Influences on pass rates in PLAB examinations:

PLAB 1 and 2 Annual Report 2021 - March 2022

Dr Matt Homer, Leeds Institute of Medical Education, University of Leeds

m.s.homer@leeds.ac.uk

Contents

List of figures

List of tables

Executive summary

This report investigates the association between PLAB1 and PLAB2 pass/fail outcomes and a range of candidate characteristics. The key findings are as follows:

PLAB1 – individual influences on first attempt pass rates

When comparing PLAB1 pass rates between various candidate characteristics, those that show the biggest effects are, for example:

- age (with older candidates less likely to pass – page 6),
- religion (with, for example, Christians having slightly lower pass rates compared to Muslims – page 8), and,
- PMQ country of origin (for countries with at least 100 candidate, Sri Lanka has the highest pass rate, and Russia the lowest– page 11).

PLAB1 – multiple influences on first attempt pass rates

In a combined (i.e. multivariate) model for PLAB1 outcomes controlling for a range of candidate characteristics (page 15 onwards).

- Only a few candidate factors show a strong predictive relationship with pass/fail decisions, and these are broadly consistent with the bivariate analysis (i.e. individual factor influences on passing).
- Older candidates are less likely to pass PLAB1 at the first attempt
- The nature of the language test is a statistically significant predictor of passing PLAB1 (both IELTS and OET have higher pass rates compared to those with alternative evidence of English language[1]).
- In terms of ethnicity, those in White group are more likely to pass PLAB1 compared to candidates classed as Asian or Asian British.
- In addition, PMQ country of origin (treated as a random effect) accounts for around 11% of (residual) variance in pass/fail outcomes.
- The 'raw' and adjusted estimates of pass rates by PMQ country of origin are very strongly related (r=0.98) suggesting that differences in pass rates by country are not driven by differences in individual candidate factor across countries that have been included in the analysis.

PLAB2– individual influences on pass rates

Candidate characteristics tend to have a stronger association with PLAB2 pass/fail outcomes compared with PLAB1. Some of the larger effects are, for example, with:

- gender (females pass rate 69.9% compared to 54.3% for males – page 23),

[1] Where candidates do not have English Language test scores, they have provided alternative evidence of their language ability. This is usually evidence of completing their medical degree in English.

- religion (69.3% pass rate for those with no religion compared to 58.7% for Muslims – page 25), and,
- PMQ country of origin (higher pass rate from South Africa, and the lowest Sri Lanka (but both of these countries have quite small candidate numbers – page 28).

PLAB2 – multiple influences on pass rates

In a combined (i.e. multivariate) model for PLAB2 outcomes controlling for a range of candidate characteristics (page 32 onwards).

- Six candidate factors/levels show a statistically significant predictive relationship with pass/fail decisions, and these are generally consistent with the bivariate analysis (i.e. individual factor influences on passing). For example, gender and age both predict PLAB2 outcomes (females and younger candidates more likely to pass).
- PMQ country of origin (treated as a random effect) accounts for around 10-13% of (residual) variance in pass/fail outcomes.
- The 'raw' and adjusted estimates of pass rates by PMQ country of origin are very strongly related (r=0.94) suggesting that differences in PLAB2 pass rates by country are, as with PLAB1, not driven by differences in the candidate factors that have been included in the analysis.
- When comparing ranking of pass rates by country between PLAB1 and PLAB2, there is some association (r=0.36 to 0.39), but these weak to moderate correlations suggest that these rankings are somewhat different across the two elements of PLAB.

Introduction

According to the GMC, '*Differential attainment (DA) is what we call the gap between attainment levels of different groups of doctors'* [2]. In the last few years, understanding the variation in assessment outcomes across different sub-groups of candidates in a range of exams has become an important focus of the GMC's work, and of that of UK medical schools more widely. There is a particular concern that any such systematic differences in performance (e.g. by, say, age, gender or ethnicity) might imply unfairness in education and training pathways, and perhaps in assessment processes.

This report consists essentially of a differential attainment analysis for PLAB. Using PLAB data from 2018 to 2021, it investigates the extent to which exam-level pass rates in both elements of PLAB vary by a range of candidate characteristics - gender, age, sexual orientation, ethnic origin, religion, disability, language test results, and Primary Medical Qualification (PMQ) country of origin.

In the remainder of this report we describe briefly the methodologies employed, then present the findings and make a few concluding remarks.

Methodology

PLAB1 candidate level data from September 2018 to August 2021 were analysed.

Candidate level PLAB2 assessment data from candidates who sat PLAB2 during the calendar years 2019 and 2020 were also analysed.

For both PLAB1 and PLAB2, only first attempts at the respective examinations were included in the analysis.

Table 1 shows the sample sizes of the PLAB data used.

Exam	No. of candidates	Description
PLAB1	25,895	This is candidate first attempts at PLAB1 generated over the period September 2018 to August 2021.
PLAB2	9,642	This is candidate first attempts at PLAB2 for candidates who had a PLAB2 exam in the period January 2019 to December 2020. The actual dates of these PLAB2 first attempts are from March 2017 to December 2020.

Table 1: Number of candidates used in each analysis

The data includes candidate pass/fail outcomes and personal characteristics (gender, age, sexual orientation, ethnic origin, religion, disability, language test type and Primary Medical Qualification (PMQ) country of origin).

In terms of methods of analysis, as simple methods as possible are used – including purely descriptive and graphical representations. The Chi-squared test of association is used to compare pass rates across groups (i.e. a set of bivariate analyses of how pass/fail decisions

[2] https://www.gmc-uk.org/education/standards-guidance-and-curricula/projects/differential-attainment/what-others-are-doing

vary across each candidate factor). The independent sample t-test is used to compare scale variables (i.e. age).

For interpreting results, the focus is on appropriate effect sizes (e.g. Phi for Chi-Squared, Cohen's d for t-tests) rather than p-values (Wasserstein and Lazar, 2016). This is particularly important given the relatively large overall sample sizes in the data (Table 1). For Phi, the usual guidelines are that 0.1, 0.3 and 0.5 correspond to small, medium and large effects. For Cohen's d, the corresponding values are 0.2, 0.5 and 0.8 – although these are best considered as 'rules of thumb' and the context needs to be taken into account.[3]

To investigate multiple influences on pass/fail decisions in a combined (multivariate) analysis, mixed modelling is carried out using the R package lme4 (Bates et al., 2015). In this multivariate modelling with pass/fail as the outcome, we treat most factors as fixed effects, but PMQ country of origin as a random (intercept) effect. This means we get separate estimates of the effect of each country of origin on pass rates – both 'raw' (i.e. without other candidate factors being taken into account) and adjusted (where these are all included and adjusted for). Pearson correlation is used to compare different model estimates for PMQ country of origin.

Findings

PLAB1 – individual influences on first attempt pass rates

The following sub-sections detail bivariate analysis of PLAB1 pass/fail decisions against personal candidate characteristics. The overall PLAB1 pass rate in this data is 72.0%.

Gender

The pass rates at first attempt by gender are very similar for PLAB1 over this period (Table 2) – 71.6% for males, and 72.4% for females (Chi-square=1.79, df=1, p=0.18, phi=0.01).

			Pass Flag (P/F		Total
			F	P	
Gender	Man	Count	3,747	9,458	**13,205**
		% within Gender	28.4%	71.6%	**100%**
	Woman	Count	3,506	9,184	**12,690**
		% within Gender	27.6%	72.4%	**100%**
Total		Count	**7,253**	**18642**	**25,895**
		% within Gender	**28.0%**	**72.0%**	**100%**

Table 2: Cross-tab of PLAB1 pass/fail versus gender

Age

There is some evidence that those passing were approximately a year younger on average compared to those failing (Table 1: t=12.80, df=25,893, p<0.001, Cohen's d=0.20).

[3] See here for a guide to interpretation

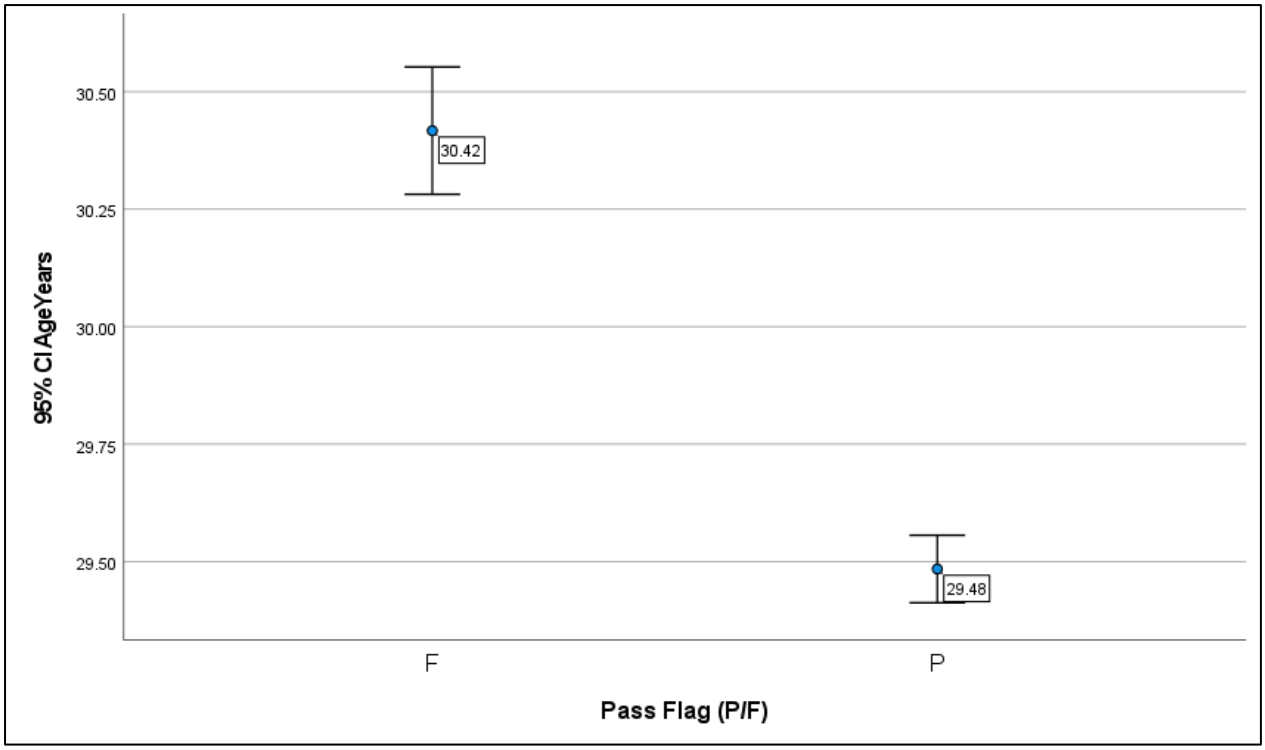

Figure 1: Error bar of PLAB1 pass/fail by mean age

Sexual orientation

The pass rates for Lesbian/Gay/Bisexual candidates are higher than for the Heterosexual/straight group – 77.2% and 71.9% respectively, but the number in the former group is relatively small (Table 3: Chi-square=4.00, df=2, p=0.14, phi=0.01).

			Pass Flag (P/F)		Total
			F	P	
Sexual orientation	Heterosexual/Straight	Count	6,809	17,430	**24,239**
		% within Sexual Orientation	28.1%	71.9%	**100%**
	Lesbian/Gay/Bisexual	Count	66	223	**289**
		% within Sexual Orientation	22.8%	77.2%	**100%**
	Not recorded/Other/Not said	Count	378	989	**1,367**
		% within Sexual Orientation	27.7%	72.3%	**100%**
Total		Count	**7,253**	**18,642**	**25,895**
		% within Sexual Orientation	**28.0%**	**72.0%**	**100%**

Table 3: Cross-tab of PLAB1 pass/fail versus sexual orientation

Ethnic origin

Overall, the relationship between ethnic origin and pass rates is quite small despite reaching statistical significance (Table 4: Chi-Square=76.2, df=5, p<0.001, phi=0.05). There is, however, some variation across particular groups – for example, the pass rate is 69.9% for Black or Black British candidates, but for Whites this is 78.0%.

			Pass Flag (P/F)		Total
			F	P	
Ethnic origin	Asian or Asian British	Count	3,618	8,736	**12,354**
		% within Ethnic Origin	29.3%	70.7%	**100%**
	Black or Black British	Count	1,698	3,943	**5,641**
		% within Ethnic Origin	30.1%	69.9%	**100%**
	Mixed	Count	134	349	**483**
		% within Ethnic Origin	27.7%	72.3%	**100%**
	Not stated	Count	483	1,444	**1,927**
		% within Ethnic Origin	25.1%	74.9%	**100%**
	Other Ethnic Groups	Count	1,083	3,328	**4,411**
		% within Ethnic Origin	24.6%	75.4%	**100%**
	White	Count	237	842	**1,079**
		% within Ethnic Origin	22.0%	78.0%	**100%**
Total		Count	**7,253**	**18,642**	**25,895**
		% within Ethnic Origin	**28.0%**	**72.0%**	**100%**

Table 4: Cross-tab of PLAB1 pass/fail versus ethnic group

Religion

There are some moderate differences in pass rates by religion – for example, 69.8% for Christians compared to 73.7% for Muslims (Table 5: Chi-Square=108.3, df=7, p<0.001, phi=0.07).

			Pass Flag (P/F)		Total
			F	P	
Religion	Buddhist	Count	173	569	**742**
		% within Religion	23.3%	76.7%	**100%**
	Christian	Count	2,062	4,760	**6,822**
		% within Religion	30.2%	69.8%	**100%**
	Hindu	Count	1,155	2,364	**3,519**
		% within Religion	32.8%	67.2%	**100%**
	Muslim	Count	3,359	9,404	**12,763**
		% within Religion	26.3%	73.7%	**100%**
	No religion	Count	191	695	**886**
		% within Religion	21.6%	78.4%	**100%**
	Other/Jewish/Missing	Count	73	152	**225**
		% within Religion	32.4%	67.6%	**100%**
	Prefer not to say	Count	181	556	**737**
		% within Religion	24.6%	75.4%	**100%**
	Sikh	Count	59	142	**201**
		% within Religion	29.4%	70.6%	**100%**
Total		Count	**7,253**	**18,642**	**25,895**
		% within Religion	**28.0%**	**72.0%**	**100%**

Table 5: Cross-tab of PLAB1 pass/fail versus religion

Disability

The pass rate is slightly lower for those candidates declaring a disability - 69.3% versus 72.1% for non-disabled, but the numbers in the disabled group are relatively small (Table 6: Chi-Square=8.89, df=2, p=0.01, phi=0.02).

			Pass Flag (P/F)		Total
			F	P	
Disability	No	Count	7,125	18,377	**25,502**
		% within Disability	27.9%	72.1%	**100%**
	Not recorded	Count	22	26	**48**
		% within Disability	45.8%	54.2%	**100%**
	Yes	Count	106	239	**345**
		% within Disability	30.7%	69.3%	**100%**
Total		**Count**	**7,253**	**18,642**	**25,895**
		% within Disability	**28.0%**	**72.0%**	**100%**

Table 6: Cross-tab of PLAB1 pass/fail versus disability

Language test type[4]

The pass rate is similar across language test type groups - 72.0% for IELTS versus 72.2% for OET, with the Alternative group a little lower (68.2%) (Table 7: Chi-Square=6.86, df=2, p=0.03, phi=0.02).

			Pass Flag (P/F)		Total
			F	P	
Language test type	Alternative[5]	Count	291	625	**916**
		% within ELT Type	31.8%	68.2%	**100%**
	IELTS	Count	3,213	8,257	**11,470**
		% within ELT Type	28.0%	72.0%	**100%**
	OET	Count	3,749	9,760	**13,509**
		% within ELT Type	27.8%	72.2%	**100%**
Total		**Count**	**7,253**	**18,642**	**25,895**
		% within ELT Type	**28.0%**	**72.0%**	**100%**

Table 7: Cross-tab of PLAB1 pass/fail versus language test type

[4] There was over 50% missing on the actual language test scores themselves, so no further analysis of these has been included in this report.

[5] These candidates will have provided alternative evidence of their language ability, usually that they completed their medical degree in English.

PMQ origin – country[6]

There are some quite large differences across countries for pass rates (Table 8: Chi-Square=449.1, df=27. p<0.001, phi=0.13).

[6] Only countries with at least 100 candidates in the original PLAB1 dataset are included separately in this cross tab analysis.

			Pass Flag (P/F)		Total
			F	P	
Country	Bahrain	Count	27	124	**151**
		% within PMQ Country	17.9%	82.1%	**100%**
	Bangladesh	Count	343	919	**1,262**
		% within PMQ Country	27.2%	72.8%	**100%**
	China	Count	350	491	**841**
		% within PMQ Country	41.6%	58.4%	**100%**
	Egypt	Count	811	2,609	**3,420**
		% within PMQ Country	23.7%	76.3%	**100%**
	Ghana	Count	56	164	**220**
		% within PMQ Country	25.5%	74.5%	**100%**
	Grenada	Count	15	89	**104**
		% within PMQ Country	14.4%	85.6%	**100%**
	India	Count	1,208	2,538	**3,746**
		% within PMQ Country	32.2%	67.8%	**100%**
	Iran, Islamic Republic Of	Count	15	90	**105**
		% within PMQ Country	14.3%	85.7%	**100%**
	Iraq	Count	71	327	**398**
		% within PMQ Country	17.8%	82.2%	**100%**
	Ireland	Count	29	100	**129**
		% within PMQ Country	22.5%	77.5%	**100%**
	Jordan	Count	59	363	**422**
		% within PMQ Country	14.0%	86.0%	**100%**
	Libyan Arab Jamahiriya	Count	98	180	**278**
		% within PMQ Country	35.3%	64.7%	**100%**
		Count	80	351	**431**

	Myanmar	**% within PMQ Country**	18.6%	81.4%	**100%**
	Nepal	**Count**	153	363	**516**
		% within PMQ Country	29.7%	70.3%	**100%**
	Nigeria	**Count**	1,199	2,811	**4,010**
		% within PMQ Country	29.9%	70.1%	**100%**
	Other	**Count**	573	1,286	**1,859**
		% within PMQ Country	30.8%	69.2%	**100%**
	Pakistan	**Count**	1,156	3,405	**4,561**
		% within PMQ Country	25.3%	74.7%	**100%**
	Philippines	**Count**	110	225	**335**
		% within PMQ Country	32.8%	67.2%	**100%**
	Russian Federation	**Count**	105	136	**241**
		% within PMQ Country	43.6%	56.4%	**100%**
	Saint Kitts And Nevis	**Count**	30	66	**96**
		% within PMQ Country	31.3%	68.8%	**100%**
	Saudi Arabia	**Count**	45	105	**150**
		% within PMQ Country	30.0%	70.0%	**100%**
	South Africa	**Count**	31	137	**168**
		% within PMQ Country	18.5%	81.5%	**100%**
	Sri Lanka	**Count**	13	112	**125**
		% within PMQ Country	10.4%	89.6%	**100%**
	Sudan	**Count**	283	661	**944**
		% within PMQ Country	30.0%	70.0%	**100%**
	Turkey	**Count**	45	220	**265**
		% within PMQ Country	17.0%	83.0%	**100%**
		Count	239	331	**570**

	Ukraine	**% within PMQ Country**	41.9%	58.1%	**100%**
	United Arab Emirates	**Count**	82	262	**344**
		% within PMQ Country	23.8%	76.2%	**100%**
	Zimbabwe	**Count**	27	177	**204**
		% within PMQ Country	13.2%	86.8%	**100%**
Total		**Count**	7,253	18,642	25,895
		% within PMQ Country	28.0%	72.0%	100%

Table 8: Cross-tab of PLAB1 pass/fail versus PMQ country of origin

Figure 2 shows pass rates in order highest to lowest – with Russia having the lowest rate (56.4%) and Sri Lanka the highest (89.6%). Note, some countries, including both of these, have relatively low numbers of candidates – see Table 8 for exact numbers.

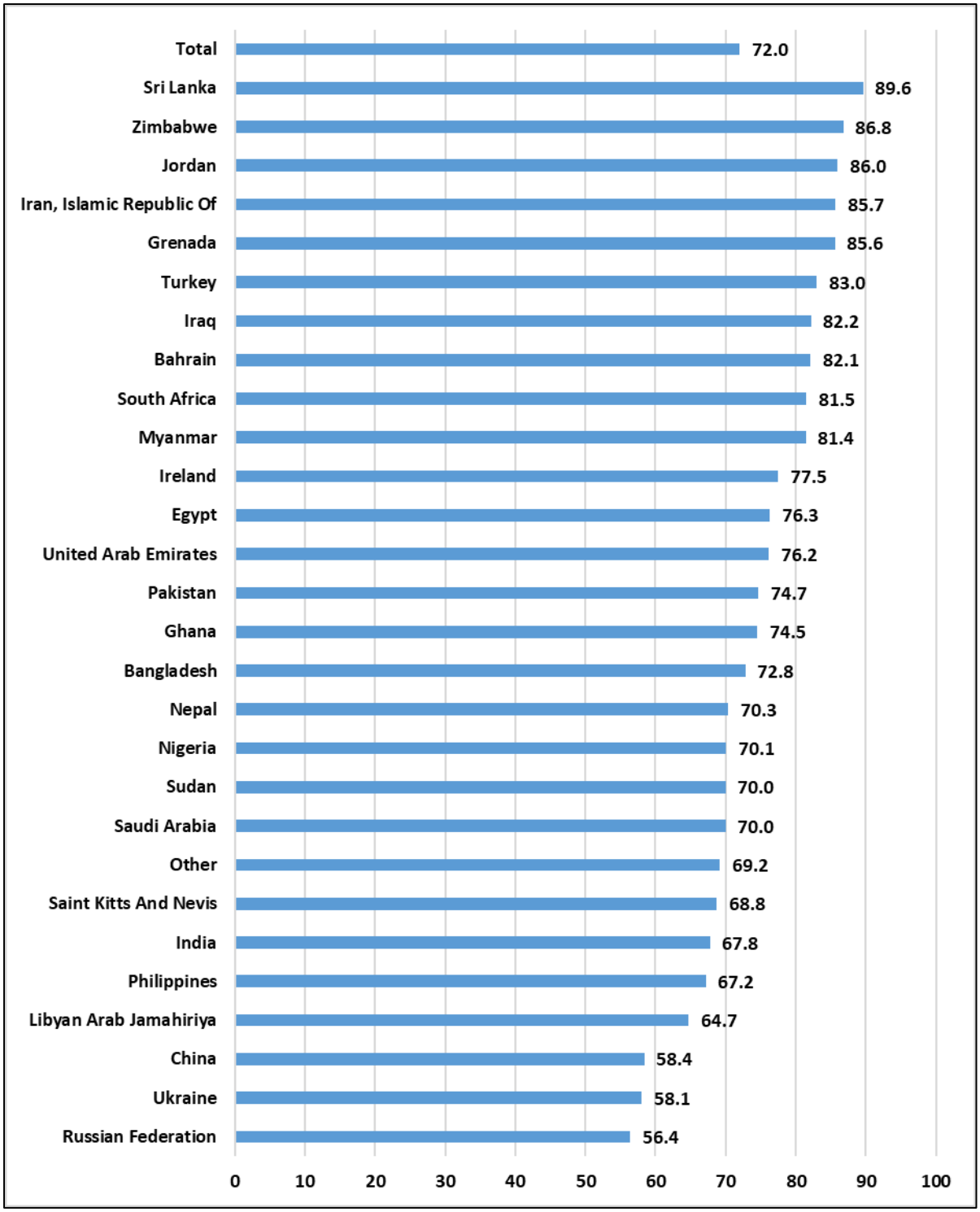

Figure 2: PLAB1 pass rates by PMQ country of origin (highest to lowest)

PLAB1 – multiple influences on first attempt pass rates

In a combined (multivariate) model, we can investigate independent influences on PLAB1 outcomes (pass/fail) using all the factors analysed separately in in the previous section. This

will give us estimates of the separate effects of each factor on the PLAB1 outcome having controlled for the others.[7] We begin with a simple, null, model for comparison.

The null model – including only PMQ country of origin

We treat PMQ country of origin as a random effect. In a (null) model just including that factor and no others, PMQ country of origin accounts for 11.0% of variance in pass/fail outcomes. This tells us that there is some variation in pass rates across countries (as we have already seen from page 11 onwards), but that country of origin alone does not pre-determine to any great extent passing or failing PLAB1.

The full (multivariate) model with predictors

On adding all the explanatory factors (in a multivariate model), this proportion of (residual) variance changes very little to be approximately 11.6% (Table 9). This tells us that having accounted for a range of factors (e.g. gender, age, sexual orientation etc), PMQ country of origin explains around a tenth of the remaining (residual) variation in PLAB1 pass/fail outcomes – but as previously, this is not a dramatically large effect.

[7] Disability was removed from the modelling as it was confounded with other factors. In addition, there were convergence problems, probably due to the range and number levels of the factors involved in combination. However, the results presented here appear reliable and consistent with earlier analyses.

Random effect (full model)	Variance	Std.Dev.	Conditional Intra-class correlation coefficient
PMQ Origin (intercept)	0.38	0.62	11.6%
Total	3.29		
Number of observations=25,895, PMQ country of origin=134			

Table 9: Random effect of PMQ country of origin on passing PLAB1

Predictor estimates in the full model

When it comes to fixed effects (i.e. categorical and scale factors influencing the chances of passing or failing – e.g. gender, age, ethnicity and so on), Table 10 shows the results of the full model for those statistically significant factors influencing PLAB1 outcomes. These are approximately in order of size of effect (strongest to weakest).

Fixed effects	coefficient=logs odds of pass versus fail	Odds of pass versus fail	std error	z value	p-value	Interpretation
(Intercept)	1.41	4.10	0.19	7.39	<0.01	Not applicable – this is the baseline value.
Age At PLAB1 First Attempt	-0.03	0.97	0.003	-11.79	<0.01	As age increases, the odds of passing decreases
English Language Test Type=IELTS – relative to alternative group	0.58	1.78	0.10	6.08	<0.01	Those taking an IELTS language test are more likely to pass versus those with alternative evidence of language ability.
English Language Test Type=OET – (relative to alternative group)	0.47	1.59	0.09	4.93	<0.01	Those taking an OET language test are more likely to pass versus those with alternative evidence of language ability.
Ethnic origin White (relative to Asian or Asian British group)	0.28	1.33	0.10	2.84	<0.01	Those in White group more likely to pass compared to Asian or Asian British group

Table 10: Fixed effects (significant at p=0.05 level) for passing PLAB1

We can see in Table 10 that only a few predictors impact sufficiently strongly on pass/fail decisions to be flagged as statistically significant in this model. These are

- *Age* - which on increasing has a negative impact (as we also saw in Figure 1).
- *English language test type* – where IELTS and OET candidates are more likely to pass PLAB1 compared to those who have provided alternative evidence (see also Table 7 for the bivariate analysis using this factor which has a similar finding).
- *Ethnic group* – White candidates more likely to pass compared to those from the Asian or Asian British group (again this result is broadly consistent with the earlier analysis in Table 4).

PMQ country of origin effects in the full model

For completeness, Table 11 shows the individual PMQ country of origin full model estimates of passing/failing PLAB1 (ordered highest to lowest odds).[8]

We see that Sri Lanka and the Palestinian Territories have the highest odds i.e. candidates with PMQs from these countries are the most likely to pass PLAB1 having accounted for other factors. The countries with the lowest pass rates are Georgia and Romania.

It should be noted that the sample sizes in some of the countries is small (median=15, min=1, max= 4,561). The countries with substantial numbers of candidates (e.g. over 100) are shown above in Table 8.

To protect candidate confidentiality, the names of countries with fewer than five PLAB1 individuals have been supressed in Table 11.

[8] These values are somewhat artificial as they use all reference group categories, so should be thought of as illustrative only.

PMQ country of origin	Sample size	Log odds	Odds	Model-based pass proportion
Sri Lanka	125	2.60	13.51	0.93
Palestinian Territories	64	2.56	12.89	0.93
Grenada	104	2.38	10.81	0.92
Iran, Islamic Republic Of	105	2.25	9.47	0.90
Zimbabwe	204	2.23	9.27	0.90
Australia	89	2.16	8.70	0.90
Tunisia	12	2.09	8.06	0.89
Jordan	422	2.07	7.91	0.89
Canada	10	2.06	7.82	0.89
Brazil	84	2.05	7.75	0.89
Iraq	398	1.99	7.30	0.88
Korea, Republic Of	12	1.96	7.12	0.88
Bahrain	151	1.94	6.93	0.87
Cuba	38	1.92	6.82	0.87
Jamaica	66	1.90	6.69	0.87
South Africa	168	1.89	6.63	0.87
Thailand	6	1.89	6.60	0.87
Colombia	13	1.88	6.53	0.87
Myanmar	431	1.86	6.43	0.87
New Zealand	22	1.85	6.39	0.86
Oman	35	1.85	6.38	0.86
Lebanon	33	1.85	6.34	0.86
Turkey	265	1.81	6.10	0.86
Japan	14	1.80	6.05	0.86
	4	1.79	5.99	0.86
Ireland	129	1.78	5.93	0.86
Afghanistan	22	1.76	5.84	0.85
Kazakhstan	15	1.74	5.68	0.85
	2	1.72	5.60	0.85
St. Vincent And The Grenadines	7	1.72	5.56	0.85
Somalia	8	1.71	5.54	0.85
Morocco	18	1.70	5.46	0.85
United States	18	1.68	5.38	0.84
	3	1.64	5.18	0.84
Chile	11	1.64	5.16	0.84
Syrian Arab Republic	91	1.63	5.12	0.84
	3	1.63	5.10	0.84
	3	1.62	5.06	0.84
	2	1.62	5.06	0.83
Barbados	27	1.60	4.97	0.83
Egypt	3,420	1.60	4.94	0.83
	2	1.59	4.90	0.83
	1	1.59	4.89	0.83

Kenya	37	1.58	4.85	0.83
	2	1.57	4.82	0.83
	5	1.57	4.80	0.83
	2	1.57	4.78	0.83
	2	1.56	4.78	0.83
	1	1.56	4.77	0.83
Trinidad And Tobago	60	1.56	4.74	0.83
Tanzania, United Republic Of	15	1.55	4.73	0.83
	1	1.55	4.73	0.83
Pakistan	4,561	1.55	4.71	0.82
	1	1.54	4.67	0.82
	1	1.53	4.61	0.82
	1	1.51	4.53	0.82
Indonesia	20	1.51	4.52	0.82
Bangladesh	1,262	1.51	4.52	0.82
	1	1.51	4.51	0.82
	1	1.50	4.50	0.82
	1	1.50	4.48	0.82
	1	1.49	4.42	0.82
	5	1.48	4.40	0.81
	4	1.46	4.30	0.81
United Arab Emirates	344	1.44	4.23	0.81
Ghana	220	1.42	4.13	0.81
Belize	6	1.40	4.07	0.80
	3	1.40	4.04	0.80
Saint Kitts And Nevis	96	1.39	4.01	0.80
Zambia	14	1.37	3.95	0.80
Algeria	24	1.36	3.91	0.80
Nepal	516	1.34	3.80	0.79
	2	1.33	3.79	0.79
Taiwan	7	1.33	3.79	0.79
Ethiopia	21	1.31	3.71	0.79
Sudan	944	1.31	3.70	0.79
Albania	19	1.31	3.70	0.79
Venezuela	6	1.30	3.65	0.79
	2	1.29	3.65	0.78
	3	1.29	3.63	0.78
Nigeria	4,010	1.28	3.61	0.78
Dominica	25	1.28	3.60	0.78
	2	1.28	3.60	0.78
Guyana	39	1.28	3.59	0.78
Antigua And Barbuda	21	1.27	3.57	0.78
	2	1.27	3.56	0.78
Argentina	10	1.27	3.55	0.78

Malawi	15	1.23	3.41	0.77
Yemen	38	1.21	3.36	0.77
Philippines	335	1.20	3.32	0.77
	1	1.19	3.30	0.77
Libyan Arab Jamahiriya	278	1.19	3.28	0.77
	1	1.18	3.26	0.77
Ecuador	8	1.18	3.25	0.76
	1	1.18	3.25	0.76
Saudi Arabia	150	1.16	3.20	0.76
	1	1.16	3.19	0.76
	4	1.15	3.17	0.76
Peru	11	1.15	3.17	0.76
	4	1.14	3.14	0.76
India	3,746	1.14	3.14	0.76
	1	1.14	3.13	0.76
Saint Lucia	20	1.13	3.10	0.76
	1	1.12	3.07	0.75
Dominican Republic	6	1.11	3.05	0.75
	3	1.09	2.99	0.75
	3	1.09	2.99	0.75
Honduras	13	1.09	2.98	0.75
Hungary	36	1.08	2.96	0.75
	3	1.04	2.84	0.74
Gambia	11	1.02	2.78	0.74
Mauritius	48	1.02	2.78	0.74
Mexico	27	1.02	2.77	0.73
	3	1.00	2.73	0.73
Malaysia	89	0.99	2.70	0.73
Uganda	38	0.96	2.60	0.72
Kyrgyzstan	31	0.95	2.60	0.72
Serbia	16	0.92	2.51	0.72
Moldova, Republic Of	9	0.84	2.31	0.70
	4	0.83	2.30	0.70
China	841	0.78	2.17	0.68
Russian Federation	241	0.75	2.12	0.68
Czech Republic	28	0.71	2.03	0.67
Ukraine	570	0.70	2.01	0.67
	5	0.68	1.98	0.66
Cameroon	31	0.67	1.96	0.66
Azerbaijan	12	0.63	1.88	0.65
Poland	46	0.56	1.76	0.64
Armenia	22	0.56	1.75	0.64
Bulgaria	17	0.46	1.58	0.61
Belarus	54	0.44	1.55	0.61

Italy	12	0.44	1.55	0.61
Georgia	46	0.43	1.54	0.61
Romania	38	0.34	1.41	0.59

Table 11: Individual PMQ country of origin estimates of odds of passing PLAB1

Comparisons of PMQ country of origin estimates between null and full models

If we compare the PMQ country estimates (i.e. the log odds column shown in Table 11) from the full model to those from the null model (i.e. with no predictors), we find a very strong correlation ($r=0.98$, $n=134$, $p<0.001$). This is strong evidence that adjusting for the candidate characteristics we have included (gender, age, sexual orientation etc) does not change the ranking of PLAB1 pass rates by countries very much.

PLAB2– individual influences on pass rates

In what follows, each sub-section details bivariate analysis of PLAB2 pass/fail decisions against personal candidate characteristics. The overall pass rate in this data is 61.8%[9].

Gender

The pass rates are much higher for females (69.9%) than for males (54.3%) for PLAB2 over this period (Table 12) (Chi-square=246.0, df=1, p<0.001, phi=0.16).

This is a much bigger gender difference than seen in PLAB1 (Table 2).

			Pass Flag (P/F		Total
			F	P	
Gender	Man	Count	2,289	2,725	**5,014**
		% within Gender	45.7%	54.3%	**100%**
	Woman	Count	1,394	3,234	**4,628**
		% within Gender	30.1%	69.9%	**100%**
Total		Count	**3,683**	**5,959**	**9,642**
		% within Gender	**38.2%**	**61.8%**	**100%**

Table 12: Cross-tab of PLAB2 pass/fail versus gender

Age

There is some evidence that those passing were approximately a year younger on average compared to those failing (Table 1: t=8.17, df=9640, p<0.001, Cohen's d=0.17).

This is a very similar age effect as seen in PLAB1 (Figure 1).

[9] These overall pass rates for PLAB2 are slightly lower than those published because the latter includes all candidate attempts not just first attempts as here.

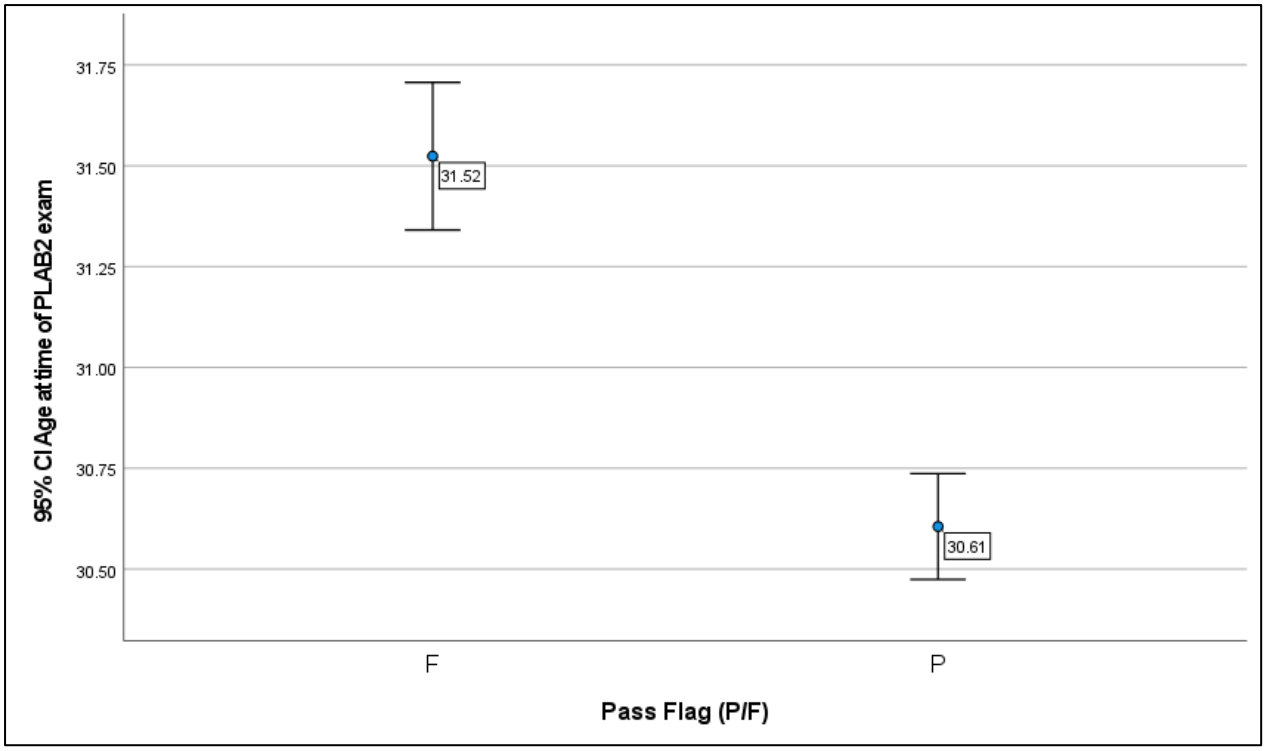

Figure 3: Error bar of PLAB2 pass/fail by mean age

Sexual orientation

The pass rates are lower for those that did not have their sexuality recorded – 51.9% for this group compared to 62.3% for heterosexuals and 67.0% for Lesbian/Gay/Bisexual (Table 13Table 3: Chi-square=24.3, df=2, p<0.001, phi=0.05).

The corresponding analysis for PLAB1 showed a slightly smaller difference across these three groups (Table 3).

			Pass Flag (P/F)		Total
			F	P	
Sexual orientation	Heterosexual/Straight	Count	3,392	5,611	**9,003**
		% within Sexual Orientation	37.7%	62.3%	**100%**
	Lesbian/Gay/Bisexual	Count	36	73	**109**
		% within Sexual Orientation	33.0%	67.0%	**100%**
	Not recorded/Other/Not said	Count	255	275	**530**
		% within Sexual Orientation	48.1%	51.9%	**100%**
Total		Count	**3,683**	**5,959**	**9,642**
		% within Sexual Orientation	**38.2%**	**61.8%**	**100%**

Table 13: Cross-tab of PLAB2 pass/fail versus sexual orientation

Ethnic origin

There is some variation in pass rates by ethnic origin (Table 14: Chi-Square=77.8, df=5, p<0.001, phi=0.09). The lowest pass rate is 57.7% for the Asian or Asian British group, compared to 72.7% for the (smaller) Mixed group.

As already seen for other characteristics, these differences are typically a little larger than they were in the corresponding PLAB1 analysis (Table 4).

			Pass Flag (P/F)		Total
			F	P	
Ethnic origin	Asian or Asian British	Count	1,779	2,425	**4,204**
		% within Ethnic Origin	42.3%	57.7%	**100%**
	Black or Black British	Count	819	1,600	**2,419**
		% within Ethnic Origin	33.9%	66.1%	**100%**
	Mixed	Count	56	149	**205**
		% within Ethnic Origin	27.3%	72.7%	**100%**
	Not stated	Count	259	369	**628**
		% within Ethnic Origin	41.2%	58.8%	**100%**
	Other Ethnic Groups	Count	629	1,084	**1,713**
		% within Ethnic Origin	36.7%	63.3%	**100%**
	White	Count	141	332	**473**
		% within Ethnic Origin	29.8%	70.2%	**100%**
Total		Count	**3,683**	**5,959**	**9,642**
		% within Ethnic Origin	**38.2%**	**61.8%**	**100%**

Table 14: Cross-tab of PLAB2 pass/fail versus ethnic group

Religion

There are some differences in pass rates by religion – for example, 69.3% for those with no religion compared to 58.7% for Muslims (Table 15: Chi-Square=96.8, df=8, p<0.001, phi=0.10).

These differences are perhaps slightly larger than those observed in the corresponding PLAB1 analysis (Table 5).

			Pass Flag (P/F)		Total
			F	P	
Religion	Buddhist	Count	159	174	**333**
		% within Religion	47.7%	52.3%	**100%**
	Christian	Count	918	1,920	**2,838**
		% within Religion	32.3%	67.7%	**100%**
	Hindu	Count	434	698	**1,132**
		% within Religion	38.3%	61.7%	**100%**
	Muslim	Count	1,852	2,637	**4,489**
		% within Religion	41.3%	58.7%	**100%**
	No religion	Count	110	248	**358**
		% within Religion	30.7%	69.3%	**100%**
	Not recorded	Count	42	26	**68**
		% within Religion	61.8%	38.2%	**100%**
	Other/Jewish	Count	24	40	**64**
		% within Religion	37.5%	62.5%	**100%**
	Prefer not to say	Count	123	184	**307**
		% within Religion	40.1%	59.9%	**100%**
	Sikh	Count	21	32	**53**
		% within Religion	39.6%	60.4%	**100%**
Total		Count	**3,683**	**5,959**	**9,642**
		% within Religion	**38.20%**	**61.8%**	**100%**

Table 15: Cross-tab of PLAB2 pass/fail versus religion

Disability

The pass rates are very similar between those with and without a disability (63.0% vs 62.0%), but the small unrecorded group have a lower pass rate (38.2%). (Table 16: Chi-Square=16.1, df=2, p<0.001, phi=0.04).

The difference in PLAB2 pass rates between disabled and non-disabled is smaller than that seen in PLAB1 (Table 6).

			Pass Flag (P/F)		Total
			F	P	
Disability	No	Count	3,594	5,853	9,447
		% within Disability	38.0%	62.0%	100%
	Not recorded	Count	42	26	68
		% within Disability	61.8%	38.2%	100%
	Yes	Count	47	80	127
		% within Disability	37.0%	63.0%	100%
Total		Count	3,683	5,959	9,642
		% within Disability	38.2%	61.8%	100.0%

Table 16: Cross-tab of PLAB2 pass/fail versus disability

Language test type

The pass rate is lower for OET than IELTS (58.8% vs 64.1% respectively) (Table 17: Chi-Square=30.0, df=2, p<0.001, phi=0.06).

There was a smaller difference in pass rates in PLAB1 for these groups (Table 7).

			Pass Flag (P/F)		Total
			F	P	
Language test type	Alternative[10]	Count	178	236	**414**
		% within ELT Type	43.0%	57.0%	**100%**
	IELTS	Count	2,015	3,595	**5,610**
		% within ELT Type	35.9%	64.1%	**100%**
	OET	Count	1,490	2,128	**3,618**
		% within ELT Type	41.2%	58.8%	**100%**
Total		Count	**3,683**	**5,959**	**9,642**
		% within ELT Type	**38.2%**	**61.8%**	**100%**

Table 17: Cross-tab of PLAB2 pass/fail versus language test type

[10] These candidates will have provided alternative evidence of their language ability, usually that they completed their medical degree in English.

PMQ origin – country[11]

There are quite large differences across countries for pass rates (Table 18: Chi-Square=226, df=25, p<0.001, phi=0.15). These differences tend to be slightly larger than those seen for the corresponding analysis in PLAB1 (Table 8).

[11] Only countries with at least 50 candidates are included separately in this cross tab analysis.

			Pass Flag (P/F)		Total
			F	P	
Country	Bahrain	Count	14	59	73
		% within PMQ Country	19.2%	80.8%	100%
	Bangladesh	Count	173	184	357
		% within PMQ Country	48.5%	51.5%	100%
	China	Count	128	120	248
		% within PMQ Country	51.6%	48.4%	100%
	Egypt	Count	515	784	1,299
		% within PMQ Country	39.6%	60.4%	100%
	Ghana	Count	16	60	76
		% within PMQ Country	21.1%	78.9%	100%
	Grenada	Count	18	71	89
		% within PMQ Country	20.2%	79.8%	100%
	India	Count	435	725	1,160
		% within PMQ Country	37.5%	62.5%	100%
	Iran, Islamic Republic Of	Count	18	36	54
		% within PMQ Country	33.3%	66.7%	100%
	Iraq	Count	55	112	167
		% within PMQ Country	32.9%	67.1%	100%
	Ireland	Count	12	51	63
		% within PMQ Country	19.0%	81.0%	100%
	Jordan	Count	31	72	103
		% within PMQ Country	30.1%	69.9%	100%
	Libyan Arab Jamahiriya	Count	52	63	115
		% within PMQ Country	45.2%	54.8%	100%

	Myanmar	**Count**	110	112	**222**
		% within PMQ Country	49.5%	50.5%	**100%**
	Nepal	**Count**	88	107	**195**
		% within PMQ Country	45.1%	54.9%	**100%**
	Nigeria	**Count**	648	1,152	**1,800**
		% within PMQ Country	36.0%	64.0%	**100%**
	Other	**Count**	247	597	**844**
		% within PMQ Country	29.3%	70.7%	**100%**
	Pakistan	**Count**	728	886	**1,614**
		% within PMQ Country	45.1%	54.9%	**100%**
	Philippines	**Count**	28	82	**110**
		% within PMQ Country	25.5%	74.5%	**100%**
	Russian Federation	**Count**	45	43	**88**
		% within PMQ Country	51.1%	48.9%	**100%**
	South Africa	**Count**	5	58	**63**
		% within PMQ Country	7.9%	92.1%	**100%**
	Sri Lanka	**Count**	31	28	**59**
		% within PMQ Country	52.5%	47.5%	**100%**
	Sudan	**Count**	133	241	**374**
		% within PMQ Country	35.6%	64.4%	**100%**
	Turkey	**Count**	33	53	**86**
		% within PMQ Country	38.4%	61.6%	**100%**
	Ukraine	**Count**	71	124	**195**
		% within PMQ Country	36.4%	63.6%	**100%**
	United Arab Emirates	**Count**	38	89	**127**
		% within PMQ Country	29.9%	70.1%	**100%**

	Zimbabwe	**Count**	11	50	**61**
		% within PMQ Country	18.0%	82.0%	**100%**
Total		**Count**	**3,683**	**5,959**	**9,642**
		% within PMQ Country	**38.2%**	**61.8%**	**100%**

Table 18: Cross-tab of PLAB2 pass/fail versus PMQ country of origin

Figure 4 shows PLAB2 pass rates by country in order highest to lowest – with Sri Lanka having the lowest rate (note some countries have relatively low numbers of candidates – see Table 18 for details).

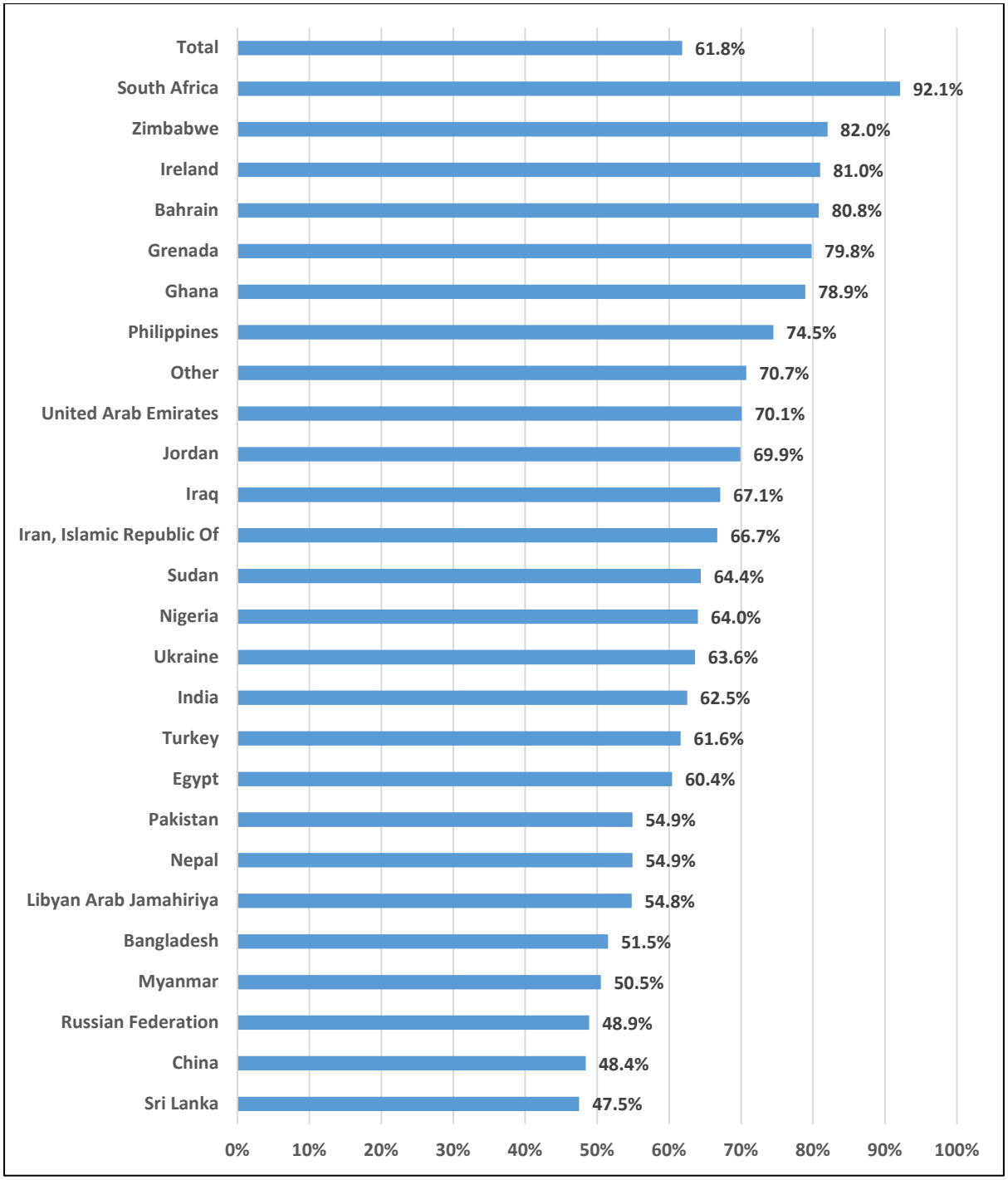

Figure 4: PLAB2 pass rates by PMQ country of origin (highest to lowest)

PLAB2 – multiple influences on pass rates

As with PLAB1, we can use a combined (multivariate) model to investigate independent influences on PLAB2 pass/fail outcomes in terms of the candidate characteristics analysed separately in in the previous sub-sections. This will give us estimates of the separate effects

of each characteristic on the PLAB2 outcome having controlled for the others.[12] Again, we begin with a simple, null, model for comparison.

The null model – including only PMQ country of origin

As for PLAB1, we treat PMQ country of origin as a random effect. In a (null) model just including that factor and no others, PMQ country of origin accounts for 10% of variance in pass/fail outcomes. This is slightly weaker effect of country of origin than seen in PLAB1 (13%, page 15), perhaps because all candidate for PLAB2 have to have passed PLAB1 so the weakest candidates (possibly more likely from particular countries) are not present in PLAB2.

In summary, there is some variation in PLAB2 pass rates across countries (as also seen in the analysis from page 28 onwards), but that PMQ country of origin alone does not pre-determine to any great extent the passing or failing of PLAB2.

The full (multivariate) model with predictors

On adding all the explanatory factors in a full model, the percentage of (residual) variation due to country is approximately 12% (Table 19). This tells us that having accounted for a range of factors (e.g. gender, age, sexual orientation etc), PMQ country of origin is slightly more important in explaining the remaining (residual) variation in PLAB2 pass/fail outcomes – but again, this is not a large effect.

[12] As with the corresponding PLAB1 analysis, Disability was removed from the modelling as it was confounded with other factors. In addition, there were convergence problems, probably due to the range and number levels of the factors involved in combination. However, the results presented here appear reliable and consistent with other analyses.

Random effect (full model)	Variance	Std.Dev.	Conditional Intra-class correlation coefficient
PMQ Origin (intercept)	0.40	0.64	12.3%
Total	3.29		
Number of observations=9,642, PMQ country of origin=116			

Table 19: Random effect of PMQ country of origin on passing PLAB2

Predictor estimates in the full model

The categorical and scale factors (fixed effects) influencing the chances of passing or failing PLAB2 in this multivariate model are shown in Table 20 – ordered approximately strongest to weakest effect. Those that are not statistically significant at the 0.05 level are not shown.

Fixed effects	coefficient=logs odds of pass versus fail	Odds of pass versus fail	std error	z value	p-value	Interpretation
Gender (female)	0.72	2.06	0.05	15.91	<0.01	Females are more likely to pass than males. This difference in performance in OSCEs has seen before e.g. (Haq et al., 2005).
Age	-0.04	0.96	0.00	-8.57	<0.01	As age increases, the odds of passing decreases. This could be related to the generalist nature of PLAB reflecting foundation practice and therefore making the assessment more challenging to those who have specialised and tend to be older.
English Language Test Type=IELTS [13]	1.11	3.03	0.16	7.03	<0.01	Those taking an IELTS language test are more likely to pass versus those with alternative evidence of language ability.
English Language Test Type=OET[Error! Bookmark not defined.]	1.00	2.72	0.16	6.30	<0.01	Those taking an OET language test are more likely to pass versus those with alternative evidence of language ability.
Sexual Orientation (not recorded) [14]	-0.29	0.75	0.11	-2.74	0.01	Those not declaring or with sexual orientation not recorded less likely to pass
Ethnic Origin (mixed) [15]	0.46	1.58	0.19	2.42	0.02	Those with ethnic origin 'Mixed' more likely to pass compared to the reference group (Asian or Asian British). No other ethnic group show a significant difference in odds of passing compare to this latter group.

Table 20: Fixed effects (significant at p=0.05 level) for passing PLAB2

Table 20 shows that six predictors (or, more precisely, levels of predictors) impact sufficiently strongly on pass/fail decisions to be flagged as statistically significant in this model.

If we compare with the bivariate analyses of PLAB2 outcomes, we see the results are quite consistent. For example, in both types of analysis females are much more likely to pass than males (compare Table 12 and Table 20). Similarly, the influences of sexual orientation (Table 13), ethnic origin (Table 14) and language test type (Table 17) in the bivariate analysis are also entirely consistent with the more complex (multivariate) analysis (Table 20). In other words, many of the effects seen in the simple cross-tab analyses remain once other factors are accounted for.

PMQ country of origin effects in the full model

For completeness, Table 21 shows the individual PMQ country of origin full model estimates of passing/failing PLAB2 (ordered highest to lowest odds). [16]

We see that Australia and Jamaica have the highest odds i.e. candidates with PMQs from these countries are the most likely to pass PLAB2 having accounted for other factors. The countries with the lowest pass rates but with substantial numbers of candidates are China and Russia.

It should be noted that the sample sizes in some of the countries is small (median=9, min=1, max=1,800). The countries with substantial numbers of candidates (e.g. over 50) are shown above in Table 18, and the results in the multivariate analysis are quite consistent with the bivariate analysis of pass rates by country. Again, broadly speaking, adjusting for other factors does not change the approximate ordering of countries by pass rates.

To protect candidate confidentiality, the names of countries with fewer than five PLAB2 individuals have been supressed in Table 21.

[13] Reference group English Language Test Type – Alternative
[14] Reference group heterosexual
[15] Reference group Asian or Asian British
[16] These should be treated as illustrative as they use reference categories.

PMQ country of origin	Sample size	Log odds	Odds	Model-based pass proportion
Australia	37	1.66	5.27	0.84
Jamaica	35	1.54	4.65	0.82
New Zealand	15	1.42	4.16	0.81
South Africa	63	1.41	4.09	0.80
Trinidad And Tobago	30	1.39	4.01	0.80
Grenada	89	1.29	3.62	0.78
Ireland	63	1.28	3.59	0.78
Bahrain	73	1.11	3.02	0.75
United States	7	1.05	2.85	0.74
Antigua And Barbados	12	0.97	2.63	0.72
Kenya	19	0.87	2.38	0.70
	5	0.85	2.35	0.70
Brazil	49	0.82	2.27	0.69
Zimbabwe	61	0.82	2.26	0.69
Barbados	14	0.79	2.20	0.69
	3	0.78	2.17	0.68
Saudi Arabia	45	0.75	2.12	0.68
	4	0.72	2.06	0.67
Cameroon	14	0.71	2.04	0.67
	3	0.70	2.01	0.67
Lebanon	8	0.68	1.98	0.66
Ghana	76	0.61	1.85	0.65
Bulgaria	7	0.60	1.82	0.65
	2	0.59	1.81	0.64
	1	0.59	1.80	0.64
	1	0.57	1.77	0.64
Saint Lucia	9	0.52	1.69	0.63
	2	0.52	1.68	0.63
Malaysia	33	0.51	1.66	0.62
	1	0.51	1.66	0.62
	1	0.51	1.66	0.62
	5	0.50	1.65	0.62
	1	0.50	1.65	0.62
	1	0.49	1.64	0.62
Iraq	167	0.49	1.64	0.62
Philippines	110	0.49	1.63	0.62
	1	0.48	1.62	0.62
	1	0.47	1.60	0.62
Oman	18	0.47	1.59	0.61
	1	0.46	1.59	0.61
	1	0.46	1.58	0.61

	1	0.45	1.57	0.61
	1	0.45	1.57	0.61
	1	0.45	1.56	0.61
	1	0.44	1.55	0.61
Iran, Islamic Republic of	54	0.43	1.54	0.61
	1	0.43	1.53	0.60
Jordan	103	0.42	1.53	0.60
Italy	6	0.42	1.52	0.60
Albania	6	0.41	1.51	0.60
Tanzania, United Republic of	10	0.41	1.50	0.60
	2	0.39	1.48	0.60
	4	0.39	1.48	0.60
Tunisia	7	0.39	1.47	0.60
	3	0.39	1.47	0.60
	5	0.37	1.45	0.59
	3	0.37	1.44	0.59
Mauritius	15	0.36	1.44	0.59
Mexico	13	0.36	1.43	0.59
	3	0.35	1.42	0.59
Argentina	8	0.31	1.37	0.58
Romania	21	0.31	1.36	0.58
Indonesia	7	0.31	1.36	0.58
	4	0.30	1.35	0.57
Guyana	16	0.29	1.34	0.57
United Arab Emirates	127	0.29	1.34	0.57
	4	0.29	1.33	0.57
Ethiopia	6	0.25	1.28	0.56
	2	0.23	1.26	0.56
	3	0.20	1.22	0.55
Syrian Arab Rep	44	0.19	1.21	0.55
Poland	23	0.19	1.20	0.55
	2	0.18	1.19	0.54
Egypt	1,299	0.16	1.17	0.54
	1	0.16	1.17	0.54
	1	0.16	1.17	0.54
Sudan	374	0.15	1.16	0.54
Malawi	9	0.15	1.16	0.54
	1	0.15	1.16	0.54
Czech Republic	10	0.14	1.16	0.54
	4	0.14	1.15	0.53
	1	0.14	1.15	0.53
	2	0.13	1.14	0.53
Algeria	10	0.12	1.13	0.53
Uganda	16	0.12	1.13	0.53

	1	0.12	1.13	0.53
	1	0.11	1.12	0.53
	1	0.08	1.09	0.52
Libyan Arab Jamahiriya	115	0.07	1.08	0.52
	1	0.05	1.05	0.51
	2	0.02	1.02	0.50
Nigeria	1,800	0.02	1.02	0.50
Armenia	12	0.01	1.01	0.50
Saint Kitts And Nevis	31	0.00	1.00	0.50
Morocco	6	0.00	1.00	0.50
Yemen	9	-0.01	0.99	0.50
Afghanistan	16	-0.03	0.97	0.49
Georgia	14	-0.04	0.96	0.49
Turkey	86	-0.06	0.94	0.49
Belarus	13	-0.07	0.93	0.48
	2	-0.07	0.93	0.48
Palestinian Territories	26	-0.09	0.92	0.48
Ukraine	195	-0.09	0.91	0.48
India	1,160	-0.10	0.90	0.47
Hungary	13	-0.12	0.89	0.47
Pakistan	1,614	-0.15	0.86	0.46
	5	-0.16	0.85	0.46
Sri Lanka	59	-0.18	0.84	0.46
Nepal	195	-0.25	0.78	0.44
Cuba	15	-0.25	0.78	0.44
	4	-0.26	0.77	0.44
Bangladesh	357	-0.27	0.77	0.43
Myanmar	222	-0.36	0.69	0.41
China	248	-0.45	0.64	0.39
Russian Federation	88	-0.47	0.62	0.38
Kyrgyzstan	9	-0.49	0.61	0.38

Table 21: Individual PMQ country of origin estimates of odds of passing PLAB2

Comparisons of PMQ country of origin estimates between null and full models

As we did with PLAB1, if we compare the PMQ country estimates (i.e. the log odds column shown in Table 21) from the full model to those from the null model (i.e. with no predictors), we find a very strong correlation ($r=0.94$, $n=116$, $p<0.001$). Again, this is strong evidence that adjusting for the candidate characteristics we have included (gender, age, sexual orientation etc) does not change the ranking of PLAB2 pass rates by PMQ countries of origin very much.

Comparisons of PMQ country of origin estimates between PLAB1 and PLAB2

We can also compare PMQ country of origin estimates between PLAB1 and PLAB2 – we find weaker relationships (null models, $r=0.26$, $n=115$, $p<0.001$; and for the full models

r=0.32, n=115, p<0.001). This is quite interesting as it suggest that there is substantial variation in rankings of pass rates by PMQ country of origin when comparing between PLAB1 and PLAB2 - despite there being some overall correlation between these.

Brief conclusion

In this analysis we have investigated influences on PLAB1 and PLAB2 pass/fail outcomes. We have typically found larger effects in PLAB2 compared to PLAB1 – a good example being gender, where males have lower pass rates in PLAB2 but are similar to females in PLAB1 (Table 12 compared Table 2). Some of these differences across test formats might be due to language differences (e.g. communication skills), which we have not been fully able to include in this work.

We have also found that pass rates do vary by PMQ country of origin, but that adjusting for a range of candidate factors does not impact to any great extent on differences in these rates. This tells us that these differences in pass rates are due to more general factors than we have been able to include in these analyses. These might include, perhaps, the degree of similarity between the health care systems and medical education systems in the PMQ country of origin to those in the UK – and again, to different levels of language proficiency.

Overall, the analysis does not raise any particular areas of strong concern that might require further specific investigation or obvious changes to PLAB policy.

It is not possible, however, to make a definitive judgment based on these analyses as to whether PLAB is 'fair' to all candidates. In fact, it is difficult to think of an appropriate methodology that would allow such a claim to be made, particularly for PLAB2[17]. It is important to emphasise that many, it not most, of the statistically significant differences that are evidenced in this work are not actually that large in terms of effect sizes. Further, in a modern approach to assessing the validity (and subsequent fairness) of an exam system, evidence must rest on a much wider range of data to that produced by investigating influences of candidate factors on assessment outcomes alone (Cook et al., 2015). With a diverse and globally present exam like PLAB, gathering such a range of evidence would be quite difficult.

[17] A more complex item-level Differential Item Functioning (DIF) analysis might produce some additional findings, but would apply more easily to PLAB1, and would likely be quite problematic in itself for some of the characteristics investigated here.

References

Bates, D., Mächler, M., Bolker, B. and Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*. **67**(1), pp.1–48.

Cook, D.A., Brydges, R., Ginsburg, S. and Hatala, R. 2015. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Medical Education*. **49**(6), pp.560–575.

Haq, I., Higham, J., Morris, R. and Dacre, J. 2005. Effect of ethnicity and gender on performance in undergraduate medical examinations. *Medical Education*. **39**(11), pp.1126–1128.

Wasserstein, R.L. and Lazar, N.A. 2016. The ASA's Statement on p-Values: Context, Process, and Purpose. *The American Statistician*. **70**(2), pp.129–133.