002;21:90–102.
4. McKinstry B. Are there too many female medical graduates? yes. *BMJ* 2008;336:748.
5. McManus IC, Sproston KA. Women in hospital medicine in the United Kingdom: glass ceiling, preference, prejudice or cohort effect? *J Epidemiol Community Health* 2000;54:10–16.
6. McNally SA. Competition ratios for different specialties and the effect of gender and immigration status. *J R Soc Med* 2008;101:489–92.
7. Lambert TW, Goldacre MJ, Turner G. Career choices of United Kingdom medical graduates of 2002: questionnaire survey. *Med Educ* 2006;40:514–21.
8. Goldacre MJ, Laxton L, Lambert TW. Medical graduates' early career choices of specialty and their eventual specialty destinations: UK prospective cohort studies. *BMJ* 2010;341:c3199.
9. Arulampalam W, Naylor R, Smith J. Doctor who? who gets admission offers in UK medical schools. *IZA Discussion Papers* 2005. No 1775.
10. McManus IC, Esmail A, Demetriou M. Factors affecting likelihood of applicants being offered a place in medical schools in the United Kingdom in 1996 and 1997: retrospective study commentary: league tables will help_commentary: some legal aspects arising from the study. *BMJ* 1998;317:1111–7.
11. Goldacre MJ, Turner G, Lambert TW. Variation by medical school in career choices of UK graduates of 1999 and 2000. *Med Educ* 2004;38:249–58.
12. Fazel S, Ebmeier KP. Specialty choice in UK junior doctors: is psychiatry the least popular specialty for UK and international medical graduates? *BMC Med Educ* 2009;9:77.
13. British Medical Association. Career and advice support. 2017 http://bma.org.uk/.
14. Richards T. The overseas doctors training scheme: failing expectations. *BMJ* 1994;308:1627–31.
15. Welsh C. Training overseas doctors in the United Kingdom: they must be given accurate information about their job prospects. *BMJ* 2000;321:253–4.

16. General Medical Council. National training survey 2013: socioeconomic status questions: General Medical Council. 2013 http://www.gmc-uk.org/education/national_summary_reports.asp (accessed 12 Mar 2017).
17. General Medical Council. National Training Survey 2013: key findings. 2013 http://www.gmc-uk.org/education/national_summary_reports.asp (accessed 12 Mar 2017).
18. Milburn A. 2014. *Elitist Britain?* The Social Mobility and Child Poverty Commission. https://www.gov.uk/government/news/elitist-britain-report-published. (accessed 28 Aug 2014).
19. Health Careers. Explore roles doctors: NHS. 2017 https://www.healthcareers.nhs.uk/EXPLORE-ROLES (accessed 12 Mar 2017).
20. McFadden D, Train K. Mixed MNL models for discrete response. *J Appl Econ* 2000;15:447–70.
21. Hole AR. Estimating mixed logit models using maximum simulated likelihood. *Stata J* 2007;7:388–401.
22. Morris S, Elliott B, Ma A, *et al*. Analysis of consultants' NHS and private incomes in England in 2003/4. *JRSM* 2008; 101:372–80.
23. Fang H, Moro A. *Theories of statistical discrimination and affirmative action: a survey*. USA: National Bureau of Economic Research, 2010.
24. Richardson HC, Redfern N. Why do women reject surgical careers? *Ann R Coll Surg Engl* 2000;82:290–3.
25. Lowell BL, Findlay A. Migration of highly skilled persons from developing countries: impact and policy responses. *Int Migr Papers* 2001;44.