Emily Unwin, Katherine Woolf, Jane Dacre and Henry WW Potts

Sex differences in fitness to practise test scores:

a cohort study of GPs

Abstract

Background
Tests of competence are written and clinical assessments taken by doctors under investigation by the General Medical Council (GMC) who have significant performance concerns. Male doctors on average perform more poorly in clinical assessments than female doctors, and are more likely to be sanctioned. It is unclear why.

Aim
To examine sex differences in the tests of competence assessment scores of GPs under investigation by the GMC, compared with GPs not under investigation, and whether scores mediate any relationship between sex and sanction likelihood.

Design and setting
Retrospective cohort study of GPs' administrative tests of competence data.

Method
Analysis of variance was undertaken to compare written and clinical tests of competence performance by sex and GP group (under investigation versus volunteers). Path analysis was conducted to explore the relationship between sex, written and clinical tests of competence performance, and investigation outcome.

Results
On the written test, female GPs under investigation outperformed male GPs under investigation (Cohen's $d = 0.28$, $P = 0.01$); there was no sex difference in the volunteer group (Cohen's $d = 0.02$, $P = 0.93$). On the clinical assessment, female GPs outperformed male GPs in both groups (Cohen's $d = 0.61$, $P<0.0001$). A higher clinical score predicted remaining on the UK medical register without a warning or sanction, with no independent effect of sex controlling for assessment performance.

Conclusion
Female GPs outperform male GPs on clinical assessments, even among GPs with generally very poor performance. Male GPs under investigation may have particularly poor knowledge. Further research is required to understand potential sex differences in doctors who take tests of competence and how these impact on sex differences in investigation outcomes.

Keywords
female; fitness to practise; general practitioners; male; performance; primary care.

INTRODUCTION

Erasure or suspension from the UK medical register can result in loss of career and income; lesser sanctions (conditions and undertakings) and warnings can also limit doctors' careers and hamper career progression.[1,2] Male doctors are nearly three times more likely to be sanctioned than female doctors.[3] The reasons underlying this difference are not clear.

Performance in academic assessments is associated with future disciplinary action,[4,5] suggesting that sanctioned doctors may have deficient medical knowledge and clinical skills. Research has also demonstrated that female doctors perform better academically,[6] particularly in clinical[7–10] rather than written assessments.[7,11,12]

Role of the General Medical Council

The role of the General Medical Council (GMC) is to ensure proper standards in the practice of medicine in the UK, thereby protecting, promoting, and maintaining the health and safety of the patient population and the community as a whole.[13] Under the Medical Act 1983, the GMC holds the powers to take action against a doctor's registration if the doctor's fitness to practise is impaired for one or more of a number of reasons, including deficient professional performance.[14,15]

The GMC triages information received about complaints that require investigation.[16] If during an investigation a doctor's professional performance is called into question, they may be required to undergo a performance assessment consisting of two main parts: a peer review and test of competence.[17] The GMC's decisions about the investigated doctor's fitness to practise, and the outcome of the investigation, are based on the doctor's performance in both parts.

At any point during an investigation it is possible for the doctor to apply to be removed from the medical register (voluntary erasure), which if granted means the doctor does not have to complete the investigation.[18] However, should the doctor wish to return to the register and practise medicine, they would have to demonstrate that they are fit to practise.

Tests of competence are designed to identify gaps in the knowledge and clinical skills of doctors under investigation for poor performance.[13] They comprise a written knowledge test (KT) and an objective structured clinical examination (OSCE). In general practice there is also simulated surgery (SS), which simulates a typical GP surgery, and assesses the doctor's clinical, management, and communication skills.[19]

To ensure tests of competence are fair and fit for purpose, they are 'piloted' with doctors with no known fitness to practise concerns and who volunteer and are paid

E Unwin, BSc, MBBS, MSc, AFHEA, PhD candidate; **K Woolf**, BSc, PhD, FHEA, associate professor in medical education and NIHR career development fellow in medical education; **J Dacre**, BSc, MD, FRCP, professor of medical education, Research Department of Medical Education, University College London Medical School, Royal Free Hospital, London. **HWW Potts**, BA, MSc, PhD, PGCert, associate professor in health informatics, Institute of Health Informatics, University College London, London.

Address for correspondence
Emily Unwin, Research Department of Medical Education, Royal Free Hospital Hampstead, London WC1E 6BT, UK.

Email: emily.unwin.12@ucl.ac.uk

Submitted: 23 July 2018; **Editor's response:** 21 August 2018; **final acceptance:** 18 September 2018.

This is the full-length article (published online 26 Feb 2019) of an abridged version published in print. Cite this version as: **Br J Gen Pract 2019; DOI: https://doi.org/10.3399/bjgp19X701789**

How this fits in

Male doctors are more likely to face disciplinary action and have their medical registration acted on than female doctors. Doctors who are referred to the General Medical Council (GMC) because of concerns regarding their performance may be required to complete a set of assessments to assess their medical knowledge and clinical skills. This study has shown that, in this highly selected population of doctors, female doctors perform better than male doctors at the written and clinical assessments, and that performance at the clinical assessment predicts whether a GP being investigated by the GMC will remain on the medical register without receiving a warning or sanction. Sex has no independent effect on the outcome. Further research is required to better understand the reasons why female GPs perform better at the clinical assessment than male GPs.

to complete them. There is no pass mark: the performance of the doctor under investigation is compared with reference groups of volunteers who have completed the same questions, and this contributes to the findings of the GMC investigation.[20] Further information about the GMC's tests of competence can be found in some of the references.[13,21,22]

Study aims

The authors hypothesised that male doctors on average have greater gaps in their medical knowledge and clinical skills than female doctors, which contributes to sex differences in disciplinary action. This study aimed to explore:

- the presence and magnitude of sex differences in performance on the written and clinical components of the GMC's tests of competence;
- whether tests of competence sex differences are comparable for doctors currently under investigation by the GMC and doctors not currently under investigation and without restrictions on their medical registration, who had volunteered to complete the tests of competence; and
- whether any relationship between sex and likelihood of disciplinary action is mediated by tests of competence performance.

METHOD

Study design, setting, and source of data

This was a retrospective cohort study of GPs in the UK involving GPs undergoing tests of competence as part of a GMC fitness to practise investigation, and GPs not currently under investigation and with no restrictions on their medical registration who volunteered to take tests of competence. Data were obtained from the GMC and the Research Department of Medical Education at University College London.

Outcome of the GMC fitness to practise investigation

This study examined whether the outcome of the GMC investigation was related to GP sex, written assessment score and SS score, and whether tests of competence performance mediated the relationship between sex and outcome.

Investigation outcome was collapsed into a nominal variable: no sanction imposed, warning or sanction imposed, and no longer registered on the List of Registered Medical Practitioners (Box 1). Warnings and sanctions were combined because warnings are recorded on the List of Registered Medical Practitioners for a period of time and can affect a doctor's career progression.[1]

Population

The study included all GPs between 2008 and 2013 under investigation by the GMC and required to complete tests of competence between 2008 and 2014. The study focused on general practice because it is the largest medical specialty and attracts the most complaints.[23]

The comparison group comprised GPs not currently under investigation and with no restrictions on their registration, who voluntarily completed the KT and/or SS assessment as part of the tests of competence pilots.

Box 1. Definition of outcome types after General Medical Council fitness to practise investigation

Outcome type	Definition
No sanction imposed	No impairment found during the investigation and no restrictions imposed on the doctor's medical registration
Warning/sanction imposed	Warning issued, but no restrictions placed on the doctor's medical registration **OR** Sanction imposed resulting in a restriction on the doctor's medical registration. Sanctions include undertakings, conditions, suspension from the LRMP, and erasure from the LRMP[3,23]
No longer registered on the LRMP	Administrative erasure from LRMP, voluntary erasure from LRMP, or doctor deceased[23]

LRMP = List of Registered Medical Practitioners.

Table 1. Mean and standard deviation of the standardised knowledge test scores and standardised simulated surgery scores by GP group and sex

	GPs under investigation		GP volunteers	
KT	Males (*n* = 94)	Females (*n* = 23)	Males (*n* = 206)	Females (*n* = 276)
Mean KT *z*-score (SD)	−5.16 (4.70)	−3.93 (3.20)	−0.01 (0.98)	0.01 (0.98)
SS	Males (*n* = 92)	Females (*n* = 23)	Males (*n* = 268)	Females (*n* = 57)
Mean SS % score (SD)	52.0 (18.2)	64.4 (18.7)	59.7 (18.4)	70.5 (15.1)

KT = knowledge test. SD = standard deviation. SS = simulated surgery.

There were two versions of the SS test in circulation during the timeframe of this study. Most doctors completed version 1, whereas version 2 was used primarily to reassess doctors. The OSCE was not used for this study because the OSCE and simulated surgery have been found to have significant overlap.[19]

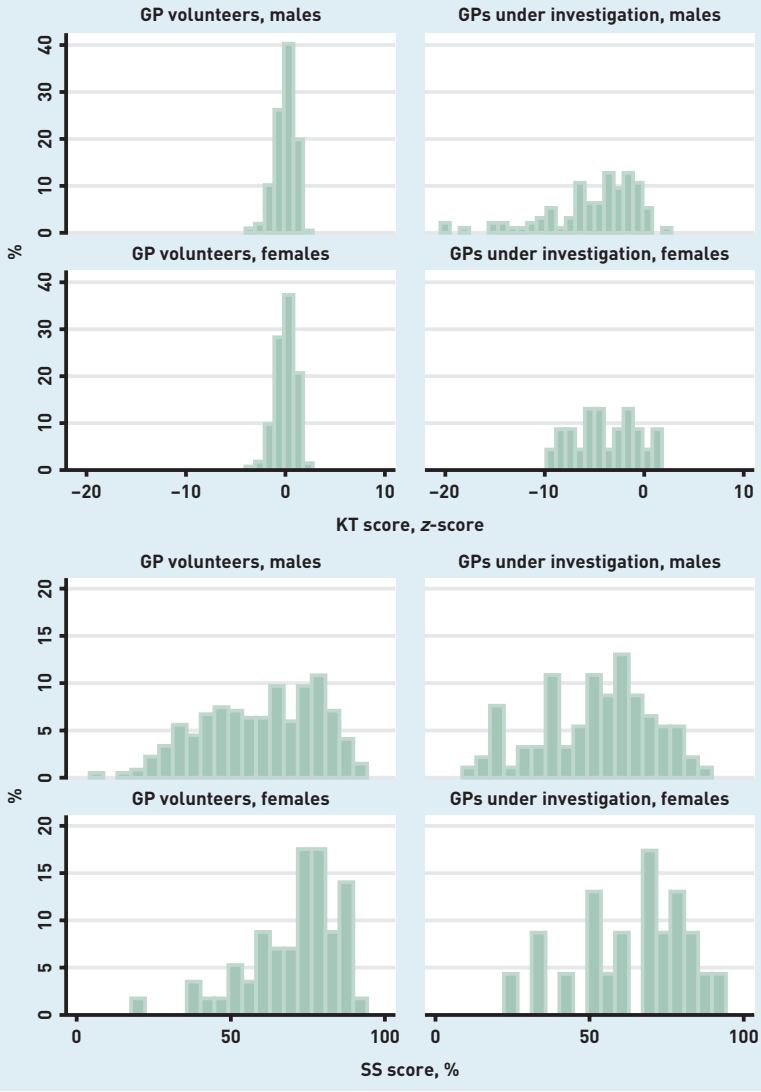


Figure 1. Distribution of standardised knowledge test (KT) and simulated surgery (SS) scores by GP group and sex.

Statistical methods

The KT scores were *z*-transformed for both GP groups (under investigation; volunteers). Analysis of variance was performed to test the performance at the KT and SS test by doctors' sex and GP group, checking for interaction effects. Multinomial and binomial logistic regression models were built, and then a path analysis model was built using multiple regression, with each variable being set as the dependent variable in turn. Paths were included in the model if they were significant at $P<0.05$. Statistical analyses used Stata (version 12/SE). The STROBE statement[24] guided the reporting.

RESULTS

Descriptive results

Doctors under investigation. A total of 120 GPs completed a KT and SS test; 24 (20%) were female and 96 (80%) were male. KT scores were missing for 3/120 (2.5%) GPs and therefore their data were excluded from further KT analysis, resulting in 117 GPs under investigation with KT scores. Five of the 120 (4%) GPs had completed version 2 of the SS test and therefore their data were excluded from further SS test analysis, resulting in 115 GPs under investigation with SS test scores.

Volunteer doctors. A total of 482 GPs (43% males) completed a KT between 2008 and 2014. Complete data were available for 325/349 GPs who completed a SS test between 1997 and 2006. Sex data were missing for 22/349, and a further 2/349 were missing complete SS test data, and were excluded.

Test of competence results

Means and standard deviations for the KT and SS test in GPs under investigation and for volunteer GPs are shown in Table 1, with distributions of scores shown in Figure 1.

Written assessment (KT). A total of 599 GPs (117 under investigation and 482 volunteers) completed a KT. There was a significant interaction between sex and GP group on the KT: $F(1595) = 5.16$, $P = 0.02$. In GPs under investigation, females obtained higher scores than males (mean difference in *z*-score = 1.23, Cohen's $d = 0.28$, $P = 0.01$), but the volunteer GPs showed no evidence of a sex difference (mean difference in *z*-score = 0.02, Cohen's $d = 0.02$, $P = 0.93$).

Clinical assessment (SS). A total of 440 GPs (115 under investigation and 325 volunteers) completed the SS test. The mean score was 59.7 (standard deviation = 18.7); range

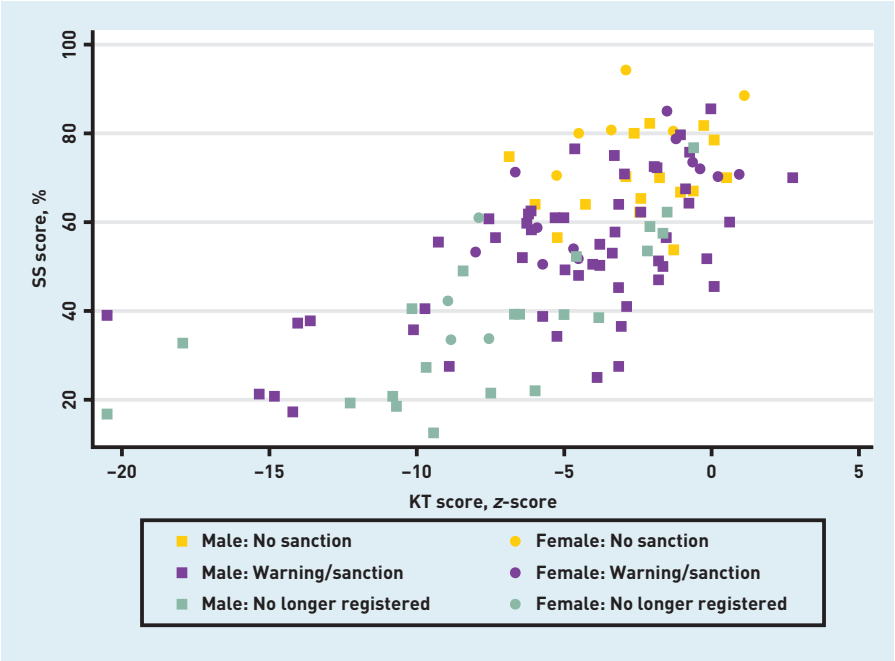


Figure 2. Written (KT) and clinical assessment (SS) scores of GPs under investigation and General Medical Council investigation outcome, by GP sex. KT = knowledge test. SS = simulated surgery.

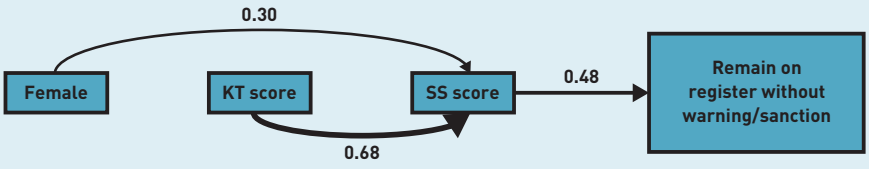
4.2 to 94.3. There was no interaction between sex and GP group: F (1436) = 0.10, $P = 0.75$. There was a main effect of sex, with females obtaining higher scores (mean difference = 11.0, Cohen's $d = 0.61$, $P<0.0001$); and a main effect of GP group, with volunteers obtaining higher scores (mean difference = 7.14, Cohen's $d = 0.39$, $P = 0.0004$).

KT and SS test scores and the GMC investigation outcome. The study then focused on the 112/120 investigated GPs with complete data. Fisher's exact test showed no sex differences in sanctions ($P = 0.61$; of 22 females, 12 were sanctioned, six were not, and four were no longer registered; of 90 males, 53 were sanctioned, 16 were not, and 21 were no longer registered). A *t*-test showed no significant sex difference on the KT ($P = 0.26$), but female GPs significantly outperformed male GPs on the SS test ($P<0.001$).

Figure 2 shows the KT and SS scores of GPs under investigation by outcome type (no sanction, warning/sanction imposed, and no longer registered). Most GPs in the 'no sanction' category occupy the upper right quadrant with high KT and SS scores; however, many GPs in the 'warning/sanction' or 'no longer registered' categories also have comparatively high KT and SS scores. GPs with low KT or SS scores are always in the 'warning/sanction' or 'no longer registered' categories. GPs with low SS scores have a range of KT scores, but those with low KT scores always have low SS scores.

For the path analysis, outcome type was collapsed into a binary variable (no sanction versus sanction/no longer registered) to increase power (univariate association with sex: $P>0.05$). The final model showed being female predicted SS score; KT score predicted SS score; and SS score predicted remaining on the register without warning/sanction (Figure 3). The effect of sex on final outcome was through higher SS performance, with no independent effect of sex controlling for test performance. Multinomial logistic regression (with three outcome categories) produced similar results.

Figure 3. Path diagram showing the relationship between sex, knowledge test (KT) score, simulated surgery (SS) score, and outcome.



Path coefficients are standardised beta coefficients from the regression analysis. Solid black arrows indicate positive effects, with the width of the arrow lines being proportional to the effect sizes. Only paths significant at $P<0.05$ are shown. A high SS score made it more likely that a GP would end up on the register without a warning/sanction following an investigation, which itself was predicted by being female, and by a high KT score.

DISCUSSION

Summary

Among GPs under investigation by the GMC, females outperformed males on the written and clinical components of the tests of competence. Among GPs not under investigation, females outperformed males on the SS test only. Among GPs under investigation, low SS score increased the likelihood of having a warning or sanction imposed, but male GPs were no more likely than female GPs to have a warning or sanction imposed despite lower test scores. Warnings or sanctions are imposed based on a range of evidence of which the test of competence is one component; the outcome 'warning or sanction imposed' does not demonstrate the severity, which may differ by test of competence score levels.

Strengths and limitations

A strength of this study was the inclusion of nearly all GPs investigated by the GMC for fitness to practise concerns who had completed a test of competence. The comparator group of doctors not under investigation enabled the examination of sex differences at different levels of performance. A weakness of the study is that the volunteer GPs are a self-selecting group who may not be representative of the overall GP population. It is also not possible

to generalise the findings to doctors in other specialties.

Unmeasured factors may have influenced the findings. Data from other specialties show that volunteer doctors differ from doctors under investigation in terms of sex, ethnicity, world region of primary medical qualification, and seniority:[20] factors known to influence performance at clinical exams.[25–27] Demographic data (aside from sex) were missing for many volunteer GPs so it was not possible to adjust for these confounders.

In the group of investigated GPs it would have been interesting to explore whether ethnicity was associated with performance or outcome, but unfortunately ethnicity data were missing for nearly one-third of these GPs. It would also have been of interest to explore the relationship between number of years since primary medical qualification (or age), sex, and performance to see whether recent changes in the sex distribution of GPs,[28,29] and the role and format of the examinations required to gain Membership of the Royal College of General Practitioners (MRCGP),[30] have influenced the performance of GPs. However, because of the small numbers of female GPs and GPs who had been qualified for 10 years or less, it was not possible to explore this further. Data were not included on the OSCE test of competence, although the SS score is a better predictor of investigation outcome.[19]

Comparison with existing literature

Female doctors generally outperform male doctors in postgraduate medical exams in general practice,[8,10,25,31] and other specialties.[7,9,32] The current study has shown that this holds even in a group with overall very poor performance; however, there was no sex difference in the written test for GPs not under investigation.

Previous research has shown that male doctors receive more sanctions after controlling for time since primary medical qualification, non-domestic primary medical qualification, and specialty.[3] However, in the current study there was no evidence to suggest a sex difference in warning/sanction rates among this highly selected group of GPs who completed a test of competence as part of an investigation. This suggests that the final decision on warnings or sanctions showed no sex bias beyond the differences in test performance.

Implications for research and practice

The reasons for sex differences are uncertain. Further research is required to examine whether male and female doctors differ in ways the current study was unable to measure and which influenced their performance.

Lack of insight — being unaware of and not addressing deficiencies — is common among investigated doctors.[33] Previous research found that male volunteer doctors tended to overestimate their written and clinical test of competence scores, suggesting less insight.[34] If these sex differences are present in doctors under investigation, more poorly performing female doctors may remove themselves from the medical register before taking a test of competence. It would be interesting to investigate further those GPs who are no longer registered following an investigation into their fitness to practise, particularly those GPs who were voluntarily erased from the UK medical register, not only in terms of demographics and performance, but also the reasons behind their decision to apply for voluntary erasure.

Female doctors in general have a more patient-centred approach and ask more psychosocial questions than male doctors, which stimulates more patient disclosure.[35] Female doctors, including those under investigation, may therefore learn more information from patients, and perform better at the clinical assessment. Females on average have been found to score higher on dutifulness (a facet of conscientiousness),[36] which is a predictor of performance.[37–40] Female doctors may therefore be higher on personality traits that lead to them maintaining their skills and knowledge, and performing better at assessments. Female doctors have also been shown to have higher person-related values.[41–43] It has been shown that performance in a clinical setting is predicted by the person-related values held by a doctor.[42] It is plausible that female GPs' higher performance is therefore due to differences in skills and attitudes, as well as knowledge.

Another factor that may have influenced doctors' examination performance is dyslexia or another specific learning difficulty. Doctors with a specific learning difficulty may also face extra challenges in the effective performance of their duties, especially if unrecognised or undisclosed.[44] It is uncertain whether there are sex differences in specific learning difficulties such as dyslexia among doctors, but, given that nearly 2% of medical students have dyslexia,[45] it would be important to explore not only how specific learning difficulties may affect learning and performance, but also how to better identify and support those

doctors with specific learning difficulties.

Organisational factors can also affect performance;[37] for example, professional isolation from peers and colleagues can limit opportunities for feedback and development.[33] A possible hypothesis is that male GPs may be more at risk of professional isolation. Evidence suggests that male GPs are more likely to work in a single-handed practice than female GPs.[46] There is also evidence to suggest that professional isolation is more prevalent in rural areas.[47]

It is unclear why it is that, among volunteer GPs, female doctors performed better on the clinical but not the written assessment. It may be that the male doctors in the self-selected group of volunteer GPs had better medical knowledge than the average male GP population.

Future work should explore whether reasons for referral to the GMC (allegation type) differ between the sexes and whether certain allegation types are associated with a higher risk of sanctions.

This study has implications for support for GPs undergoing a fitness to practise investigation and future work could explore how the medical profession could better support those doctors undergoing an investigation.

Funding

Katherine Woolf is funded by a National Institute for Health Research (NIHR) career development fellowship. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

Ethical approval

This study is part of a research project that has received ethical approval from the University College London (UCL) Research Ethics Committee (Project ID: 5025/001).

Provenance

Freely submitted; externally peer reviewed.

Competing interests

Emily Unwin is supported by a UCL Impact studentship. Katherine Woolf is supported by an NIHR Fellowship and is a Higher Education Funding Council for England (HEFCE)-supported staff member at UCL Medical School. She received funding from the General Medical Council (GMC) for two unrelated research studies during the course of the current study. She is educational advisor to the Membership of the Royal Colleges of Physicians of the United Kingdom. Jane Dacre was the President of the Royal College of Physicians during the course of the current study, and was a GMC council member and chaired the GMC Education and Training Committee (2008–2012). She was the principal investigator for the Fitness to Practise contract with the GMC. Henry WW Potts is an HEFCE-supported staff member at UCL. He has received research funding from the GMC.

Acknowledgements

The authors acknowledge the GMC for facilitating access to the GMC data and for reviewing the final version of the manuscript. They also acknowledge Alison Sturrock and Samantha Henry at the Research Department of Medical Education (RDME), UCL, for assisting access to the RDME data and for assisting with queries about the dataset.

Discuss this article

Contribute and read comments about this article: **bjgp.org/letters**

REFERENCES

1. Community Research. *The effects of having restrictions on practice or warnings. Research report for the GMC*. 2015. https://www.gmc-uk.org/-/media/documents/the-effects-of-restrictions-or-warnings-research-report-final_pdf-63538542.pdf (accessed 18 Feb 2019).
2. Rimmer A. GMC warnings should be scrapped, says MDU. *BMJ* 2014; **349:** g6893.
3. Unwin E, Woolf K, Wadlow C, Dacre J. Disciplined doctors: does the sex of a doctor matter? A cross-sectional study examining the association between a doctor's sex and receiving sanctions against their medical registration. *BMJ Open* 2014; **4(8):** e005405.
4. Yates J, James D. Risk factors at medical school for subsequent professional misconduct: multicentre retrospective case-control study. *BMJ* 2010; **340:** c2040.
5. Papadakis MA, Arnold GK, Blank LL, *et al*. Performance during internal medicine residency training and subsequent disciplinary action by state licensing boards. *Ann Intern Med* 2008; **148(11):** 869–876.
6. Ferguson E, James D, Madeley L. Factors associated with success in medical school: systematic review of the literature. *BMJ* 2002; **324(7343):** 952–957.
7. Dewhurst NG, McManus C, Mollon J, *et al*. Performance in the MRCP(UK) examination 2003–4: analysis of pass rates of UK graduates in relation to self-declared ethnicity and gender. *BMC Med* 2007; **5:** 8.
8. Pope L, Hawkridge A, Simpson R. Performance in the MRCGP CSA by candidates' gender: differences according to curriculum area. *Educ Prim Care* 2014; **25(4):** 186–193.
9. Richens D, Graham TR, James J, *et al*. Racial and gender influences on pass rates for the UK and Ireland Specialty Board Examinations. *J Surg Educ* 2016; **73(1):** 143–150.
10. Shaw B, Fox J, Brown J, *et al*. An investigation of factors affecting the outcome of the clinical skills assessment (CSA) in general practice specialty training. *Educ Prim Care* 2014; **25(2):** 91–95.
11. Menzies L, Minson S, Brightwell A, *et al*. An evaluation of demographic factors affecting performance in a paediatric membership multiple-choice examination. *Postgrad Med J* 2015; **91(1072):** 72–76.
12. Rushd S, Landau AB, Khan JA, *et al*. An analysis of the performance of UK medical graduates in the MRCOG Part 1 and Part 2 written examinations. *Postgrad Med J* 2012; **88(1039):** 249–254.
13. Sturrock A, Conlon L, Hatch D, Dacre J. The GMC's Tests of Competence: what happens now? *Clin Teach* 2006; **3(1):** 9–12.
14. General Medical Council. *Our role and the Medical Act 1983*. 2019. https://www.gmc-uk.org/about/legislation/medical_act.asp (accessed 18 Feb 2019).
15. General Medical Council, Medical Schools Council. *Professional behaviour and fitness to practise: guidance for medical schools and their students*. London: GMC, Medical Schools Council, 2016.
16. General Medical Council. *How we investigate concerns*. 2019. https://www.gmc-uk.org/concerns/information-for-doctors-under-investigation/how-we-investigate-concerns (accessed 18 Feb 2019).
17. General Medical Council. *Performance assessments*. 2019. https://www.gmc-uk.org/concerns/doctors_under_investigation/performance_assessments.asp (accessed 18 Feb 2019).
18. General Medical Council. *Guidance on making decisions on voluntary erasure applications*. 2014. http://www.gmc-uk.org/voluntary_erasure_guidance.pdf_25416412.pdf (accessed 18 Feb 2019).
19. Jayaweera HK, Potts HWW, Keshwani K, *et al*. The GP tests of competence assessment: which part best predicts fitness to practise decisions? *BMC Med Educ* 2018; **18:** 2.
20. Mehdizadeh L, Sturrock A, Myers G, *et al*. Doctors who pilot the GMC's Tests of Competence: who volunteers and why? *Postgrad Med J* 2014; **90(1070):** 675–679.
21. General Medical Council. *What will your assessment consist of?* 2019. https://www.gmc-uk.org/concerns/information-for-doctors-under-investigation/performance-assessments/investigation-and-tribunal-directed/what-will-your-assessment-consist-of (accessed 18 Feb 2019).
22. Dacre J, Potts HW, Sales D, *et al*. The development of a new method of knowledge assessment: tailoring a test to a doctor's area of practice. *Acad Med* 2009; **84(8):** 1003–1007.
23. General Medical Council. *The state of medical education and practice in the UK 2014*. 2014. https://www.gmc-uk.org/-/media/documents/somep-2014-final_pdf-58751753.pdf (accessed 18 Feb 2019).
24. von Elm E, Altman DG, Egger M, *et al*. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med* 2007; **4(10):** e296.
25. Esmail A, Roberts C. Academic performance of ethnic minority candidates and discrimination in the MRCGP examinations between 2010 and 2012: analysis of data. *BMJ* 2013; **347:** f5662.
26. Mehdizadeh L, Sturrock A, Dacre J. Are the General Medical Council's Tests of Competence fair to long standing doctors? A retrospective cohort study. *BMC Med Educ* 2015; **15:** 80.
27. Woolf K, Haq I, McManus IC, *et al*. Exploring the underperformance of male and minority ethnic medical students in first year clinical examinations. *Adv Health Sci Educ Theory Pract* 2008; **13(5):** 607–616.
28. Bostock N. The rise of women in general practice. *GP Online* 2018; **8 Mar**: https://www.gponline.com/rise-women-general-practice/article/1458988 (accessed 18 Feb 2019).
29. General Medical Council. *The state of medical education and practice in the UK 2017*. 2017. https://www.gmc-uk.org/-/media/about/somep-2017/somep-2017-final-full.pdf?la=en&hash=B6AD13C9D672F7FCD927498A3F50BB0A2A4286F2 (accessed 18 Feb 2019).
30. Hutt P. The new MRCGP exam. *BMJ* 2006; **332:** DOI: https://doi.org/10.1136/bmj.332.7551.sgp199-a.
31. Siriwardena AN, Irish B, Asghar ZB, *et al*. Comparing performance among male and female candidates in sex-specific clinical knowledge in the MRCGP. *Br J Gen Pract* 2012; DOI: https://doi.org/10.3399/bjgp12X649142.
32. Unwin E, Potts HWW, Dacre J, *et al*. Passing MRCP (UK) PACES: a cross-sectional study examining the performance of doctors by sex and country. *BMC Med Educ* 2018; **18(1):** 70.
33. Holden JD, Cox SJ, Hargreaves S. Avoiding isolation and gaining insight. *BMJ* 2012; **344:** e868.
34. Mehdizadeh L, Sturrock A, Myers G, *et al*. How well do doctors think they perform on the General Medical Council's Tests of Competence pilot examinations? A cross-sectional study. *BMJ Open* 2014; **4(2):** e004131.
35. Roter DL, Hall JA, Aoki Y. Physician gender effects in medical communication: a meta-analytic review. *JAMA* 2002; **288(6):** 756–764.
36. Costa PT, Terracciano A, McCrae RR. Gender differences in personality traits across cultures: robust and surprising findings. *J Pers Soc Psychol* 2001; **81(2):** 322–331.
37. Cohen D, Rhydderch M. Measuring a doctor's performance: personality, health and well-being. *Occup Med (Lond)* 2006; **56(7):** 438–440.
38. Chamorro-Premuzic T, Furnham A. Personality predicts academic performance: evidence from two longitudinal university samples. *J Res Pers* 2003; **37(4):** 319–338.
39. Ferguson E, James D, O'Hehir F, *et al*. Pilot study of the roles of personality, references, and personal statements in relation to performance over the five years of a medical degree. *BMJ* 2003; **326(7386):** 429–432.
40. Lievens F, Coetsier P, De Fruyt F, De Maeseneer J. Medical students' personality characteristics and academic performance: a five-factor model perspective. *Med Educ* 2002; **36(11):** 1050–1056.
41. Clack GB, Head JO. Gender differences in medical graduates' assessment of their personal attributes. *Med Educ* 1999; **33(2):** 101–105.
42. Inglehart MR, Brown DR. Gender differences in values and their impact on academic achievement [abstract]. 10th Annual Meeting of the International Society of Political Psychology, San Francisco, CA, 4–7 July 1987.
43. Woolf K. The academic underperformance of medical students from ethnic minorities. PhD thesis. London: University College London, 2009.
44. Locke R, Scallan S, Mann R, Alexander G. Clinicians with dyslexia: a systematic review of effects and strategies. *Clin Teach* 2015; **12(6):** 394–398.
45. Shrewsbury D. Trainee doctors with learning difficulties: recognizing need and providing support. *Br J Hosp Med (Lond)* 2012; **73(6):** 345–349.
46. Kelly E, Stoye G. *Does GP practice size matter? GP practice size and the quality of primary care*. London: Institute for Fiscal Studies. 2014.
47. Richard HM, Farmer J, Selvaraj S. Sustaining the rural primary healthcare workforce: survey of healthcare professionals in the Scottish Highlands. *Rural Remote Health* 2005; **5(1):** 365.