

Fairness of decisions to refer doctors to The
Medical Practitioners Tribunal Service Interim
Orders Tribunal

Prepared for the General Medical Council
By UCL Medical School
July 2018



Contents

	Page
Executive summary	3
Introduction	5
Experimental phase	6
Qualitative phase	16
Conclusion	36
References	38
Appendix 1- Study information sheet for UCL participants	39
Appendix 2: UCL case examiners instructions	41
Appendix 3: Standardised form for UCL case examiners	44
Appendix 4: Demographic information about GMC case examiners over period of active cases	45
Appendix 5: GMC case examiners invitation	46
Appendix 6: UCL examiners interview questions	47
Appendix 7: GMC examiners interview questions	48

Fairness of decisions to refer doctors to The Medical Practitioners Tribunal Service Interim Orders Tribunal

Executive summary

To investigate reasons behind why non-white and overseas doctors are consistently over represented in the GMC Fitness to Practise procedures, the GMC commissioned a series of projects looking at 'fairness of decisions'. This project concentrates on the decision making process around whether case examiners refer or do not refer doctors to the Interim Orders Tribunal (IOT). It specifically looks at whether there was any bias against the over represented groups. It also investigates the reasoning behind the decision making process in terms of decision making theories. It does not analyse the definition of a 'fair decision' or look for specific reasons for possible bias.

There were two parts to this study. The first part was a pilot experimental study which aimed to test whether a study of this design would be a practical way of shedding light on whether GMC case examiner decisions were influenced by the ethnicity or country of primary medical qualification (PMQ) of doctors under investigation when referred doctors to the IOT. Due to the pressure of their responsibility for ongoing cases, GMC case examiners were unavailable to take part thus a compromise on a preferred method was used. In this compromise, UCL staff were given GMC case examiner training and re-made GMC decisions on 120 real cases. The UCL decisions were then compared to the GMC decisions. Before making this comparison, it was necessary to check for any evidence of bias in UCL decision making to ensure the comparison was valid.

To see whether the ethnicity or country of primary medical qualification (PMQ) influenced the UCL case examiner decisions, the cases were duplicated and the ethnicity or country of PMQ changed on those duplicates without the knowledge of the case examiners. White doctors were presented as being from a black and minority ethnic (BME) group (or vice versa), and UK PMQ doctors were presented as have a non-UK PMQ (or vice versa). If decisions for the same cases were systematically different when the doctors were presented as white or BME, or as having a UK or non-UK PMQ, this could suggest bias. If there were no differences, this could suggest a lack of bias.

After establishing that there was no evidence of bias in UCL decisions, the decisions of the UCL case examiners were then compared to the real GMC case examiner decisions. If the UCL case examiners did not show bias, and their decisions were not systematically significantly different from the GMC decisions with respect to doctors grouped by ethnicity and by PMQ, this could indicate a lack of bias in the GMC decisions. If the UCL case examiners did not show bias and their decisions were systematically different from the GMC decisions with respect to doctors grouped by ethnicity and by PMQ, this might indicate bias in the GMC decisions.

The results found no evidence that the UCL case examiners' decisions were influenced by the ethnicity or PMQ of the doctors under investigation, suggesting UCL case examiner decisions were unbiased. GMC case examiners referred slightly fewer white UK graduates than UCL case examiners,

but referred similar proportions of BME UK graduates. GMC case examiners also referred slightly fewer non-UK doctors than UCL case examiners, but referred similar proportions of UK graduates.

Interpreting these results is not straightforward. If the GMC case examiners were biased in favour of white doctors (i.e. under-referred them compared to UCL case examiners, whose decisions were seemingly unaffected by the ethnicity of the doctors), it would reasonably be expected that GMC case examiners would also be biased in favour of UK doctors. The pattern of results however suggests that if anything, GMC case examiners under-referred non-UK doctors, the majority of whom are BME doctors, compared to UCL case examiners, which is difficult to explain by reference to GMC case examiner bias.

Importantly, the results should also be treated with caution for a number of methodological reasons. This (experimental design) part of the study was a pilot to test whether it was practically possible to conduct a study of this design that would answer the research questions, and therefore involved a relatively small number of cases. This small number negatively affects the reliability of the findings. This is particularly important since the results were only weakly statistically significant, and in fact were no longer so after a statistical correction (although that correction can sometimes result in the rejection of a truly significant result, and therefore should itself be treated with some caution).

Another limitation is that several changes were made to case examiner training and guidance over the time frame of cases being analysed, and thus UCL case examiners were using different criteria from the GMC case examiners to assess the older cases. In addition, UCL case examiners had not been selected against a set of competencies using the selection criteria used for selecting GMC case examiners and were undertaking the task for the first time, whereas GMC case examiners were much more experienced. Perhaps partly as a result of their inexperience, UCL case examiners overall referred many more doctors compared to GMC case examiners. With the relatively small number of cases involved, this over-referring could affect the overall findings. It is unknown how risk-averse the various different GMC case examiners were. Furthermore, it is likely that some cases were easier to make decisions about than others, and it could be that case difficulty related to the ethnicity and/or PMQ of the doctor.

The practical difficulties involved in redacting and falsifying cases made it impossible due to the resources that would be required to progress to a full study with a large number of cases. Based on these results from the pilot, it is not possible to say with any degree of certainty whether or not the decisions made by the GMC case examiners to refer a case to the IOT were influenced by the ethnicity or PMQ of the doctor.

The qualitative phase was clearer to interpret. There is clear evidence of good practice in case examiner training and the guidance provided to case examiners. There is clear use of systemic processing when writing their justification. There was evidence of the use of heuristic processing in decision making of GMC case examiners, this was in keeping with the guidance and training provided to case examiners. There was no evidence that the slightly lower referral rate for white doctors and overseas-trained doctors was due to bias on the part of the case examiners.

Introduction

The GMC has evidence that suggests non-white doctors and overseas trained doctors are over represented in their Fitness to Practise investigations. This project was designed to investigate an aspect of this finding.

It used mixed methods to review the fairness of the decisions made to refer doctors to The Medical Practitioners Tribunal Service Interim Orders Tribunal (IOT). The purpose was to determine whether differences in the proportion of cases of non-white doctors/overseas trained referred to IOT (as compared to white/UK trained doctors) involved any bias on the basis of the information at the time a complaint was received.

The two main research questions were:

- 1) Consider if the greater likelihood of BME UK qualified doctors being referred compared to other groups of doctors can be justified by the guidance for referral to IOT based on the information available to the decision maker at the time.
- 2) Consider if the greater likelihood of doctors with a PMQ from outside the UK being referred compared to other groups of doctors can be justified by the guidance for referral to IOT based on the information available to the decision maker at the time.

To try to answer these questions we performed two parts to this research project; an experimental and qualitative phase.

The experimental phase was a pilot quantitative study which looked at potential differences in referral rates to IOT based on ethnicity and country of primary medical qualification (PMQ) and aimed to test whether a study of this design in this area could practically answer the research questions.

The qualitative study used decision science frameworks to explore the cognitive processes that influenced case examiner's decisions to make (or not) a referral to IOT looking in particular for any evidence of bias in the decision making process.

Experimental phase

Introduction

The aim of this part of the research was to see whether GMC case examiner decisions to refer to the IOT were influenced by the ethnicity or the PMQ of the doctor.

A good test of this would have been to give GMC case examiners –without them realising - duplicate cases: one with the doctor's true ethnicity or PMQ and one with the doctor's ethnicity or PMQ falsified, and then see whether - over a sample of cases - case examiner decisions were the same regardless of what they believed the doctor's ethnicity or PMQ to be.

This was not possible due to the pressure of their responsibility for ongoing cases, so we adopted an alternative approach: training UCL employees to act as case examiners and remake the GMC's decisions, and then comparing the UCL and GMC decisions about white and BME UK graduates, and UK and non-UK graduates. For example, if we found that UCL case examiners referred the same proportion of BME and white doctors as the GMC, that could indicate the GMC decisions were not influenced by a doctor's ethnicity. The potential problem with this approach is that it could indicate that the UCL case examiners were biased in the same way as the GMC examiners.

We therefore added an extra step to attempt to counter this problem:

- First, we assessed whether UCL case examiners were influenced by a doctor's ethnicity or PMQ (i.e. whether the UCL case examiners were biased).
- Then we looked at whether the UCL decisions were systematically different or similar to GMC decisions with respect to doctors grouped by ethnicity and by PMQ.
- From this, we inferred whether the GMC decisions were biased.

For example, if we found that the UCL case examiner decisions were not influenced by a doctor's ethnicity/PMQ, AND the UCL decisions were similar to the GMC decisions, it would give us stronger grounds for saying that the GMC decisions were unlikely to be biased (although we still could not say that definitely).

The original plan was to do a two stage approach to this research; with an initial pilot study looking at sample of 120 cases and helping to establish the practicality of the study design before proceeding to a much larger sample. After detailed discussion with the GMC about the challenges that were identified in the pilot phase and therefore the practicalities of doing the second phase, it was agreed that it would not be possible to perform the analysis beyond the initial 120 cases.

Methods

Participants

Employees of UCL Medical School acted as case examiners. Invitations to participate in this research project were sent to all UCL medical school staff in November 2015. (Appendix 1). Volunteers were asked to review 2 sets of 12 cases but were unaware of the research questions or methodology of

this project. They all attended the standard GMC two days of training for case examiners in how to undertake IOT decisions in January 2016 and had refresher training before given their cases.

To ensure each UCL case examiner’s decision was made independently and could be treated as such in the statistical analysis, and to help ensure confidentiality, volunteers were instructed not to discuss any aspects of the cases. There was one exception: in recognition of the fact that cases could be distressing, each volunteer was provided with a named member of the research team with whom they could discuss the content of cases but they were asked not to discuss anything that might influence the case examiner’s decision. Volunteers were given clear instructions on how to approach the cases (Appendix 2). All of the guidance documentation, examiner training presentations, cases and forms to document their findings were provided on an encrypted memory stick. (Appendix 3)

Two of the volunteers were not able to complete the initial batch of ethnicity cases; one left UCL, the other was not able to complete due to time commitments. One examiner who completed the ethnicity cases was not able to review the PMQ cases. Another email was sent to all UCLMS staff in November 2016 looking for additional volunteers. We recruited another 7 case examiners; they attended two days of the same training in January 2017. One examiner was not able to complete the PMQ cases in the time frame.

Table 1: Demographics of the UCL case examiners.

Case Examiners		Ethnicity cases	PMQ cases
Total number of staff participating		10	9
Sex	Male	2	1
	Female	8	8
Age	Under 20 years	0	0
	Age 20 to 29	5	5
	Age 30-39	4	4
	Age 40-49	1	0
	Age 50 and above	0	0
Broad ethnicity category*	White	6	6
	Mixed/Multiple ethnic groups	2	1
	Asian/Asian British	1	1
	Black/ African/Caribbean/Black British	1	1
	Other ethnic group	0	0

The demographic information about GMC case examiners over the time period that cases were active can be found in Appendix 4.

Materials

The GMC provided 120 cases. Those 120 cases were randomly selected by the GMC from the entire pool of cases closed between 2007 and 2014. The random selection was done as follows:

To compare white and BME UK graduates, all UK PMQ cases were first grouped by ethnicity (white and BME), and then within each ethnic group, cases were further categorised into those referred and those not referred. This gave a total of four groups (white UK referred n=779, white UK not referred n=534, BME UK referred n=228, BME UK not referred n=160). From each of those four groups, 150 cases were selected using a random number generator in Excel. The first 15 of each of those 150 randomly selected cases were then chosen for the research. This resulted in a total of 60 cases (15 white UK referred, 15 white UK not referred, 15 BME UK referred, 15 BME UK not referred). Those 60 cases were referred to as the 'ethnicity group'.

To compare UK and non-UK graduates, all cases were grouped by PMQ (UK and non-UK). The UK graduates chosen for the ethnicity group (above) were rejected to avoid case duplication. Within each PMQ group, the cases were categorised into those referred and those not referred, giving a total of four groups (UK referred n=857, UK not referred n=394, non-UK referred n=1348, non-UK not referred n=715). The same procedure was then followed to choose 15 cases from each of the groups, resulting in a total of 60 cases (15 UK graduates referred, 15 UK graduates not referred, 15 non-UK graduates referred, and 15 non-UK graduates not referred). Those 60 cases were referred to as the 'PMQ group'.

In both the ethnicity and the PMQ groups, two of the originally chosen files needed to be replaced. For example, one case was specifically about English language difficulties which would have made it obvious they were unlikely to have attended a UK medical school. The removal of cases was agreed by the GMC and UCL.

The ethnicity group contained 28 white UK graduates of whom 43% had been referred to the IOT, and 32 BME UK graduates of whom 53% had been referred to IOT¹. The PMQ group contained 30 UK graduates (of whom 47% had been referred) and 30 non-UK graduates (outcome data was missing for 2 cases in the materials shared for the research).

We treated the 60 ethnicity cases and the 60 PMQ cases as completely separate groups. This meant we were able to look at whether ethnicity influenced case examiner decisions, and whether PMQ influenced case examiner decisions, separately.

We asked UCL case examiners to re-make the GMC's decisions (without them knowing what the real GMC case examiner decision had been) for the 60 ethnicity cases and the 60 PMQ cases.

¹ It had originally been decided to have 30 white and 30 BME doctors, with 50% referral rates in each group but there were not enough white UK graduates who had been referred in the GMC's pool of cases.

We assessed whether the UCL case examiners were influenced by the doctor's ethnicity as follows:

- Without the UCL case examiners' knowledge, we created duplicates of the 60 cases and reversed the ethnicity: white doctors were presented as BME doctors and BME doctors were presented as white doctors. All other aspects of the case were identical.²
- This resulted in 120 ethnicity cases for case examiners to make decisions about: 60 where the presented ethnicity was the doctor's true ethnicity and a duplicate 60 cases where the presented ethnicity was reversed.
- We allocated cases at random to UCL case examiners, with the proviso that no examiner could have the two versions of the same case. Each UCL case examiner received 12 cases.
- To assess whether the UCL case examiners were influenced by ethnicity, we compared the referral rates for the same cases when the doctor was presented as white and when the doctor was presented as BME. If ethnicity was not systematically influencing case examiners decisions (i.e. there was no bias), the same cases should have the same outcomes regardless of what case examiners believed the doctors' ethnicity to be.

Having established whether or not UCL case examiner decisions were influenced by ethnicity, we then compared the UCL case examiner referral rates for white and BME doctors with the GMC referral rates for white and BME doctors.

We then repeated this process for the 60 PMQ cases, presenting UK graduates as non-UK graduates, and non-UK graduates as UK graduates.

Analysis

To consider the similarity between UCL and GMC case examiner decision-making we conducted McNemar's test. We did this separately for four groups of cases dependent on the presented and true ethnicity of the doctor (presented white true white, presented BME true BME, presented BME true white, presented white true BME). We repeated this for the PMQ cases.

To consider whether UCL case examiner decision-making was influenced by the ethnicity/PMQ of a doctor, we conducted a logistic regression (0=did not refer, 1=referred) on the true ethnicity/PMQ of the doctor (white vs BME/UK vs non-UK), whether the ethnicity/PMQ presented was the same or different to the true ethnicity/PMQ, and the interaction between those two predictor variables. Statistical analyses were conducted in SPSS v21.

² Each case had a GMC coversheet which provided the doctors' name and demographic details. Three administrators familiar with GMC work re-typed the name and ethnicity or PMQ information on each sheet, for example changing 'White British' to 'Asian or Asian British'. Each doctor was also given a false name for the purposes of the study. The names were created to match the presented ethnicity/PMQ of the doctor, for example a White British doctor was given the fake name Nicholas Howells and when the same doctor was falsely presented as Asian or Asian British he was given the fake name Sandeep Kapur.

Results

Ethnicity cases

Comparison between UCL and GMC case examiner decisions

It is clear from Figure 1 that UCL case examiners referred significantly more cases than GMC case examiners in general. Across all 120 cases, UCL case examiners referred 81/120 (67.5%) cases, compared to the GMC referral rate of 29/60 (48.3%) (McNemar's test $p = 0.002$).

Figure 1 suggests that UCL case examiners referred more white doctors than the GMC, but they referred a more similar proportion of BME doctors. It can be seen below in Table 3 that this is unlikely to be because UCL case examiners were biased against white doctors.

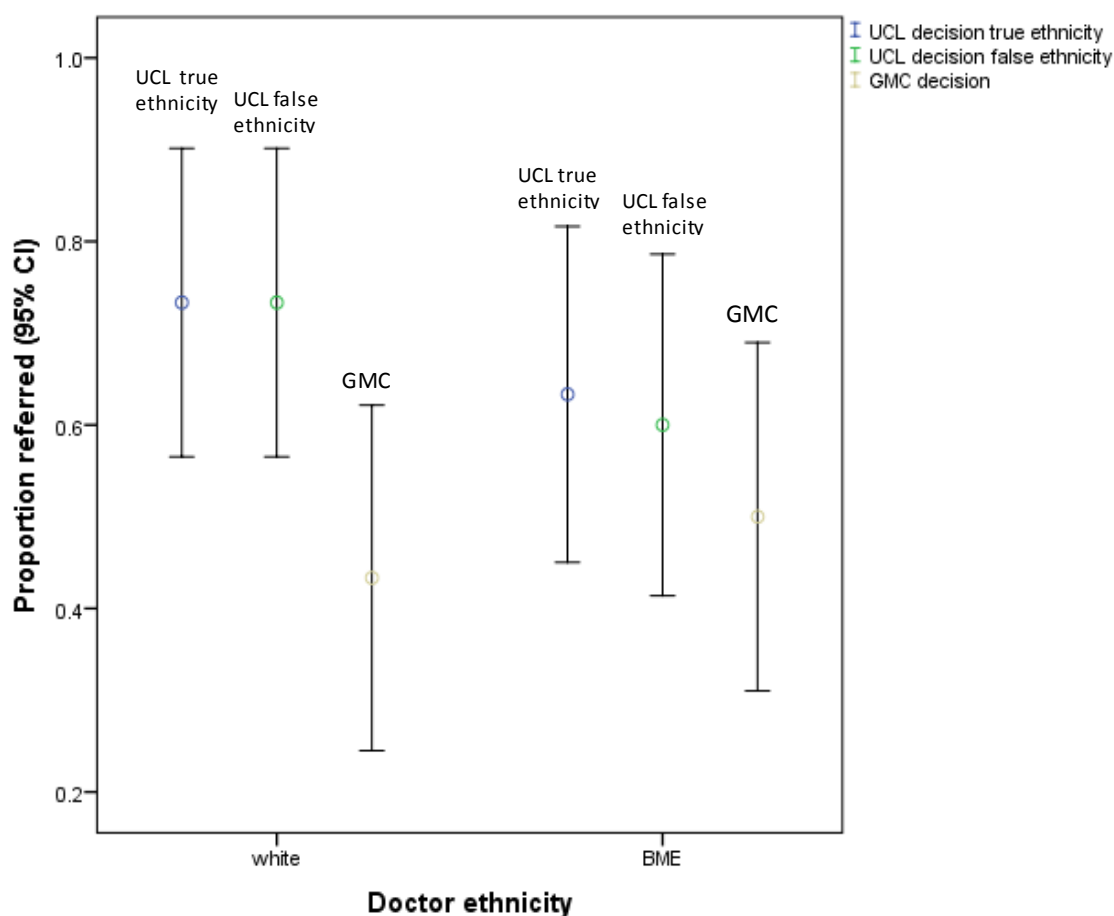


Figure 1: Referral rates for white and BME doctors by ethnicity

McNemar's test allows us to statistically compare the proportion of cases referred by UCL and the GMC, for white and BME doctors. The results are presented in Table 2.

For white doctors, the McNemar tests are significant at the 0.5 level, indicating a significant difference between UCL and GMC case examiner decisions (white doctors presented as white $p=0.022$; white doctors presented as BME $p=0.035$).

For BME doctors, the test was not significant at the 0.5 level (BME presented as white $p=0.774$; BME presented as BME $p=0.549$).

The difference for white doctors is not large, and indeed it is not statistically significant with a Bonferroni-corrected p value of 0.0125 (the correction takes into account the probability of getting a statistically significant result by chance alone when performing multiple statistical tests –see Box 1).

Box 1: Bonferroni correction for multiple comparisons

Experiments are designed to test hypotheses using statistical methods, for example to test whether GMC case examiners were influenced by doctors' ethnicity in their decisions to refer to the IOT. Statistical significance testing allows us to estimate the probability that a result is due to random chance, and therefore whether a hypothesis can be supported.*

The level of probability generally accepted as statistically significant is 5% or $p=0.05$. This means 19 times out of 20 the result of our experiment is real and not just due to random chance or to put it another way, 1 time out of 20 it is in fact just due to random chance. A more stringent level often set is 0.1% ($p=0.001$), which means 999 out of 1000 times the result would be real, and only 1 time out of 1000 would it be due to random chance. A result that is significant at 0.001 is therefore 'more' significant (less likely to be due to random chance) compared to one significant at 0.05. These significance levels are effectively arbitrary - the difference between $p=0.049$ (accept the hypothesis) and $p=0.051$ (reject the hypothesis) is probably meaningless in real life - however they are widely used.

A problem with significance testing is that the more statistical tests you do, the more likely you are to get a result that is due to random chance. The most common approach to tackle this problem is to do a Bonferroni correction for multiple comparisons, which simply means that you set a more stringent significance level. If you do two tests, you halve the $p=0.05$ significance level to $p=0.025$; if you do four tests, you quarter it to $p=0.0125$. This correction can however be overly conservative, which means you run the risk of rejecting a truly significant result by mistake. That is why it is important to show both the corrected and the uncorrected results.

A non-significant corrected result suggests that the result is due to random chance; however there may still be some doubt about rejecting the finding completely due to the overly-conservative nature of the correction. In other words, even if the uncorrected result were significant, the fact that it is no longer significant after correction casts considerable doubt upon its reliability but does not mean it can be definitely rejected.

Furthermore, when interpreting any statistical result it is always important to consider aspects of how the experiment was carried out which may affect its accuracy.

*It is more correct to say that experiments are designed to test whether a null hypothesis can be rejected. A null hypothesis is the opposite of the hypothesis e.g. whether the GMC case examiners were *not* influenced by doctors' ethnicity in their decisions to refer to the IOT. Statistical significance testing asks whether we can reject that null hypothesis, i.e. say it is *not* true that the GMC case examiners were *not* influenced by a doctor's ethnicity. Scientists reject null hypotheses rather than proving hypotheses true because in science we can only ever rule things out not prove things to be true.

Table 2: Agreement and disagreement between GMC and UCL case examiner decisions by the true and presented ethnicity of the doctor.

Doctor ethnicity			GMC decision		Total	McNemar's test p value
			Did not refer	Referred		
Presented white True white	UCL decision	Did not refer	6	2	8	0.022
		Referred	11	11	22	
		Total	17	13	30	
Presented BME True white	UCL decision	Did not refer	5	3	8	0.035
		Referred	12	10	22	
		Total	17	13	30	
Presented BME True BME	UCL decision	Did not refer	7	4	11	0.549
		Referred	7	12	19	
		Total	14	16	30	
Presented white True BME	UCL decision	Did not refer	7	5	12	0.774
		Referred	7	11	18	
		Total	14	16	30	
Total	UCL decision	Did not refer	25	14	39	0.002
		Referred	37	44	81	
		Total	62	58	120	

Were UCL case examiner decisions influenced by the ethnicity of the doctor?

The UCL case examiner referral rates were very similar regardless of the presented or true ethnicity of the doctor. This suggests that UCL case examiners decisions were not influenced by the ethnicity of the doctor. (Table 3) Although UCL case examiner referral rates for white doctors were about 10% higher than for BME doctors regardless of whether those doctors were presented as being white or BME; a logistic regression showed that UCL case examiner decisions to refer were not statistically significantly related to a doctor's ethnicity ($p=0.437$), to whether the doctor's presented ethnicity was the same or different to the true ethnicity ($p=0.791$), or to the interaction term ³($p=0.864$).

³ This refers to the statistical interaction between a doctor's true ethnicity and their presented ethnicity, for which there are four options (true white presented white, true BME presented BME, true white presented BME, true BME presented white). If whether or not a doctor were referred depended *both* on their true *and* their presented ethnicities, the interaction term would be significant. If only depended on the true *or* their presented ethnicity, the main effects would be significant. In this situation, neither main nor interaction terms are significant suggesting referral did not depend on a doctor's ethnicity, true or presented.

Table 3: UCL case examiner decisions by a doctor's presented ethnicity and true ethnicity.

	Presented white True white	Presented BME True white	Presented white True BME	Presented BME True BME	Total
Number referred (proportion)	22 (0.73)	22 (0.73)	18 (0.60)	19 (0.63)	81 (0.68)
Total number	30	30	30	30	120

PMQ cases

Nine of ten UCL case examiners completed the task, providing data on 108/120 cases. GMC case examiner decisions were missing for 4/120 cases in the material provided for the research. In total therefore we could analyse data from 104 cases (52 original cases and 52 duplicates with the PMQ reversed).

Comparison between UCL and GMC case examiner decisions

As with the ethnicity cases, UCL case examiners significantly referred more cases than GMC case examiners: UCL case examiners referred 76/104 (73.1%) cases, compared to the GMC referral rate of 28/58 (48.2%) (McNemar's test $p < 0.001$).

Figure 2 suggests that UCL case examiners referred more non-UK doctors than the GMC, but they referred a more similar proportion of UK doctors. This is unlikely to be due to UCL case examiner decisions being influenced by the PMQ of the doctor.

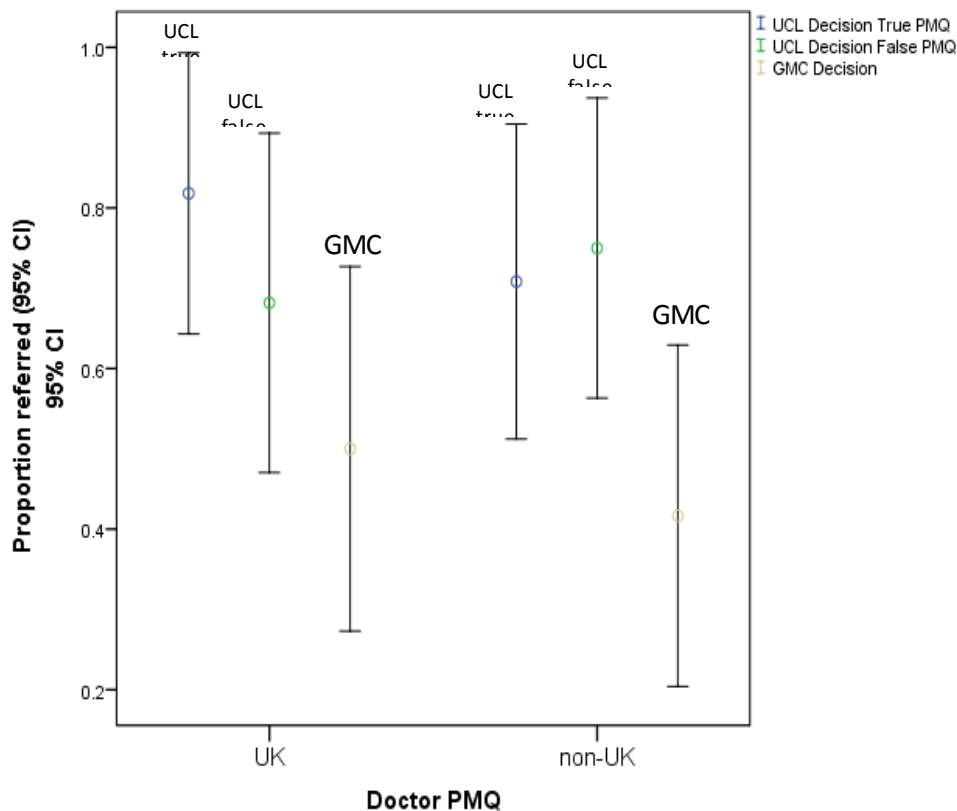


Figure 2: Referral rates for UK and non-UK doctors by PMQ

As before, we conducted McNemar’s test to statistically compare the proportion of cases referred by UCL and the GMC, for UK and non-UK doctors. For non-UK doctors, the McNemar tests are significant, indicating a significant difference between UCL and GMC case examiner decisions (non-UK doctors presented as non-UK doctors $p=0.039$; non-UK doctors presented as UK doctors $p=0.021$).

For UK doctors, the test was not significant (UK presented as UK $p=0.146$; UK presented as non-UK $p=0.227$). The difference is not large, and as with the ethnicity result, it is not statistically significant with a Bonferroni-corrected p value of 0.0125.

Table 4 Agreement and disagreement between GMC and UCL case examiner decisions by the true and presented PMQ of the doctor.

			GMC Decision		Total	McNemar’s test p value
			Did not refer	Referred		
Presented UK True UK	UCL decision	Did not refer	4	3	7	0.146
		Referred	9	9	18	
	Total	13	12	25		
Presented non-UK True UK	UCL decision	Did not refer	4	3	7	0.227
		Referred	8	10	18	
	Total	12	13	25		
Presented non-UK True non-UK	UCL decision	Did not refer	6	2	8	0.039
		Referred	10	9	19	
	Total	16	11	27		
Presented UK True non-UK	UCL decision	Did not refer	5	1	6	0.021
		Referred	9	12	21	
	Total	14	13	27		

Were UCL case examiner decisions influenced by the PMQ of the doctor?

The UCL case examiner referral rates were very similar regardless of the presented or true PMQ of the doctor, suggesting that UCL case examiners decisions were unlikely to be influenced by the PMQ of the doctor. A logistic regression showed that UCL case examiner decisions to refer were not statistically significantly related to a doctor’s PMQ ($p=0.536$), to whether the doctor’s presented PMQ was the same or different to the true PMQ ($p=0.536$), or to the interaction term ($p=0.654$).

Discussion

Based on these results, it is not possible to say with any degree of certainty whether or not the decisions made by the GMC case examiners to refer a case to the IOT were influenced by the ethnicity or PMQ of the doctor.

The results showed GMC case examiners referred slightly fewer white and non-UK doctors compared to UCL case examiners. The UCL case examiners showed no evidence of bias. Interpreting these results is not straightforward. If the GMC case examiners were biased in favour of white doctors (i.e. under-referred them compared to UCL case examiners, whose decisions were seemingly unaffected by the ethnicity of the doctors), it would reasonably be expected that GMC case examiners would also be biased in favour of UK doctors. The pattern of results however suggests that if anything, GMC case examiners under-referred non-UK doctors, the majority of who are BME doctors, compared to UCL case examiners, which is difficult to explain in reference to GMC case examiner bias.

This was a pilot phase of a study to determine if the experimental design could practically answer the research questions. Due to the challenges experienced in the pilot, it was impractical to proceed to a larger study with more cases. The pilot involved a small number of cases which negatively affects the reliability of the findings, and the findings are also subject to a number of important additional caveats:

- It is likely that cases varied in how 'risky' they were – in particular this may be an issue for the white and UK trained doctors who were very few in the larger pool of GMC cases.
- It is likely that the UCL case examiners - and indeed the GMC case examiners - varied in how likely they were to refer a case of a particular level of riskiness (in assessment, we could call this how 'hawkish' or 'dovish' they were). While the random allocation of cases to UCL case examiners aimed to counter the effects of both case riskiness and examiner hawkishness, the relatively small number of cases and examiners means it is likely these problems were not completely overcome. This was particularly so for the PMQ cases where one UCL case examiner did not complete and where GMC decisions were missing for two cases in the material provided for the research. It is unknown how risk-averse the various different GMC case examiners were.
- UCL case examiners were not selected against a set of competencies using the selection criteria used for selecting GMC case examiners and were undertaking the task for the first time, whereas GMC case examiners were much more experienced and work within systems that provide for support on difficult cases. Perhaps partly as a result of their inexperience, UCL case examiners overall referred many more doctors compared to GMC case examiners. With the relatively small number of cases involved, this over-referring could affect the overall findings.

Several changes were made to case examiner training and guidance over the time frame of cases being analysed, and thus UCL case examiners were using different criteria from the GMC case examiners to assess the older cases.

Qualitative phase

Introduction

The purpose of this part of the project was to explore case examiners' views about guidance and training they received and how they make decisions to refer (or not) cases to IOT. By doing so we were looking for any evidence of bias.

The objectives were to:

- describe how GMC guidance is used to arrive at a decision.
- describe views about the quality of guidance and training received.
- identify challenges of making the decision to refer or not refer doctors to an IOT.
- identify factors that influenced decisions to refer or not refer doctors to an IOT.

Background

One of the roles of case examiner, is to decide if a case needs referring to an IOT hearing, who then consider suspending or restricting a doctor's practice while the investigation continues. This qualitative study explored the cognitive processes that influenced case examiner decisions to refer (or not) to an IOT hearing. We used the decision sciences to inform the design and analysis of this study.

The decision sciences are the field of study about human judgement and decision making; drawing on theories and evidence from the behavioural sciences, economics and mathematics to describe how all individuals make judgements and decisions.

We know that non-white and overseas doctors are more likely to be referred to the GMC's Fitness to Practise proceedings, in particular by public bodies such as healthcare providers and the police, and to receive sanctions on their registration. This is likely to be due to a multitude of factors, this study is one of a number of independent studies to check that decision making is free from bias.

Case files consist of information that has been collated in relation to the allegation some of which can include letters from the police, correspondence between a doctor's colleagues or the trust, patient's clinical notes or reports from proceedings that have already taken place. The volume of information can vary substantially between cases, ranging from between a few pages to over 2000 pages. A case examiner's task is to review these documents on the GMC's electronic case management system and judge whether temporary restrictions on a doctor's registration may be necessary.

For this research study, thousands of documents had to be copied from the electronic database into case files and, in the case files prepared for this research study, repetition of documents was not uncommon.

Information processing and selective attention

Actively making sense of information is known as Information Processing. The information processing approach has driven much of the research on judgement and choice (1). It is still one of the dominant frameworks to describe how all individuals make judgements and decisions from the information in the world (1, 2)

In this approach, the brain has a given infrastructure that makes sense of and stores information from the outside world (3) but the brain cannot consciously process all of the complex information. Consequently people are highly selective about the information they attend to as the brain has limited capacity for conscious attention. (4, 5) This happens at the unconscious level so most of the time people are unaware of how they process information.

These general principles of information processing and selective attention apply to the way GMC case examiners review cases. The decisions they make are based on the information they judged to be most relevant i.e. a sub-set of information that they consciously attended to and retained in memory rather than all of the information available. This principle would be true of any individual who was asked to review a GMC case and make a decision. Therefore part of the case examiner's role is to be skilled in judging relevant from less relevant information presented to form the basis of their decision.

How people make decisions: a dual processing model

There are two main ways people process information in order to make judgements and decisions

- 1) heuristic processing
- 2) systematic processing strategies

The way people use these two systems are considered standard features of human decision making and are represented in dual processing models of information processing (1, 5, 6, 7). Heuristic thinking (also known as system one) is characterised by a fast and frugal process via the use of mental shortcuts or 'rules of thumb'. Judgements and decisions are made quickly and easily because little, if any, conscious thought is involved. It is largely automatic, for example, people readily conclude that a baby dressed in pink is a girl, or they may buy a camera that is recommended by a trusted friend rather than spend time searching for the latest model or best value for money.

Systematic processing strategies (also known as system two) involve a slower and deliberative thought process. It requires conscious effort to scrutinise information and evaluate the advantages and disadvantages of the consequences of several options. For instance, a patient experiencing a dilemma about whether to have risky surgery or not, may reach their decision systematically. However, it is not correct to think that all high stakes decisions are made systematically, and the two processing strategies should not be thought of as a dichotomy where one is better than the other.

Heuristics are considered to be 'smart' and effective strategies that mostly lead to good judgements and decisions (8, 9). Their use should not be thought of as inherently problematic or as a failure of

people's cognitive limitations (9). Furthermore, one study found that clinical decisions in one particular context were more robust when made on principles of heuristics as opposed to principles of optimisation (10). It is not well understood how people switch between a heuristic and systematic processing strategy, but some influential factors are likely to include the quality of information, its accuracy and vividness, personal relevance of the information and comprehension ability of the decision maker (5).

Case examiners, like all individuals, will engage in some level of heuristic and systematic processing when deciding whether to refer cases to an IOT or not.

How bias occurs in the decision process

Bias can be defined as a systematic error in thinking (11) which has been found to occur in people's thinking in predictable ways. That is, people make the same types of incorrect judgements when a particular heuristic strategy is used (11,12,13). These cognitive errors are referred to as biases. Most of this work is attributed to an extensive research programme by Kahneman and Tversky in the 1970's, where many types of heuristics and their characteristics were first identified. The dominant view of how bias occurs in a decision process is due to an over reliance or a faulty use of a rule of thumb (i.e. heuristic). While this view (that bias is linked to inappropriate use of heuristics) has since been challenged (9, 10), it is accepted that people's thinking is susceptible to bias even if it is unclear on why it happens.

Biased information processing can potentially occur at any stage of the IOT reviewing process. Case examiner decisions are based on a review of qualitative documents, without a mark that denotes a pass or fail boundary. Instead case examiners must piece a narrative together and make a number of judgements before summarising their opinion into a decision of what should happen next. It is unclear as to whether this type of exercise allows greater room for personal biases and idiosyncrasies than if data was presented quantitatively or marked by machines. There is likely to be differences between the stringency of case examiners as well. Evidence demonstrates that individual assessors in any field differ in their marking behaviour, referred to as the 'hawk-dove effect' with strict assessors known as 'hawks' and lenient assessors known as 'doves' (14, 15). There is currently no consensus on how to ensure decisions made via qualitative assessments are rigorous, however frameworks that guide the collection and interpretation of qualitative data have been suggested to support the defensibility of these types of decisions (16).

The GMC employ a range of tools and good practices with regards to their organisational decision making processes (17). Their case examiners are well trained and highly skilled in their decision making role, but nevertheless are susceptible to the same pitfalls in their reasoning as everyone else. Heuristics are recognised as a universal feature of human cognition that GMC case examiners will draw upon to make their decisions, alongside more systematic methods.

The extensive work by Kahneman and Tversky led to identifying a large number of heuristics that have been identified that may lead to bias. Heuristics can be grouped into three main types that include

- 1) anchoring
- 2) availability
- 3) representativeness

Table 5 summarises each type of heuristic and how it can result in making a biased decision.

Heuristic type	Description	Possible consequences
Anchoring	Tendency to fixate on specific aspects of information early on.	Judgement based on first impressions and never revised in light of later information. Jumping to conclusions. Premature closure of thinking.
Availability	Tendency to judge things/situations as more frequent if easy to remember. Recently encountered information remains fresh in memory. Recency effect.	Judgement based on evidence that's easiest to remember. Novices more easily influenced by availability heuristic than experienced decision makers.
Representativeness	Tendency to look for confirming evidence to support the hypothesis, rather than look for disconfirming evidence to refute. Referring to prototypes/mental templates that represent the typical and overlook less typical.	Judgement based on stereotypes/prejudice. Preserving a weak hypothesis.

Table 6: Commonly used heuristics by all individuals (12)

Methods

Participants

At the start of this project in 2015 there were 18 case examiners working for the GMC. All case examiners received an invitation e-mail about the study sent by a GMC Research Manager. (Appendix 5) It was anticipated that interviewing 50% of the sample would generate a sufficient amount of data to answer the research questions, and so we recruited the first nine people who expressed an interest to participate with a view to further recruitment at a later date if necessary. Our preliminary data analysis suggested themes were becoming repetitive after interviewing people, so no further interviews were conducted among GMC case examiners.

In a separate experimental study (also described earlier in this report), staff members working in UCL Medical School were trained as mock case examiners to review a random sample of the GMC's cases and decide whether to refer each one to IOT or not. We decided to interview all UCL experimental participants about their views and experiences of making these decisions as novices, with a view to

understanding how their decision making processes compared with those of the experts (GMC case examiners).

Procedure

Those who agreed to volunteer contacted LM directly who arranged a mutually convenient time for an interview to take place. Interviews were conducted on a 1:1 basis, face to face with UCL staff and mostly via telephone with GMC staff. The interview schedule (Appendix 6 and 7) consisted of preliminary demographic questions followed by describing the process of making decisions, how they implement the GMC's guidance as well as specific examples that demonstrate their reasoning around easy and difficult cases.

Interviews lasted between 30-60 minutes and were audio recorded and transcribed verbatim by a professional transcriber. For UCL participants, prompts were used in interviews as a substantial time had elapsed from the decisions they had made in phase one and their interview.

Data analysis

Using NVivo software version 11, a thematic analysis was conducted to categorise the transcript data and summarise into main themes. Two researchers coded transcripts to enhance reliability of the analysis.

Results

Sample characteristics

The demographic characteristics of the 9 GMC case examiners that were interviewed is summarised in Table 7. One was involved in delivering training for new case examiners and two referred to themselves as senior case examiners. To become a case examiner requires a highly competitive selection process. The majority had come through this route, with the exception of one who had previous roles within the GMC before taking on the role of case examiner. Consequently they are a highly skilled group of individuals usually with a substantial amount of prior decision making expertise either from their clinical work or legal background.

Table 6 demographics of GMC interviewed case examiners

Sex	Type of Case Examiner	Ethnicity	PMQ region (if applicable)	Years of experience in role
1 Male	Clinical	White	UK	10
2 Male	Clinical	White	UK	3
3 Male	Non-clinical	Black British	n/a	3
4 Male	Clinical	White	EEA	10
5 Female	Clinical	White	UK	3.5
6 Male	Non-clinical	White	n/a	3
7 Male	Clinical	White	UK	11
8 Female	Non-clinical	White	n/a	<1
9 Female	Clinical	White	UK	3

Twelve UCL members of staff were interviewed. Seven people reviewed 20 cases each and six people reviewed 10 cases each. Six people were medically qualified. A comparison with GMC case examiners is summarised in Table 8.

Table 8: Comparison of GMC with UCL case examiners

	Sex	Ethnicity	Case examiner type
GMC (n = 9)	M = 6 F = 3	White = 8 BME = 1	Clinical = 6 Non-clinical = 3
UCL (n = 12)	M = 2 F = 10	White = 8 BME = 4	Clinical = 4 Non-clinical = 8

The thematic analysis generated five main themes, and each theme was associated with a number of sub-categories. The main themes included:

- 1) GMC practice
- 2) Reasoning processes
- 3) Fairness and objectivity
- 4) Challenges of role
- 5) Good decision making

Theme 1: GMC practice

Interviewees expressed their views on practices within the GMC with regards to their decision making role. Both GMC and UCL staff highlighted the positive and negative aspects of the case examiner training as well as the guidance document produced by the GMC to aid their decisions.

1.1 Case examiner training

Prior to the year 2014, new case examiners were provided with the GMC guidance document, 'at desk training' followed by a limited period of mentoring. Since 2014, a formalised two day training course has been delivered. Day one introduces the Fitness to Practise investigation process, role of case examiner, implementing GMC guidance and fairness in decision making. Day two is more focused on the practicalities of making decisions, with case based discussions and learning how to write up their decisions. This two day course is combined with a period of formal mentoring consisting of detailed scrutiny and feedback on decisions written by new case examiners.

Overall, interviewees were satisfied with the training they had received and one described it as "quite comprehensive". One interviewee suggested that there was only so much to be learned through formal training after which practice is required.

"I think it is satisfactory but it's like anything, you know the more you do it the better you should get at it. And it's not the kind of thing you can learn in isolation, you have to be able to meet with other case examiners when you're learning the job to be able to discuss it through." (GMC white male, clinical)

There was a period of supervised case examining where new GMC case examiners would not be left to make decisions themselves without their mentor's (an experienced GMC case examiner) input. It was described as an apprenticeship and a probationary period that was lengthy until their mentor deemed it was appropriate for them to make decisions independently.

UCL staff who were trained as case examiners experienced the two day training delivered by a senior GMC case examiner. They did not receive mentoring and went straight into reviewing cases and making decisions on their own. UCL staff found the practical elements of the training most memorable and helpful.

1.2 Guidance document on making IOT decisions

The GMC have produced a document for case examiners that offers guidance on when to refer a case to an IOT. This includes information on the role of IOT, factors that should be considered in a case and instructions on how to record their decision. Interviewees thought this was a useful piece of guidance that illustrated the spectrum of cases that indicate when a referral is necessary.

“I think the guidance does two things... it shows you the spectrum, it shows you what the lower end of the spectrum and is unlikely to be referred ... although obviously you have to take into account the circumstances of the particular case ... and it shows you the top end of the spectrum where it will be referred.” (GMC white female, non-clinical)

Case examiners used the guidance document as a framework for their thinking and writing up of their decisions. They are trained to pay particular attention to section 6 of the document that represents the “key component” of factors when considering referral or not. These factors relate to the significance of the risk the allegations present to patient safety, public confidence in the medical profession and/or doctor’s own welfare.

“It talks about the factors that are relevant when making the decision; it talks about quantifying the risk to patients. So you know talk about the risk to patients first of all, talk about the risk to public confidence, and then talk about the welfare of the doctor, and break the decision down into those parts.” (GMC white male, clinical)

“Risk to members of the public, and risk to public confidence. If we’ve got a doctor who’s been arrested for a really serious ... and there’s a really serious allegation, you know the public would expect the GMC to do something about it.” (GMC white female, clinical)

When writing up a justification for their decision, case examiners would select the most relevant aspects of the guidance document to quote from. This was to make sure their decisions were clearly grounded in the GMC’s guidance document.

“It’s very useful sometimes to absorb the information in a particular case, and then actually look at the guidance and to structure the decision that we write in the format that you might expect based on the guidance.” (GMC white male, clinical)

“And then I would read that and then try and use that terminology in the actual write up, so that I kind of kept to the wording and didn’t veer off into opinion.” (UCL white female, non-clinical)

The guidance document was viewed to have some limitations. Case examiners wanted clear examples to be included in the guidance that would help them resolve challenging cases, eg when a case does/does not clearly meet the referral criteria or how to prioritise certain pieces of evidence over others.

“Where a referral doesn’t fit becomes more an area you have to exercise judgement based on precedent both in what you’ve done previously and in discussion with colleagues.” (GMC white female, clinical)

“It’s useful in terms of thinking about the framework for the structure of the decisions and the reasons that you’d have to consider, but it isn’t that clear in terms of you know how you weigh up potentially sort of conflicting pieces of evidence.” (GMC white male, clinical)

A few people felt that the guidance document could be reduced in length, though most participants did not comment on the length.

“It’s quite wordy. I certainly use it quite a lot, it’s quite wordy. I think most of it might be redundant from my point of view...most of the stuff in it is not relevant to any given case.”
(GMC white male, clinical)

“It’s quite long, and actually what you need is that, the statements on if you decide yes, which ones. That’s kind of the meat of the entire document, really.” (UCL white female, clinical)

Theme 2: Reasoning processes

The way people reviewed cases and arrived at a decision differed depending on the nature of allegations and the complexity of information provided. There were times when they described how easy and straightforward the reviewing process was. In these cases, there was evidence that heuristics were used to simplify the thought process which allowed them to reach a clear decision with confidence. Other times there was much greater deliberation involved and a feeling of being torn between referring and not referring. Such cases, their decisions were probably made systematically via a process of weighing up the risk of referring vs not referring to an IOT.

2.1 A heuristic approach

There were plenty of examples when case examiners made effective use of heuristics (rules of thumb). These can be thought of as ‘smart and effective strategies’. Some rules of thumb were specifically rooted in GMC policy, while others were more generic that could inform the decisions made by anyone (and not just case examiners).

At the organisation level, the guidance document as well as a standard default position helped to simplify the decision process and ensure that their decisions reflected GMC policy. At the forefront of their minds was section 6 of the guidance document. Case examiners would apply this section as a criteria to judge whether each case presented a significant risk to patients, public confidence or doctor’s welfare. For cases that obviously met one or more criteria, the decision was made in a quicker and more confident manner than when it was less clear if the criteria for referral was met.

“Is there a danger to patients, is there a danger to the GMC and the profession, is it in the doctor’s best interest to consider getting them more supervised one way or another. And if you stick to that it’s very straightforward - 99.99% of the time it’s very straightforward.”
(GMC white male, clinical)

Another smart heuristic used was to refer to a default position that was encouraged by the GMC. Case examiners had been trained to err on the side of making a referral if they felt conflicted about the right course of action in order to give the IOT tribunal the opportunity to review the available

information and assess the risk. This was a simple and effective way to regain clarity, again originating from guidance at the organisational level.

“I referred that one, again because there’s a tendency to default towards referring when it’s very difficult.” (GMC white male, clinical).

“Because if you’re that close to referring it, from a personal perspective, then you should refer it.” (GMC male, non-clinical)

“And normally, given the fact that I got the impression that if you were in doubt, refer. That was the general feeling I’d gotten. Nobody said it, but that was the feeling I got from the GMC.” (UCL white female, non-clinical)

There were examples when case examiners made use of universal heuristics i.e. those that affect the judgements and decisions of all individuals. For instance, experience had taught case examiners that patterns of similar features exist between cases. There were occasions when their decision was informed by how they dealt with a similar case in the past. In such cases, a representativeness heuristic may have been used to reach a decision.

“Cases normally fall into some very clear-cut patterns, and you know the usual pattern is a doctor who’s alleged to have done some serious clinical misdemeanour. And that you know ... it’s a pattern, so you only really need to go to small parts of the guidance for that.” (GMC white male, clinical)

A danger with using the representative heuristic is the tendency that all individuals have to search for evidence that proves their hypothesis rather than disproves. This can mean that decisions are based on inaccurate stereotypes. The way case examiners can avoid this is by maintaining an open mind throughout the reviewing process, an example of good practice is below.

“I would have to say the thing that makes my job interesting is there’s no stereotype – every case needs to be considered on its own merits, so I will never open a case and think ‘barn door - IOT’ ‘barn door – not IOT.’” (GMC white female, clinical)

Case examiners had been trained to distinguish credible from less credible evidence to inform their decision. Where cases included expert reports or documents from the police, these were described as potentially credible pieces of evidence that case examiners would actively hasten to find and review among files. Such evidence may have encouraged the use of an anchoring heuristic at some point in the reviewing stage.

“You just try and pick out the areas of the documents that will underpin what you need. So yeah things like a trust investigation or if somebody says ... if the police say we’ve arrested this person on suspicion of sexual assault and we’ve attached a witness statement from the complainant, you’ll go and look at the witness statement.” (GMC white male, non-clinical)

2:2 Systematic strategies

Some cases were described as more complicated and challenging than others. These were cases that usually consisted of a high volume of information and where it was less obvious as to whether allegations met the criteria in the guidance for referral or not. On these occasions, systematic approaches probably informed their decisions to refer or not rather than heuristics. Case examiners would speak about how much longer it would take to make a decision because they had to deliberate over the risks of referring versus not referring. The examples below illustrate how conflicted they could feel about the decision process and how they experienced assessing potential risks.

“If you were to be safe you could send everything couldn't you? If you were to be ultra-safe then then the case examiner would say it's getting that balance between...it's just weighing up is it so serious that we have to send this doctor to potentially have his or her registration restricted. And like I said, it can be a number of factors that we weigh up.” (GMC white male, non-clinical)

“You're making a decision based on what might happen, what the evidence suggests did happen, but it's not been tested, nothing's been proved. So that's a pretty draconian power, so you should really only refer it if you think it's necessary. Conversely, if you don't refer you've got to pretty certain that that risk isn't there... so you have to weigh the risk and be satisfied that it's necessary, and the only way I think you can do that is if you read everything and then consider it in that structured way.” (GMC white male, non-clinical)

“The one where the allegation was a sexual assault on a mentally ill relative ... it was very difficult as I remember it because ... you know it could have been anything from completely made up you know psychotic delusions, to someone taking advantage of a vulnerable child. And obviously that's a huge spectrum ... and I referred that. I remember thinking well you know if that is literally made up ramblings of a psychotic person then that doctor's family circumstances are already absolutely disastrous, and the referral to an IOP would not make things any better.” (GMC white male, clinical)

“And I found that quite difficult, because it kind of felt like all the procedures are in place, the guy's getting help and getting better, and yet it's come through for a decision. I found that challenging. So I would deliberate on those quite a lot.” (UCL BME male, clinical)

Theme 3: Fairness and objectivity

Case examiners were aware that certain groups of doctors were over represented in fitness to practise investigations, but a firm view that specific demographic characteristics of a doctor (such as sex, ethnicity, country of qualification) did not negatively influence their decisions to refer to IOT or not. All had received training in unconscious bias and how such factors can influence their decisions. Case examiners have experienced a highly competitive selection process and were recruited due to some prior experience of making high stakes professional judgements. Some implied that their prior experience enabled them to review cases in a fair and objective manner.

“You know we are recruited because of our ability to make impartial balanced judgements. So I suppose you could argue you can’t train that into somebody... if people are inherently going to be biased towards certain characteristics I’m not sure you can train that out of somebody.” (GMC white female, clinical)

“I think all of us are senior enough and are experienced enough to know if we are being overtly prejudiced” (GMC, white male, non-clinical)

3.1 Efforts be fair

Case examiners emphasised that they made a concerted effort to maintain fairness and objectivity during their decision making process of each case. They had strategies to maintain a fair process which included focusing on the evidence (or lack of) to justify their written decision or acknowledging when their objectivity may be compromised.

“I like to think that who the doctor is doesn’t matter to me in making my decision, it’s about what are the nature of the allegations, what is the nature of the evidence, and on those two, what is the nature of the risk, therefore what is the appropriate outcome based on what we’ve done in other similar cases.” (GMC white female, clinical)

“So one of the ways I try and get round that is to make sure that when I re-read that decision everything is linked to something pretty ... some strong evidence. And to me there are certain things where I just don’t look at them because I don’t care, so things like where a doctor got his or her medical qualification makes no difference to me. That has no bearing on my final decision so I don’t look ... things like the doctor’s age – I don’t look.” (GMC white male, non-clinical)

“If I felt that I couldn’t deal with a case (emotionally) I would get it reassigned.” (GMC, white male, clinical)

3.2 Preconceived views

Case examiners had a preconceived view about some doctors over others and some cases did evoke emotional reactions that may or may not have affected the decisions they reached. There was a view that locum doctors about whom serious concerns are raised, present more of a risk to patient safety given the nature of their work conditions i.e. moving frequently and potentially being poorly supervised. Consequently, cases that concern locum doctors may be more likely to be referred to an IOT than others. If so, case examiners were keen to emphasise their decisions concerning locum doctors about whom serious concerns are raised, were not a reflection of unconscious bias. Rather they were acting in accordance with their training and organisation’s legislation in focussing on risk.

“I did mention earlier the issue about I feel more reassured when a doctor is in stable governanced and scrutinised employment, the doctor’s in training for example with the deanery, their NHS trust ... compared to someone who was a perennial locum perhaps just working short time. And I think that does ... my personal view is that does impact on risk because if you’re not long with an employer for them to get to know you, that gives the

doctor a lot more opportunity if they're not very good or not very well behaved." (GMC, white female clinical)

When asked whether they had noticed if any groups of doctors are more referred to IOT than others, one response was:

"I think it's more difficult for people who work as locums, ... I never really thought about this before you asked me. But um ... considering as I said your question, when you've got somebody who doesn't work in an organisation as such, so they haven't got a stable job and they tend to move around a lot, then I can see why the GMC might sort of you know try to be more cautious with them." (GMC white male, clinical)

There was a perception among some UCL participants that the training they received highlighted locum doctors about whom serious concerns are raised, present a greater risk to patient safety. Whilst this was not documented in training slides or guidance document, this message may have come through during the discussions of sample cases.

"From our training we were told always to look out for whether doctors were locum... I feel it was implied in the training – it is more likely that you would refer a locum just because of, because they have the power to sort of go anywhere." (UCL BME female, non-clinical)

On other occasions, case examiners did recognise when they felt a certain way about a case and were open about labelling these as their "personal biases" that they had to be mindful of. Again, whether the assumptions made in the examples below impacted on the decision cannot be proven, but it is possible and something that could be addressed in case examiner training.

"And the cases that annoy me are where colleagues obviously should have known better. And for some reason they didn't do what they would normally do - and they knew they weren't doing it the way they should do it, and they did it wrongly for no good reason." (GMC white male, clinical)

"Am I more lenient on the very young trainees? Possibly not lenient because you just cannot let somebody off if they've done something truly dreadful. So we get a referral of a doctor in their very first year and they've just not actually clinically been very good, I'm probably going to be slightly more lenient than say to a consultant who's not very good because he's had years of training ... do you see what I mean? I think those feelings probably do affect me." (GMC white female, clinical)

"The motivation of the complainants sometimes causes me difficulty. I had a case where the complainant made an allegation that she had to go to a walk-in clinic or an out of hours service or something like that, and she made statements that weren't relevant to the doctor - so that the place was spooky and it was deserted. And I do find some of those quite difficult. And I guess you know everybody has some kind of bias, and my bias is I'm not a victim, I don't play the role of a victim, I'm pretty good at standing up for myself. So I find myself needing to be very careful when I'm dealing with people who consider themselves to be victims. When I say 'victim' I mean in the sense that it is somebody else's responsibility, and they take no responsibility for themselves. So that's my bias." (GMC white female, clinical)

Theme 4: Challenges of role

The reviewing and decision making process was mostly described as a difficult and demanding process. Some of the factors that made the role of case examiner challenging included the organisation of cases, quality of evidence and how to manage feelings of uncertainty.

4.1 Organisation and amount of information

The format of case files differed between those that GMC staff reviewed compared with cases UCL staff reviewed. GMC staff accessed case files directly from an electronic database. UCL staff accessed electronic PDF versions of case files that were downloaded and put together specifically for the research by administrators employed by the GMC for this project. The latter format meant UCL case examiners were reviewing files that were far less streamlined and included lots of duplicated documents. This was the main factor that made decision making difficult for UCL staff.

“Just such a shame that the documentation is never put together in a logical format...it would have been a lot more straightforward to have a timeline at the beginning. Then maybe people wouldn’t read the documentation. But at the same time, that would have made the decision-making process a bit easier to get around your head.” (UCL white female, non-clinical)

Nevertheless, the actual GMC case examiners did also suggest that the volume and sequencing of documents was sometimes a hindrance in their ability to review cases efficiently. A high volume of documents made it more difficult for case examiners to select the most relevant information needed to underpin their decision.

“Other times you can get a very short paragraph saying I’ve got real concerns about this doctor – then they just send you 100 pages of documents and you have to go through and pick out ...it makes it harder because they’ve kind of just given you a dump of stuff because maybe they’re too busy, and then you have to figure out what’s relevant and what isn’t.” (GMC white male, non-clinical)

“It may be that the timeline is very complicated...it may be that the dates aren’t quite right. So occasionally you do have to take notes and try and sort things out to try and get the actual proper narrative straight in your head.” (GMC white male, clinical)

“Some of the cases involving masses of documents with different opinions about the doctor’s actions – those can be equally challenging just because of the sheer volume.” (GMC white male, clinical)

Other times it was the lack of information at that point in an investigation which made the process hard.

“The difficult ones are where you have two lines, and the allegation is so serious that it can’t merely be dismissed. And you know an allegation of a serious criminal offence is one, an allegation of a serious error, a clinical error, is another. Where you could have a few lines, no investigation has been carried out, but you may have a patient who’s sadly died, or you may have a serious criminal allegation. You can’t just say ‘I’ll tell you what, I’ll make that decision in 3 months when you’ve come back with more information’ – you have to make it. So the more information you have, more often than not the easier the decision is.” (GMC white male, clinical)

“I don’t think you can say one case is harder than another, it’s just that ... and it’s not that a case is harder because there’s more content, or a case is easier because there’s less content – sometimes if there’s less it’s actually harder to be sure that you’re making the right decision because you’ve not got enough information.” (GMC white male, clinical)

4.2 Quality of evidence

The quality rather than quantity of information case examiners had to review was significant in forming judgements about the case. Everyone spoke about information in terms of evidence and emphasised that their decisions had to be evidence based.

It was clear that some types of information were judged to be more credible evidence than others, when used to justify their decisions. Generally, official documents from the police, NHS trust and GMC were more influential than emails between various parties and character witness statements in support of the doctor (although this did depend on context of the case).

“Always consider we need to have enough information for the panel to make an informed decision on the matter. So if we don’t have enough information then we can’t make a referral. We don’t go on fishing expeditions as such.” (GMC white male, clinical)

“You know you have to kind of come down on the side of well is this credible or is this just somebody having a pop because you know there’s been a breakdown in the relationship and they’re seeking revenge if you like.” (GMC white female, non-clinical)

“There are obviously some documents that are more important than others, so if a trust has sent us an investigation ... if a trust has thought it’s serious enough that they’ve actually conducted an internal investigation, they’ll send that to us, and that will be very helpful. Because you’ll read it and you’ll figure out, and some of them will include things like witness statements. And again it gives you that confidence to think this isn’t just you know this doctor ... there is a concern.” (GMC white male, non-clinical)

4.3 Managing uncertainty

Another major challenge was making a decision whilst feeling uncertain. This is not surprising given that case examiners make decisions to manage a potential risk. According to the decision sciences literature, risk and uncertainty go hand in hand, and these are referred to as complex decisions.

There was uncertainty around understanding the content in documents, the trueness of allegations and how severe the risk to patient, public, and/or doctor was.

“If you felt like it was more a dispute between colleagues.... So if I felt like it was very much a personality clash, I didn’t take that as seriously.” (UCL white female, non-clinical)

“The ones that I do slightly struggle with is when little old ladies write in and say ‘the doctor had his hands all over me’. And then you read she might have a cardiovascular problem...so the complaint might be looking as though there were serious sexual allegations concerning the doctor, and then you think what motivation might the doctor have, other than trying to examine a lady with a cardiovascular problem.” (GMC white female, clinical)

Case examiners would manage and resolve their uncertainty by discussing with other case examiners or taking a break and revisiting the case later. It was viewed as good practice, and part of GMC policy, for case examiners to consult one another to manage and resolve any uncertainty around a particular case. A case examiner may speak to any available colleague at the time, or approach someone they perceived to be more experienced than themselves. They also would discuss challenging cases in a group setting at case based discussing meetings.

“Occasionally you might think am I just being a bit too hard on this bloke for some reason, or am I not being hard enough, am I not taking it seriously enough. And you know you can argue it either way maybe. So you have a chat, what do you think, we’ll have a little vote on it or something, which way do you think we should go on this. And sometimes people can remind you of things that you’ve forgotten.” (GMC white male, clinical)

“I will run those past usually a lay case examiner because we’ve got the two perspectives here, we’ve got the lay case examiners who will bring if you like the non-medical perspective, what someone who wasn’t clinically trained might consider normal –the man in street almost type situation ...” (GMC white female, clinical)

“There will be occasions where I will read the information and I will walk away from the decision for an hour because I’m not certain.” (GMC white male, non-clinical)

Theme 5: Good decision making

Decisions were not necessarily right or wrong, yet there was a clear sense that the process was important. The outcome of their choice or whether others agreed with their choice was unimportant. A good decision was one where the reasons were transparent, and this enabled them to maintain a confidence in their decisions.

5.1. Transparent reasoning

It was crucial that their decisions were evidence based and justified clearly to others in the write up so others could understand how they reached a conclusion.

The consequences of a referral to IOT for the doctor's health and career were an important consideration for all case examiners. Case examiners made an effort to be as fair and objective as possible and would only make a referral if they were sure they could base their decision on evidence and/or the guidance document.

"At the end of my decision I'd read them through and I'd make sure that they make sense in the form of you know A leads to B and therefore B leads to C, and therefore C leads to D and so on ... without any loose ends really." (GMC white male, clinical)

"You can't just say 'referral is indicated' - you have to explain in detail why...it's all about the reasoning. So a doctor has to understand why we're sending him or her to a panel. They might not agree, but they have to understand why we've reached that decision. They may well go to panel and say the information the case examiner's got wasn't accurate. That's fine, doesn't mean we made the wrong call. But they need to be able to understand when they read it – this is why I've been sent." (GMC white male, non-clinical)

"They look to see if our reasoning was clear and full enough. So basically they're not there to say oh you should or you shouldn't have done this, the audit looks at whether we've addressed the allegations, referenced the evidence, and whether our reasoning is ... there's enough reasoning for somebody to understand how we got to our decision." (GMC white male, clinical)

5.2 Confidence

A transparent reasoning process allowed GMC case examiners to maintain confidence in their own decisions, even in cases that had made them feel particularly conflicted. UCL case examiners appeared less confident than actual case examiners, probably due to the novelty of the role. The clinicians among UCL staff spoke with more confidence and certainty about the decisions they had made than non-clinicians, probably due to their experience of making tough clinical decisions.

"When you get an IOP [now called IOT] that's yours and you have to make the call ... and that's part of the reason for you know the intensive training and the intensive recruitment is that we get people as case examiners that have that confidence and that can ... the key for me with all of this is that the reasoning has to be ... it's all about the reasoning." (GMC white male, clinical)

"It's not to say that I don't recognise the gravity of the decisions that we're making, but I think you can't necessarily be effective in continuing to make different decisions about different cases if you're ruminating about something that you've looked at previously." (GMC white male, clinical)

"I think it plays on your mind while you're doing it, but I think you have to make a decision and I think by having those three questions in your mind, and by informing yourself as much as possible with the evidence that's presented and trying to put your prejudice aside, sticking to the evidence, sticking to the three questions, you have to make a decision. And I

think it's playing on your mind while you're evaluating it. But not when you've reached your decision." (UCL white female, clinical)

"Because the problem with this job is you have to be confident in the decisions that you're making – everyone's going to make a wrong decision on occasion, but you've got to be confident that you're making a right decision most of the times in the right way, and that actually because most decisions are seen by two case examiners, that you can actually pick up when another case examiner might have dropped a clanger and not realised it." (GMC white male, clinical)

Discussion

Nine GMC case examiners and 12 UCL case examiners took part in 1:1 interviews between January 2016 and August 2017. A thematic analysis of the qualitative data revealed five themes and related sub-themes, as summarised in Table 9.

Table 8: Summarised themes

Theme	Sub-theme
1. GMC practice	1.1 Case examiner training 1.2 Guidance document for making IOT decisions
2. Reasoning approaches	2.1 Heuristic 2.2 Systematic
3. Fairness and objectivity	3.1 Efforts to be fair 3.2 Personal assumptions towards case 3.3 Personal assumptions towards case examiners
4. Challenges of role	4.1 Organisation of cases 4.2 Quality of evidence 4.3 Managing uncertainty
5. Good decision making	5.1 Transparency 5.2 Confidence

The findings in relation to the four research questions were:

1. How is GMC guidance document on referrals to IOT used to make decisions?

- Key component was section 6 outlining the three criteria of risk to patient safety, public confidence in GMC and doctor's interest.
- Three criteria used as a smart and effective strategies to simplify decision making process, enabled them to categorise cases in line with GMC policy.
- Frames the way they assess the risk of referring/not referring a case to IOT.
- Provides structure for the way they record their decisions in writing, terminology borrowed.
- Decisions were described as simple when cases clearly met one of the criteria for referral but challenging if less clear whether criteria are met.
- Specific feedback from case examiners
 - Condense into shorter document
 - Clarify with examples when threshold for risk is met, not met and on the borderline

2. Views about quality of case examiner training

- Largely positive impressions from GMC staff to the changes made to formalise training.
- Sessions with senior case examiner described as comprehensive and engaging.
- GMC staff found mentoring with senior case examiner as most essential part of training.
- GMC staff believe structured sessions have limited use and decision making is best learned on job and improves with experience.
- Theoretical content of their case examiner training on investigation process, IOT panels and legal aspects could be reduced.
- Specific feedback from case examiners
 - UCL felt more of structured training session should focus on practical elements of making and writing up decisions.
 - GMC welcome regular feedback on their decisions e.g. are they written transparently, how do their ratios of referral to non-referrals compare with other case examiners, outcome at IOT.

3. Challenges of making the decision to refer or not refer doctors to an interim orders tribunal

- Three main challenges identified by GMC and UCL staff i) organisation of cases ii) judging credibility of evidence iii) how to resolve uncertainty.
- Case documents were not always organised as well as they could be to facilitate an efficient decision making process.
 - Case with high volume of documents could be streamlined better. An absence of chronological order and duplication of documents were main hindrances.
- An expectation that their decisions must be underpinned by evidence, but guidance on what constitutes 'credible' vs less credible evidence was lacking.
- Feelings of uncertainty were linked to greater speculation about trueness of allegations, content presented and severity of risk to patient safety, public confidence and/or doctor.
- GMC staff resolved their uncertainty predominately by conferring with other case examiners, either via someone close to hand or seeking out a particular person
- Evidence that conferring with others sometimes changed their thinking and final decision, perceived to be a positive.
- Other approaches to resolving uncertainty were to err on the side of caution and refer to enable the IOT tribunal to review the available information and assess risk, taking a break and revisiting a case, both of which may or may not lead to sub-optimal decision making.

4. How case examiners make decisions to refer (or not) to IOT and what influenced their decisions?

- Like all individuals, case examiners will have made some decisions very quickly and automatically by using heuristics e.g. guidance document or referring out of caution. Other times they would take longer to deliberate systematically over whether the case met the threshold for referral or not.

- It is likely that they will have made heuristic and systematic judgements within the same case review.
- Nature of allegations and complexity of case affected whether they approached a case heuristically or systematically.
- GMC staff use heuristics that were grounded in GMC guidance and were appropriate
- GMC staff seemed to use representative heuristic (pattern recognition) more so than UCL, stemming from their greater expertise.
- UCL staff seemed to use anchoring and availability heuristics i.e. fixating on impression from first few pages and/or documents that were particularly striking.
- Recognition of potential impact on doctor's career and wellbeing if they referred to IOT, evidence of responsible and compassionate approach.
- Backgrounds of individuals shaped the way they thought and felt about cases e.g. clinicians drew upon their work experience as a point of reference, non-clinicians would put themselves in position of members of public.
- Majority of GMC staff recognised they had personal biases but did not think these impacted on their decisions.
- Case examiners thought differently about cases regarding locum doctors about whom serious concerns are raised, than non-locums, and possibly those with poor English language skills. This is not necessarily problematic and is rooted in GMC training and legislation that focuses on risk.
- Some cases evoked emotional reactions that revealed the assumptions/judgements case examiners can make when trying to make sense of allegations. This is inevitable but may be worth considering in training how their feelings can influence their decisions.
- Decisions were not right or wrong, but there was a sense of what made a decision "good".
- A good decision was reasoned well and not based on whether others agreed with their choice.
- Decisions that were based on evidence and were written up transparently were reasoned well.
- Well-reasoned decisions gave them confidence in their role and enabled them to disconnect from emotionally challenging cases.

Conclusions

In the pilot **experimental part** the two main research questions were:

- 1) Consider if the greater likelihood of BME UK qualified doctors being referred compared to other groups of doctors can be justified by the guidance for referral to IOT based on the information available to the decision maker at the time.
- 2) Consider if the greater likelihood of doctors with a PMQ from outside the UK being referred compared to other groups of doctors can be justified by the guidance for referral to IOT based on the information available to the decision maker at the time.

It is not possible to say with any degree of certainty whether or not the decisions made by the GMC case examiners to refer a case to the IOT were influenced by the ethnicity or PMQ of the doctor from the results of this pilot.

The results found no evidence that the UCL case examiners' decisions were influenced by the ethnicity or PMQ of the doctors under investigation, suggesting UCL case examiner decisions were unbiased. GMC case examiners referred slightly fewer white UK graduates and non-UK graduates compared to UCL case examiners. However they did not differ in their referral rates of BME UK graduates.

Interpreting these results is not straightforward. If the GMC case examiners were biased in favour of white doctors, it would be expected that GMC case examiners would also be biased in favour of UK doctors. The pattern of results suggests that if anything, GMC case examiners under-referred non-UK doctors (the majority of who are BME doctors) compared to UCL case examiners, which is difficult to explain in reference to GMC case examiner bias. We are unable to explain the reasons behind these differences however we urge caution with interpretation of this finding as the sample size in this pilot was small and the pilot identified the impracticality of proceeding to a larger sample. Further the cases were randomly allocated so were not stratified for level of severity/risk, there were differences in the selection of and the level of decision making experience between UCL and GMC case examiners and there have been significant changes to case examiner training and guidance over the time frame of the cases.

In the **qualitative study** the research question was mainly 'to determine whether differences in the proportion of cases of non-white doctors/overseas trained referred to IOT (as compared to white/UK trained doctors) involved any bias on the basis of the information at the time a complaint was received.'

We used decision science frameworks to analyse and review the themes generated by interviews with case examiners. This provided clear evidence of good practise in case examiner training and the guidance provided to case examiners. There was also clear use of systemic processing when writing their justification. Although there was evidence of the use of heuristic processing in decision making of GMC case examiners, this was in keeping with the guidance and training given to the case

examiners. There was no evidence that the difference in the proportion of non-white doctors/overseas referred was due to bias on the part of the case examiners.

References

1. Payne J & Bettman, J (2004): In Blackwell handbook of Judgement and Decision making. Chapter 6. Walking with the Scarecrow: The Information-Processing Approach to Decision Research (2004).
2. Baron, J. (2008). Thinking and Deciding. New York. Cambridge University Press. 4th edition.
3. Newell, A. & Simon, H. (1972). Human Problem Solving. Englewood Cliffs, NJ: Prentice -Hall.
4. Broadbent DE (1958) Perception and communication, Pergamon, London.
5. Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39, 752-766.
6. Chen, S., & Chaiken, S. (1999). The heuristic-systematic model in its broader context. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 73-96). New York: Guilford.
7. Croskerry, P. (2002), Achieving Quality in Clinical Decision Making: Cognitive Strategies and Detection of Bias. *Academic Emergency Medicine*, 9: 1184–1204.
8. Tversky and Kahneman (1974) Judgment under uncertainty: heuristics and biases. *Science*, 185, 1124-1131.
9. Gigerenzer, G (2008). Why heuristics work. *Perspectives on Psychological Science*. Vol 3 (1) 20-29.
10. Wegwarth, O., Gaissmaier, W & Gigerenzer G (2009). Smart strategies for doctors and doctors-in-training: heuristics in medicine *Medical Education* 2009: 43: 721–728.
11. Kahneman, D. and Tversky, 1972. Subjective probability: a judgment of representativeness. *Cognitive Psychology* 3, 430-454.
12. Kahneman D, Slovic P Tversky A (1982) *Judgement under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.
13. Tversky, A & Kahneman, D. (1973) Availability: a heuristic for judging frequency and probability *Cognitive Psychology*, 5, 207-232.
14. Denney, ML., Freeman, A & Wakeford, R (2013). MRCGP CSA: are the examiners biased, favouring their own sex, ethnicity and degree source? *BR J Gen Pract* DOI: 10.3399/bjgp13X674396
15. McManus IC, Thompson M, Mollon J. (2006). Assessment of examiner leniency and stringency ('hawk-dove effect') in the MRCP(UK) clinical examination (PACES) using multi-facet Rasch modelling. *BMC Med Educ*. Aug 18;6: 42.
16. Cook et al (2016) When Assessment Data are Words: Validity Evidence for Qualitative Educational Assessments. *Academic Med*. 91(10):1359-1369.
17. General Medical Council (2017) Understanding how organisations ensure that their decision making is fair
https://www.gmc-uk.org/-/media/documents/Understanding_how_organisations_make_their_decision_making_fair.pdf
[72291898.pdf](https://www.gmc-uk.org/-/media/documents/Understanding_how_organisations_make_their_decision_making_fair.pdf)

Appendix 1: Study information sheet for UCL participants

Study title: Decisions made to refer doctors to The Medical Practitioners Interim Orders Tribunal panel.

This study has been approved by the UCL Research Ethics Committee (Project ID Number 6281/002)

Research team:

Dr Alison Sturrock (Principle Investigator)

Dr Leila Mehdizadeh (Co-Investigator- currently on maternity leave)

Dr Katherine Woolf (Co-Investigator)

Dr Antonia Rich (Co-Investigator)

Dr Rowena Viney (Co-Investigator)

Ms Marcia Rigby (Project Manager)

Background:

The General Medical Council (GMC) have commissioned UCL medical school to conduct a study on decision making during the investigation process of doctors who have fitness to practise concerns. When reviewing a case, the GMC can decide whether a doctor should be reviewed at The Medical Practitioners Tribunal Service (MPTS) Interim Orders Tribunal Panel (IOT) whilst allegations about a doctor's fitness to practise are resolved. The role of the panel is to consider whether it is necessary to restrict a doctor's registration in order to protect patients while an investigation is being carried out. The panel make an assessment of the risk that the doctor might pose based on the available information.

Why am I being contacted?

UCLMS staff members are invited to take part in a decision making task and a one to one interview. Any staff member with decision making responsibilities in their current UCL role is eligible to apply.

What does the study involve?

Each participant will review cases of doctors who have been referred to the GMC. Your task will be to read the cases closely and decide whether or not you think each one should be referred to an Interim Orders Tribunal or not. For each case, there will be a short form for you to fill out to record your decision and reasons why. We expect this task to take up to a maximum of two hours per case (some will take less and some may take longer).

To help you make these decisions, you will be given training by a GMC trainer. This will be a two day training session. You will be required to attend on both dates.

We would also like to interview you after you have made the decisions to discuss your experience of making these decisions and how you implemented the GMC's guidance. These will be conducted one to one with a researcher and will last a maximum of 60 minutes. The interviews will be audio-recorded to ensure your views are accurately represented. All data will be anonymised.

Is participation voluntary?

It is entirely your choice whether you participate or not. There will be no adverse consequences for you by not participating. If you choose to participate in the study and later decide you wish to withdraw you can do so at any time without giving a reason.

As a compensation for your time, UCL will pay you £250 through your salary. As this is paid through your salary you will be taxed on this payment. The advantage of paying you in this way is that you will not have to declare any extra earnings to HMRC.

What happens to the data?

A statistical analysis will be performed to determine if there are any significant differences between participants' decisions and the actual decisions made by the GMC on the same cases. Interview recordings will be transcribed by a third party and analysed by the research team to generate themes from the data.

Confidentiality and anonymity

All data will be collected and stored in accordance with the Data Protection Act 1998. Data will be reported in an anonymised manner so no one beyond the research team will be able to recognise your identity from it. Data will be kept strictly confidential and stored on a shared drive with access restricted to the research team only. The person transcribing the recordings will erase any versions of the recordings and transcripts on their computers soon after they have given transcriptions to the research team.

What happens next?

If you would like more details about this study and/or are interested in participating then please contact Ms Marcia Rigby on m.rigby@ucl.ac.uk by 5pm on Friday 16th December.

Appendix 2: UCL case examiners instructions

Instructions for UCL Case Examiners

Thank you for taking part in this GMC-commissioned research project. Your time and effort is essential to the success of the project and is much appreciated.

This document contains all the information you need to take part, however if you have any questions please contact the team: medsch.iop-study@ucl.ac.uk

Your task

You will be allocated a total of 12 case files and 12 forms – one per case.

For each case you will need to:

- **Complete the Case Information Box** on the relevant “Record of IOP decision” form.
- **Review all information** within the case file.
- **Decide whether the doctor in that case should or should not be referred to a GMC Interim Orders Panel.**
- **Record your decision and reasoning** on the relevant “Record of IOP decision” form.

Materials you will need

- 1 Encrypted USB stick containing case files (pdfs) and 1 Excel spreadsheet.
- 1 “GMC’s IOP guidance” form (paper or electronic version).
- Blank “Record of IOP decision” forms (electronic Word documents) – one per case file.

Standardised procedure

It is vital that all UCL case examiners follow a standardised set of instructions when reviewing and deciding about your allocated cases. This will enable the research team to draw conclusions about the decisions made by all case examiners. If each case examiner does something different it will invalidate the results.

Below are the set of standardised instructions. Please follow these exactly in the order that they appear.

Standardised procedure for UCL case examiners

Open up the Excel spreadsheet on your USB stick. You will see it contains a case reference number and doctor name for each case file you have been allocated.

- Open the first pdf case file corresponding to the first line on the Excel spreadsheet. The case file is password-protected. The password is the doctor's full name WITHOUT spaces in lowercase. For example the password to open a Dr John Smith's case file would be johnsmith.
- The first page of the pdf case is called "Index to case bundle front sheet". This summarises the doctor's details.
- Open the Word document "1. Record of IOP decision.docx" and complete the Case Information Box for the case you are about to start reviewing. Do not complete any other part of this document until you have read through the entire case file.
- Save this Word document as "1. Record of IOP decision_your surname"
- Now read through the entire pdf case file.
- Decide whether you think this case should be referred on to a panel or not. You may refer to the "GMC's IOP guidance" form.
- Complete Section 2 of the Word Document "1. Record of IOP decision_your surname" and save it on your encrypted USB stick.
- Close the pdf case file.

Repeat all the steps above until you have reviewed and recorded your decision for each of your allocated case files.

Points to note

Redaction: Patient information has been redacted in **Black**. References to the identity of the doctor under investigation have been redacted in **Blue**.

Opening case files: You should only have one case open at a time on your computer. Please do not open multiple case files simultaneously.

Confidentiality: This is sensitive research commissioned by the General Medical Council and contains highly confidential information. The cases you are reviewing are real. You are reminded of the confidentiality agreement that you have signed to take part in this study. It is vital that you keep your allocated case files strictly private. Do not allow anyone else to see them.

You must also never speak to anyone about the cases you have seen, including fellow participants in the study. The only person you may speak to about your cases is your mentor (see below).

Your mentor: Some cases may contain distressing or shocking information which you may feel you would like to talk to someone about it. You have been allocated a mentor for the project, and your mentor will arrange regular 'catch-ups' at which you can talk about anything to do with the project, including details of the cases. Please note your mentor cannot advise you in decision-making. Attending catch-ups is not required but we encourage it, or if you want to talk about anything to do with the project. You can also contact any of the project team (contact details below) to discuss any project-related issues, and can contact UCL's Employee Assistance Service to talk to an independent person http://www.ucl.ac.uk/hr/occ_health/eap.php

Length of cases: Some cases are more complex than others and vary in the amount of information that have been collated. You will notice some case files also contain duplicated information e.g. the same letter may appear twice or doctor's CV and employment history may appear twice. It is for **you** to decide how closely you pay attention to each of the documents in a file. Part of your task is to evaluate how relevant you think each document is in helping you reach your decision about whether to refer the case to panel or not.

Thank you again for your time and effort

Alison Sturrock a.sturrock@ucl.ac.uk

Leila Mehdizadeh l.mehdizadeh@ucl.ac.uk

Katherine Woolf k.woolf@ucl.ac.uk

Marcia Rigby m.rigby@ucl.ac.uk

Antonia Rich a.rich@ucl.ac.uk

Rowena Viney r.viney@ucl.ac.uk

Appendix 3: Standardised form for UCL case examiners

1. Please complete the Case information BEFORE reviewing the case:

Case information	
Your name	
Case number	
Doctor's name	
Doctor's ethnicity	
Doctor's PMQ country	
Doctor's sex	

2. Please complete the following AFTER reviewing the case.

Summarise concerns of case

Do you want to send this case to an Interim Orders Panel? Please select one option. **Yes** **No**

If you selected 'Yes' what did you base your decision on? Please select one or more option. If you selected 'No' leave this blank and move on to final part.

Risk to patient safety

Public confidence in medical profession

Interest of the doctor

Please expand on <u>why</u> you chose to <u>refer or not</u> refer this case to an Interim Orders Panel.
--

Appendix 4: Demographic data of GMC case examiners over the time period of active cases

	Date from	Date to	Gender	BME/Non-BME	Age at start of role
CaseExaminer1	20/01/2005	29/04/2016	Female	Non-BME	50+
CaseExaminer2	30/01/2006	03/04/2017	Male	Non-BME	50+
CaseExaminer3	01/04/2004	31/12/2099	Male	Non-BME	50+
CaseExaminer4	01/04/2004	03/08/2017	Male	BME	40-49
CaseExaminer5	11/02/2013	31/12/2099	Male	BME	50+
CaseExaminer6	17/10/2011	31/12/2099	Male	Non-BME	30-39
CaseExaminer7	01/04/2004	30/04/2013	Female	Non-BME	50+
CaseExaminer8	08/07/2004	15/10/2010	Female	Non-BME	30-39
CaseExaminer9	30/01/2006	08/01/2017	Male	Non-BME	40-49
CaseExaminer10	01/04/2004	31/12/2013	Male	Non-BME	50+
CaseExaminer11	06/01/2005	31/12/2099	Male	Non-BME	30-39
CaseExaminer12	01/04/2004	31/12/2012	Female	Non-BME	40-49
CaseExaminer13	01/05/2012	31/12/2099	Male	Non-BME	40-49
CaseExaminer14	23/05/2012	31/12/2099	Female	Non-BME	50+
CaseExaminer15	15/10/2012	31/12/2099	Male	Non-BME	40-49
CaseExaminer16	03/01/2013	17/09/2017	Female	Non-BME	40-49
CaseExaminer17	21/01/2013	31/12/2099	Male	Non-BME	30-39
CaseExaminer18	06/03/2013	31/12/2099	Male	Non-BME	40-49
CaseExaminer19	17/03/2014	31/12/2099	Female	Non-BME	20-29
CaseExaminer20	17/03/2014	31/12/2099	Female	Non-BME	40-49
CaseExaminer21	01/04/2014	31/12/2099	Female	Non-BME	40-49
CaseExaminer22	31/03/2014	17/09/2017	Male	Non-BME	40-49

Appendix 5: GMC case examiners invitation

Dear (case examiner's name)

Invitation to take part in GMC funded research project about your experiences of making decisions to refer doctors to Interim Orders Panels

I am writing to invite you to take part in a confidential interview study about your decision making role with regards to doctors that are complained about to the GMC. This research study is being carried out by University College London Medical School and is funded by the GMC.

You are being invited to participate as you currently work as a case examiner for the GMC and make decisions about whether doctors who have been complained about, should be referred on to Interim Orders Panel or not.

The aim of this research is to explore;

- how you decide to refer a doctor's case to Interim Orders Panel or not.
- how you implement the GMC's guidance to arrive at a decision.
- your views about the quality of guidance and training received to support your role.

Your participation is entirely voluntary, if you agree to take part you will be asked to participate in a **30-minute one-to-one interview** with a researcher face-to-face or over the phone at a time and place that suits you.

Interviews will be audio-recorded but data will be kept strictly **confidential**, and will be **anonymised** before being analysed and written up. **All data will be collected and stored in accordance with the Data Protection Act 1998.** This project has been approved by the UCL Ethics Committee ref: 6281/002

If you are interested in taking part or have questions about the research, please get in touch with me on KClapton@gmc-uk.org.

It is up to you to decide whether to take part or not; choosing not to take part will not disadvantage you in any way. If you do decide to take part you are still free to withdraw your data at any time until 26th February 2016 (when the report will be written), without giving a reason. Please see attached information sheet for more information.

Best wishes,
Kerrin Clapton (Research Policy Manager)

Appendix 6: UCL interview questions

Preliminary

- Can you tell me about your job - what is your usual role at UCL?
- Approximately how many cases have you made decisions about?
- What training did you receive about making these retrospective decisions to refer to IOP?

General main questions

- What information did you take into account most when making a decision to refer to IOP or not?
- How did you use the GMC's guidance on referral to an IOP?
- How helpful did you find the guidance?

Specific questions to explore their reasoning about easy and difficult cases

- In the cases you reviewed was there a case when you felt certain that a doctor should be referred to IOP? Why was that?
- Was there a case when you felt certain that no referral was necessary? Why was that?
- If you were unsure about whether to refer to IOP or not, how did you resolve this?
- Were some cases easier to make a decision about than others? Why?
- **Has there been an occasion when a case has made you feel a certain way that may have impacted on your decision?**

Closing

Anything else you wish to comment on about this decision making role?

Appendix 7: GMC interview questions

Preliminary

- Can you tell me about your job - what is your role at GMC?
- Roughly how many cases have you made decisions about?
- Tell me about any training you received to help you make decisions about whether to IOP.

General main questions

- Talk me through how you decide whether or not to refer a doctor to an IOP? What happens first? Then...?
- What information do you take into account most when making a decision to refer to IOP or not?
- How do you use the GMC's guidance on referral to an IOP?
- How helpful do you find the guidance?

Specific questions to explore their reasoning about easy and difficult cases

- Can you describe a time when you felt certain that a doctor should be referred to IOP?
- Can you describe a time when you felt certain that no referral was necessary?
- Do you spend more time on some cases than others, if so why?
- What do you do if you are unsure about the appropriate decision to take?
- What makes some cases easier to make a decision about than others?
- Has there been a time when you later felt uncomfortable about a decision you had made?
 - Is it possible to be objective about IOP decisions?
 - Have you noticed any groups of doctors are more referred to IOP than others?

Closing

Anything else you wish to speak about with regards to your decision making role?