Forecasting the impact of the mode of use of the UKCAT on medical school entrant demographics

Paul Tiffin, Lisa Webster & John McLachlan

October 2013

Executive Summary

The results of previous analyses suggest that the manner in which the UKCAT is used in the admissions process may influence the demographic characteristics of entrants.[1] Consequently we use these findings in order to develop forecasts of what may happen to the demographics of the medical workforce if medical schools were to change their usage style of the test. Our results suggest that if consortium medical schools currently using a 'borderline' or 'factor' approach to the UKCAT scores switched to using the test results as a 'threshold' for interview or offer there would be modest increases (roughly 5-7%) in the proportion of males admitted to 'standard entry' medical courses. In addition, if medical schools currently using a 'factor' approach to the UKCAT scores changed to a 'threshold' approach there may be an appreciable increase in the proportion of ex-state school students admitted to 'standard entry' medical courses (roughly in the order of 10%). There is some evidence to suggest that the level of the threshold selected will also play a role in dictating the impact of adopting this usage style. It should be noted that these forecasts assume that the link between the demographics and the style of UKCAT usage are largely causal. These forecast also assume that no other significant secular trends or 'shocks' will intervene.

Full report

In order to estimate the potential impact on the demographics of UK medical graduates we first looked at where the mode of UK test usage is associated with a significant difference in the demographics of medical school entrants. In order for us to be reasonably confident that these differences are going to translate into long term trends they should be *consistent* across cohorts, especially over the most recent waves of intake. It is important to note that these findings reported only pertain to mainstream medical courses, not graduate entry or other 'widening access' schemes. Such entrants were excluded in order to reduce the risk of bias (which would be compounded due to the relatively high rates of missing sociodemographic data in entrants to these courses).

When we look again at the findings from the 2009 cohort we see that there are two statistically significant differences, *independently* predicted by the use of the test at the institution in key demographics.[1] These are male sex, and 'below average advanced educational attainment' (equivalent to ABB or below at A Levels). In addition there were three other sociodemographic variables that were independently predicted by test usage style, of borderline statistical significance (i.e. p value between .04 and .06). These were socioeconomic status, school type attended and 'English as a second language' (EASL). However, at this point we are more interested in the *raw* (unadjusted) differences in proportions of the entrants with different educational and sociodemographic variables. This is because the reason underlying the difference is not important at this point- only that there is consistent significant difference between those entrants to the three types of medical school, according to their usage style of the UKCAT test. To recap these styles are categorised as:

Borderline: UKCAT score used in borderline cases, as a tie-breaker, or (more rarely) as a "rescue" mechanism to offer interviews to candidates whose applications would otherwise rate poorly—that is, weak use of the test or few candidates affected

Factor: UKCAT used as a factor in deciding to interview or offer, or both (moderate use of the test). In practice this means around 4-33% of the decision to offer an interview or place was based on UKCAT scores. My previous report to the UKCAT Board highlighted that, as expected, a greater weighting factor makes acceptance generally more difficult but favours higher scoring candidates.

Threshold: A threshold score used to decide whether to offer an interview (relatively strong use of test). Thresholds usually range from a total UKCAT score of 1900 to 2730. My previous report to the UKCAT Board highlighted that, as expected, higher thresholds makes acceptance generally more difficult but favour higher scoring candidates. This style appears to be the only one that mitigates against the natural disadvantage experienced by a number of Widening Participation (WP) groups during the admissions process. Moreover, as highlighted in my previous report, admissions outcomes to 'threshold' universities are surprisingly predictable; given the basic demographic and educational characteristics of a candidate we can correctly predict the outcome of a single application event in around 73% of cases. Thus, endogenous factors (variables incapable of being modelled, such as what colour tie you wore at interview) are likely to play a relatively minor role in the selection process.

There has been a tendency towards a stronger use of the UKCAT as a component of the admissions process amongst consortium medical schools, over the period studied; few schools are using it in a 'borderline way, whilst more were using UKCAT score as a threshold as time went on. Moreover, overall, the weight placed on the UKCAT has tended to increase. This is depicted in Table 1.

Year of entry	'Borderline'	'Factor'	Med. Wgt %(range)	'Threshold'	Med.Threshold Score (range)
2007	N=14	N=6	8.5% (4-33%)	N=1	2350 (NA)
2008	N=7	N=9	15% (7-33%)	N=6§	1820 (1800-2200)
2009	N=5	N=11	13% (7-33%)	N=6***	2450 (2420-2680)
2010	N=6	N=9*	16.5% (2-48%)	N=7***	2625 (2570-2730)
2011	N=3	N=9**	20% (2-48%)	N=10	2570 (1900-2940)

Table 1. UKCAT usage style over time, in terms of the number of universities using the approach with median values (with range) for threshold and factor weightings. Data are obtained from the annual survey.

Notes to Table 1:
* As weighting based on rankings in three cases these were excluded
** As weighting based on rankings in two cases these were excluded
*** One med school excluded as threshold based on ranking and in one case
§ One med school excluded as threshold based on ranking and in one case a threshold based on minimum subtest scores was translated into an approximate overall test score threshold

For the purposes of this study we defined an entrant as one where an unconditional offer has been made firm (including a changed course offer [UCAS code UCCF] in some rare cases). We then look at which medical schools have significantly different proportions of entrants. We can visualise this for the 2010 entrants in the following figure taken from a previous paper [1] (see Figure 1).

Defining entrants in this way there should be a unique relationship between a candidate and a university. Thus, we checked for duplicates. A small number of these were identified. For example, in the data relating to the 2011 entrants 16 pairs of duplicates were found. These appeared to be accepting more than one 'UF' offer at same university! These were related to university code 503. Consequently one set of duplicates for each pair of observations were dropped.

In order to build up a picture of which differences between university types are consistent enough to support plausible forecasts we managed, checked and cleaned data for cohort that entered medical school from 2007 to 2011 (inclusive). Thus, we were able to use a simple logistic regression to test whether the sociodemographic status of an entrant (as defined here) could be predicted by the category of university they were accepted to ('borderline'/'factor'/'threshold'). In effect, the category of university was treated as a factor variable, having three levels. Consequently the 'baseline' category of university had to be switched at least once (e.g. from 'borderline' to 'factor') in order to evaluate all the possible comparison pairings (e.g. 'borderline' vs 'factor'). The findings from this exploration are depicted in Table 1. For simplicity only the significance of the inter-group difference are shown. Unlike previous model building process (where 'true model nesting' had to be ensured) listwise deletion was not used to deal with missing data (i.e. deletion of observation where any values were missing). Thus, as long as the dependent (outcome) variable was not missing the data was analysed using this simple logistic regression approach in order to maximise the information available.

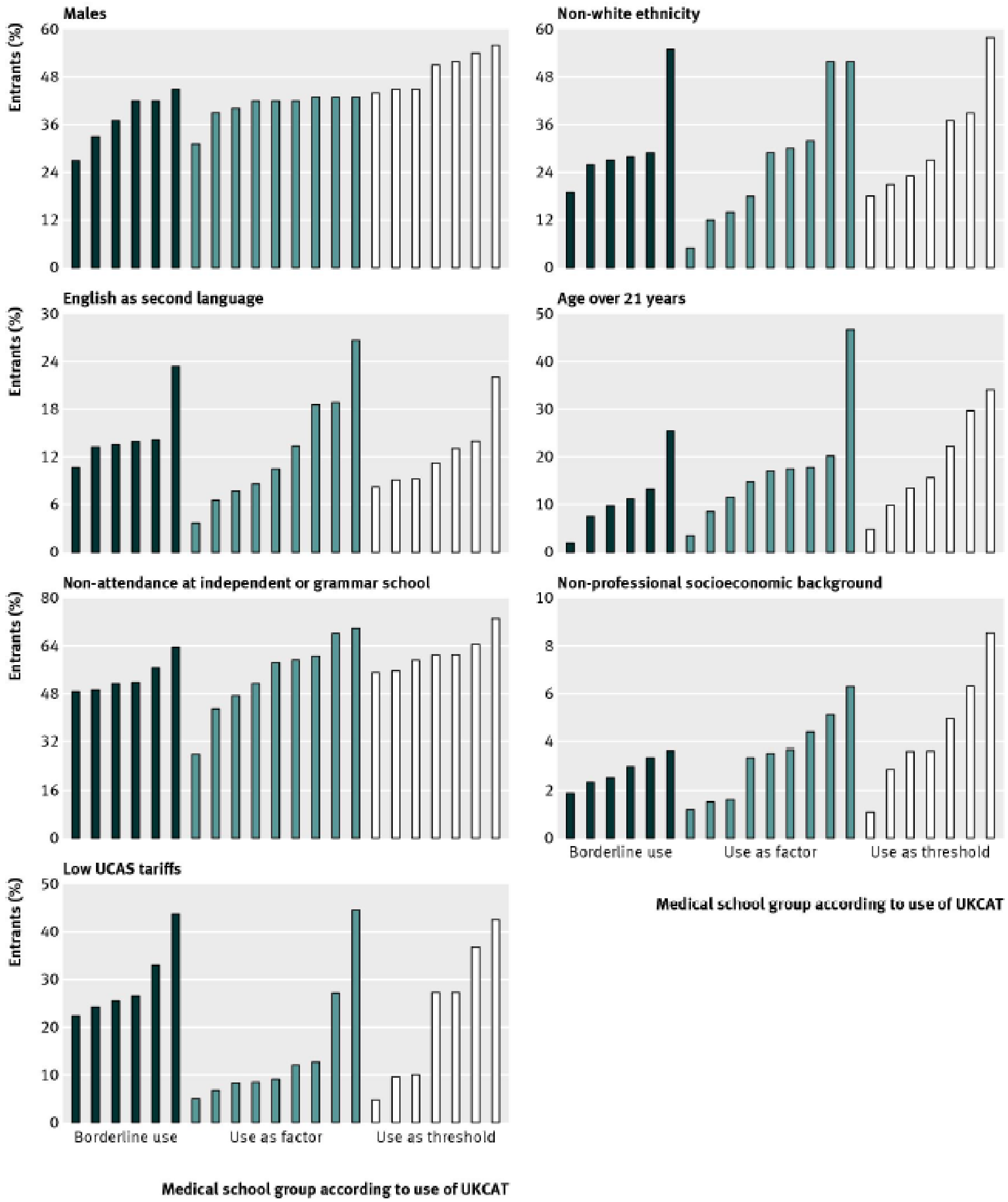

Figure 1 Percentage of medical school entrants who were male, reported non-white ethnicity, reported speaking English as a second language, were aged over 21 years at application, had not attended an independent or grammar school, reported a non-professional socioeconomic background (according to socioeconomic classification system of National Office for Statistics)25, and who obtained relatively low UCAS tariffs (equating to grades ABB or below at A level examinations, or Scottish or Irish equivalents) for each of 22 participating institutions, grouped by use of UKCAT in admissions process.

UCAS=Universities and Colleges Admissions Service; UKCAT=UK clinical aptitude test

2007 entrants			
	borderline vs factor	borderline vs threshold	factor vs threshold
Male	Ns	.02	Ns
EASL-not avail	NA	NA	NA
>20 years	P<.001	P<.001	Ns
stateschool	.01	Ns	.06
Low A levels	Ns	P<.001	P<.001
2008 entrants			
	borderline vs factor	borderline vs threshold	factor vs threshold
Male	Ns	.01	P<.001
EASL-not avail	NA	NA	NA
>20 years	Ns	Ns	Ns
stateschool	P=.01	P=.01	P<.001
Low A levels	P<.001	Ns	.02
2009 entrants			
	borderline vs factor	borderline vs threshold	factor vs threshold
Male	Ns	P=.006	P<.001
EASL	Ns	Ns	Ns
>20 years	P=.01	.001	Ns
stateschool	P<.001	Ns	P<.001
Low A levels	P<.001	Ns	P<.001
2010 entrants			
	borderline vs factor	borderline vs threshold	factor vs threshold
Male	ns	P=<.001	P=<.001
EASL	ns	P=.06	Ns
>20 years	P=<.001	P=<.001	Ns
stateschool	ns	P=.001	P=<.001
Low A levels	P=.03	P=<.001	P=<.001
2011 entrants			
	borderline vs factor	borderline vs threshold	factor vs threshold
Male	ns	P=.002	P=<.001
EASL	P=.005	P=.008	Ns
>20 years	NS	Ns	NS
stateschool	P=<.001	P=.1	P=<.001
Low A levels*	P=.005	Ns	P=.003

Table 2. P values derived from a simple logistic regression with the sociodemographic characteristics of medical school entrants as the outcome variable. This highlights where significant inter-group differences exist for the period evaluated. *N=only 65

In order to visualise and understand where differences are consistent in nature we constructed a series of matrices for each characteristic of interest. These can be seen in Table 2.

Males	**borderline vs factor**	**borderline vs threshold**	**factor vs threshold**
2007		✓	- (p=.1)
2008		✓	✓
2009		✓	✓
2010		✓	✓
2011		✓	✓
EASL	**borderline vs factor**	**borderline vs threshold**	**factor vs threshold**
2007		NA	NA
2008		NA	NA
2009			
2010		-(p=.06)	
2011	✓	✓	
Older (>20 years at application)	**borderline vs factor**	**borderline vs threshold**	**factor vs threshold**
2007	✓	✓	
2008			
2009	✓	✓	
2010	✓	✓	
2011			
School type	**borderline vs factor**	**borderline vs threshold**	**factor vs threshold**
2007	✓		-(.06)
2008	✓	✓	✓
2009	✓		✓
2010		✓	✓
2011	✓	-(p=.1)	✓
Low A levels	**borderline vs factor**	**borderline vs threshold**	**factor vs threshold**
2007		✓	✓
2008	✓		✓
2009	✓		✓
2010	✓	✓	✓
2011	✓		✓

Table 3. ‘Consistency’ matrices. Ticks indicate where a significant inter-group difference exists for the sociodemographic characteristics of medical school entrants. Note: English as a second language (EASL) was not available for 2007-8.

From these ‘consistency matrices’ we see that the only differences really consistently observed over the years are:

1. Differences in the proportion of *male* entrants between universities that use the UKCAT test as a ‘threshold’ vs other styles of usage.
2. Differences in the proportion of entrants from state (non-grammar) schools between universities that use the UKCAT test as a ‘factor’ vs other styles of usage.
3. Differences in the proportion of entrants with below average A level attainment between universities that use the UKCAT test as a ‘factor’ vs other styles of usage.

However, we can discount difference (3); the absolute numbers in the 2011 entry cohort recorded as having below average A level attainment is so small (N=65 nationally) as to render forecasting meaningless. This is partly no doubt an artefact of missing data and also a consequence of excluding ‘non-standard’ entry courses. In addition, the well recognised phenomena of A level ‘grade inflation’ may have played a role in that relatively low grades, such as AAB, are now ‘worth less’ in real terms and therefore are less likely to be acceptable as entry grades. Moreover, it should be noted that the observed and likely impact (small but significant) on medical school demographics of ‘widening participation’ and graduate entry courses has been published in a previous report. [2] Also, note that admission of older (>20 years at application) is not able to be modelled as the most recently available data suggests no intergroup differences between universities in this demographic category (as occurred in at least one other previous year). It is not immediately clear why this is but it may be that older applicants are increasingly targeting non-standard entry courses which have greater availability in recent years.

Focussing on these key differences between types of universities we wish to:

a) Estimate the likely magnitude of any difference in proportion of males/state school students
b) Estimate our level of uncertainty regarding this estimate (in terms of standard error)

Forecasting the impact of UKCAT usage style on the 'maleness' of UK medical entrants

Firstly, we evaluated the proportions and absolute numbers of male entrants to each type of medical school:

Year of entry**	'Borderline'	'Factor'	'Threshold'
2007 (B=14; F=6; T=1)	41.45% (1123/2709)	43.68% (494/1131)	48.55% (134/276)*
2008 (B=7; F=9; T=6)	42.08%(611/1452)	40.33% (642/1592)	46.76% (634/1356)
2009 (B=5; F=11;T=6)	42.95% (341/794)	40.98% (913/1315)	49.10% (655/1334)
2010 (B=6; F=9; T=7)	38.85% (399/1027)	41.55% (745/1793)	49.73% (822/1653)
2011 (B=3;F=9;T=10)	42.07% (204/490)	42.54% (799/1864)	49.39% (1040/2103)

Table 4. Percentage (%) and proportions of male entrants admitted to each type of medical school over five successive cohort. All inter-group differences statistically significant at p<.05 on a logistic regression unless otherwise stated.

*NS, p=.1 for intergroup difference between proportion of males admitted to 'factor' versus 'threshold' universities (note only one of the latter type in 2007)
**B=No. of universities using the test in borderline way; F= No. of universities using the test as a factor; T= No. of universities using the test as a threshold.

We can also graph these percentages in order to visualise any trends (see Figure 2).

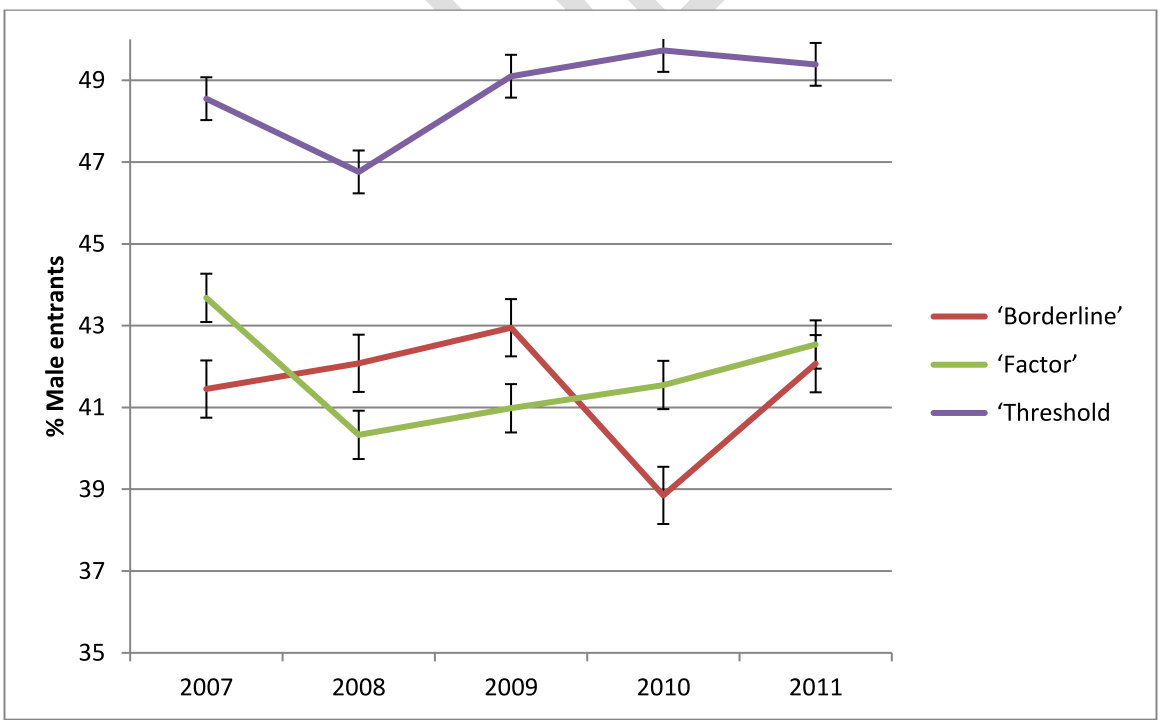

Figure 2: Trends for proportion of medical entrants who are male for each type of medical school (according to UKCAT usage style). Standard error bars are shown for the data points in each series.

As can be seen from Figure 2, there are rather stable (and significant) differences in the proportion of males that are admitted to universities using the UKCAT as a threshold score compared to the other two types of university. Assuming no underlying trend we can say that, on average 'threshold universities admit around 7% (7.06% to be more precise) more men than women.

How certain can we feel about this estimate? It should be noted that the standard error (SE) bars on Figure 2 are generated by Microsoft Excel for each data point and represent the SE for each point within its series (in effect the SD divided by the square root of N [5 in this case] for each series). There are other ways of generating SEs (and hence 95% confidence intervals) in this particular situation. For example, as either a simple SE of a proportion (this is the proportion of males multiplied by the proportion of females divided by total N and the square root taken). Another option would be to create a weighted SE for each group of medical schools so that the SE of larger schools would be given more importance that smaller ones. However, if we look at the simple SE for each group at the latest time point (2011 entrants) we will have a reasonably good idea how confident we can be in our forecast. In this case our standard error for the threshold group is calculated as a proportion of .0108, equating to around 1%. Thus, we can 95% confident that the proportion of males entering medical schools using the UKCAT as a threshold in 2012 was somewhere between 51% and 47% (hopefully the data, when available, will support this!). We can repeat this for the other two types of university; pooling the data for these two latter groups we derive a standard error of almost precisely 1%. Therefore we can be 95% confident that the proportion of male entrants to this type of medical school will lie roughly between 40 and 44%. It should be noted that these estimates ignore variation in proportions between medical schools in the same group, but at this point we are interested in forecasting *overall* effects on medical school demographics rather than individual university intakes, where there would be a greater degree of uncertainty. Assuming independence (i.e. no underlying connection between these two proportions we could be 95% confident that, for the foreseeable future (should this trend remain) the difference between male intake to a 'threshold' type university and the other types will be somewhere between 3 and 11% (if an extreme value for both is observed) but is actually likely to be nearer to the mean of 7%.

Translating these proportions into absolute numbers; should 'non-threshold' medical schools use this approach to the UKCAT then it could result in an average additional male intake to consortium universities of around 165 men per year. Over five years this might equate to 825 extra male doctors. Naturally this assumes a causal relationship between the style of UKCAT usage and the proportion of male entrants. However, given what we already know about the sociodemographics predicting test performance the link is plausible.

Of course these predictions only apply to mainstream courses and non-graduate entrants, although these individuals make up the majority of medical students.

Forecasting the impact of UKCAT usage style on the proportion and number of state school medical entrants

As before, we evaluate the proportions and absolute numbers of state school entrants to each type of medical school. Note, that in case, our denominators will differ according to the proportion of missing information on school type attended (Table 5):

Year of entry**	'Borderline'	'Factor'	'Threshold'
2007 (B=14; F=6; T=1)	57.16% (1540/2694)	61.58% (694/1127)	55.31% (151/273)
2008 (B=7; F=9; T=6)	52.00% (689/1325)	47.28% (668/1413)	57.01% (699/1226)
2009 (B=5; F=11;T=6)	50.77% (361/711)	43.10% (802/1861)	52.57% (594/1130)
2010 (B=6; F=9; T=7)	53.65% (551/1027)	50.98% (914/1793)	60.31% (997/1653)
2011 (B=3;F=9;T=10)	52.00% (195/375)	38.92% (525/1349)	47.23% (725/1535)

Table 5. proportions of state school entrants admitted to the different types of consortium medical schools for the period 2007 to 2011 (inclusive)

Although significant intergroup differences exist at most time-points the data tell a different story as can be seen in Figure 3 (below).

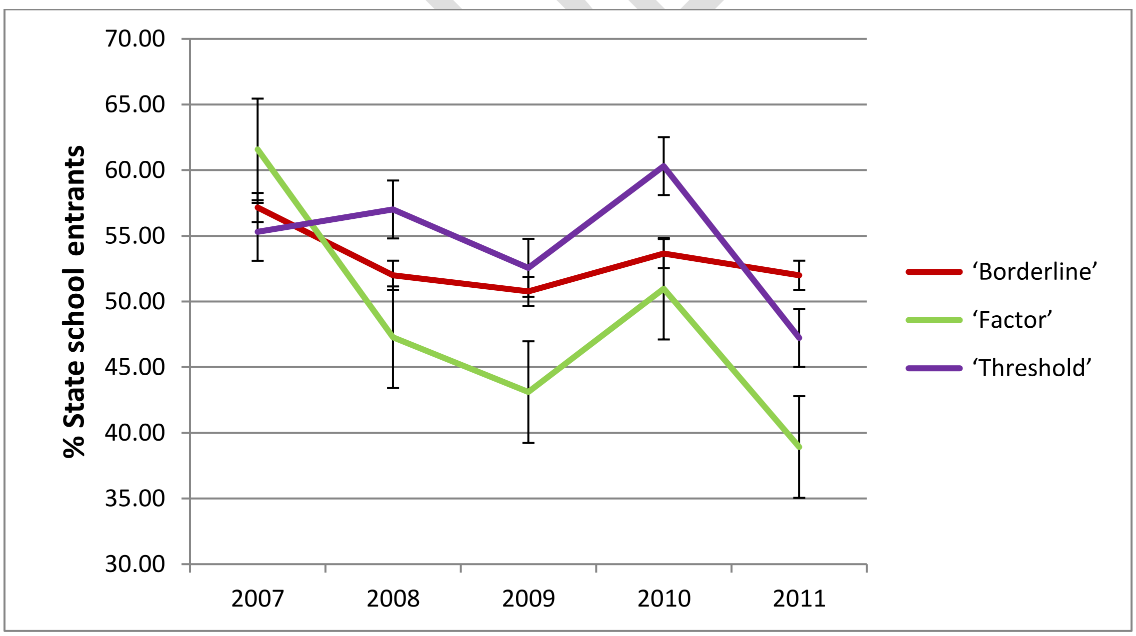

Figure 3: Trends for proportion of medical entrants who report being from a state school (excludes grammar school) for each type of medical school (according to UKCAT usage style). Standard error bars are shown for the data points in each series.

The data depicted in Figure 3 is, in some ways, more challenging to interpret and to use in forecasting than that for male entrants. Firstly, whilst those universities with a weak ('borderline') use of the UKCAT have consistently admitted around 52% of entrants from state school there appears to be some secular trends at work for the other two types of medical school. The 'factor' and 'threshold' universities have tended to 'mirror' each other, albeit at different levels. The exception to this is for 2007, though this can be discounted given that only one institution at the time was using a 'threshold' approach to the UKCAT score. All groups show some spike in the proportion of state school students admitted in 2010, although the overall trend has been for a lower percentage of such individuals to enter medical school. It is not immediately clear why those medical schools using the UKCAT as a 'factor' in the admissions process consistently admit a lower proportion of state school students compared to those using a threshold approach. It may be that usage is acting as an 'instrumental' variable with the latter more committed generally to widening participation. Additionally, the use of a UKCAT threshold could be considered generally a more robust utilisation style and the only one previously found to have mitigated against the observed disadvantage of attending state school when applying to medical school.

In terms of forecasting it is relatively safe to say that, if (and only if) the association between UKCAT usage style and the proportion of state school entrants is causal then switching from a 'factor' to a 'threshold' style of usage around 9% (the average difference between the groups from 2008 to 2011 was 9.21% with an SD of only 0.62%). Indeed the standard error for this series of observed differences in proportions for 2008 to 2011 was only 0.31 (in effect the SD divided by the square root of the number of data points (i.e. 2)). Thus, we can be 95% confident that the observed difference between the two group of medical schools in terms of proportion of state school students will be between 8.59% and 9.83%, all things being equal. Thus we can say switching from a 'factor' to a 'threshold' model may result in approximately 122 state school students per year entering consortia medical school (approximately 610 state schooled individuals over five years).

Thus, although relatively few consistent inter-group differences were observed over the period 2008-2011 two were robust enough to make plausible forecasts relating to the potential impact that changing the category of UKCAT usage style may have on the demographics of the medical profession.

<u>Weighting of the UKCAT Scores in the Admissions Process</u>

It is also worth noting that the use of the UKCAT score as either a 'threshold' or a 'factor' in the admissions process is far from a simple categorisation. Within each class of medical school the UKCAT score can be dealt with very differently. For example, the UKCAT is sometimes being used as a threshold for interview or as a threshold to make an offer. In my previous I observed that higher thresholds and factor weightings makes acceptance generally more difficult but favour candidates obtaining higher scores on the UKCAT (which intuitively makes sense). Nevertheless, there is relatively little information available to usefully tease apart the effect of differing approaches to weighting *within* each class of university on final entrant demographics. However, it may be useful to test out whether there is some indication of an effect of weighting on the demographics of entrants in the two key categories that we have focussed our forecasting efforts on (male sex and state education) whilst preserving institutional anonymity. When we use the most recently available data (2011 entrants) for these analyses we observe that the weighting of the factor has no discernible impact of either the odds of being a male entrant or from a state school background. However, there is at least some indication of a (statistically non-significant trend) towards a higher threshold predicting greater odds of an entrant being from one of these two groups within the intake of universities using this particular approach the UKCAT scores. The results of these two univariate (simple) logistic regression analyses are depicted in Table 6. We can interpret the results in Table 6 by saying that, on average, raising the UKCAT threshold score by 100 points may increase the odds of an entrant being male, or being from state school by roughly 3%. Of course, given the complexity of the situation and the lack of study power in this instance we must interpret such findings very cautiously, though they may provide some indications of the underlying averaged effects at work.

Outcome	Odds Ratio	95% Confidence Intervals	P
Entrant being male	1.03	.99 to 1.07	.1
Entrant reporting being from state school	1.03	.98 to 1.08	.3

Table 6. Results of a univariate logistic regression to investigate whether the level of the UKCAT threshold score (here used as total score/100 to assist interpretation of the ORs) used in admissions is associated with the raw (unadjusted) odds of being either a male or state schooled entrant.

<u>Further Reflections</u>

Given the overall trend for an increasingly strong use of the UKCAT test in the admissions process (outlined in Table 1) we may also ask “what might be the consequences of universally using a high UKCAT threshold score in the medical school admission process?”. Firstly, we might speculate whether setting a high minimum A-Level (or equivalent) as well as a high minimum UKCAT score may lead to ‘cookie cutter’ entrants- i.e. increase the level of homogeneity within medical students. This may or may not be desirable given the attributes required for the widely differing roles within medicine and is a question for medical school admissions boards to consider carefully. Secondly, it may be that as the pool of ‘desirable’ applicants shrinks to those with the requisite A level grades *and*

UKCAT scores the proportion of candidates declining place offers increases, especially for those medical schools that may be perceived by candidates as less desirable. Modelling such an effect would be difficult and some metric of institution 'perceived desirability' would be needed. However, what may be useful is to tentatively consider what such medical students may look like. We can do this very crudely by comparing the basic characteristics of those entrants to the University with the highest threshold compared with the lowest threshold.
The results are depicted in Table 6.

Variable	% Male	State school	Mean standardised A level tariff (sd)	Mean UKCAT Score (sd)	Non-white	Non-professional background
High Threshold	50%	53%	.29 (.74)	2874 (225)	23%	6%
Low Thresh.	43%	51%	.49 (.53)	2575	25%	3.5%

Table 6. Demographics of the entrants to the two universities with the highest and lowest UKCAT threshold scores respectively. Note A level performance is given as a standardised score (mean=0 sd=1) for all medical applicants to standard consortium medical courses in that year).

As expected we note a higher proportion of males in the 'high threshold' medical school' and a slightly higher rate of state school individuals. We also note that, although UKCAT score is higher (as expected) A level performance is slightly lower, Indeed when we look at a non-parametric correlation test (Spearman's Rho) we can observe a positive relationship between threshold level and UKCAT score (rho=.4, p<.001) and a very slight (but borderline significant) inverse correlation with standardised A level tariff (rho=.05, p=.04). This suggests that medical schools setting high thresholds do not necessarily obtain candidates with better A level performance, and indeed, there may be a slight trend in the opposite direction. Therefore, it might be that although raising a threshold might be expected to increase the homogeneity of medical school entrants, this may be a different kind of homogeneity to that obtained by raising 'the academic bar' in terms of A level (or equivalent) achievement.

Conclusions

1. Consortium medical schools currently using a 'borderline' or 'factor' approach to the UKCAT scores may wish to consider to switching to a 'threshold' approach if they wish to address the current sex-inequalities apparent in the medical admissions processes. This may result in modest but appreciable increases in the proportion of males admitted to 'standard entry' medical courses.
2. Consortium medical schools currently using a 'factor' approach to the UKCAT scores may wish to consider to switching to a 'threshold' approach if they wish to address the under-representation of state schooled

individuals. This may result in an appreciable increase in the proportion of ex-state school students admitted to ‘standard entry’ medical courses, roughly in the order of 10%. There is some evidence to suggest that the level of the threshold selected will also play a role in dictating the impact of adopting this usage style.

3. Two ‘health warnings’ should be noted.
    a. These forecasts assume that the link between the demographics and the style of UKCAT usage are largely causal. Case studies of situations where an institution changed its usage style could support or refute this assumption.
    b. These forecast assumes that no other significant secular trends or ‘shocks’ will intervene (for instance the impact of education, societal or economic influences) and observed recent trends are likely to continue.

References

1. Tiffin PA, Dowell JS, McLachlan JC. Widening access to UK medical education for under-represented socioeconomic groups: Modelling the impact of the use of the UKCAT in the 2009 cohort. BMJ. 2012; **344**: e1805.
2. Mathers J, Sitch A, Marsh JL, Parry J. Widening access to medical education for under-represented socioeconomic groups: population based cross sectional analysis of UK data, 2002-6. BMJ. 2011; **342**: d918.