BJGP British Journal of General Practice

Sex differences in fitness to practise test scores: cohort study of GPs

Journal:	*The British Journal of General Practice*
Manuscript ID	BJGP-2018-0140.R1
Manuscript Type:	Article {quantitative - not RCT}
Date Submitted by the Author:	13-Sep-2018
Complete List of Authors:	Unwin, Emily; University College London Medical School, Research Department of Medical Education (RDME) Woolf, Katherine; University College London Medical School, Research Department of Medical Education (RDME) Dacre, Jane; Royal College of Physicians Potts, Henry; University College London, Institute of Health Informatics
Keywords:	Gender, Performance, General Practitioners

Sex differences in fitness to practise test scores: cohort study of GPs

Emily Unwin[1], Katherine Woolf[1], Jane Dacre[1,2] and Henry W W Potts[3]

[1] Research Department of Medical Education, University College London Medical School, Royal Free Hospital, GF 664, Rowland Hill Street, London NW3 2PF UK

[2] Royal College of Physicians, 11 St Andrews Place, London, NW1 4LE UK

[3] Institute of Health Informatics, University College London, 222 Euston Road, London, NW1 2DA UK

Abstract

Background

Male doctors on average perform more poorly in clinical assessments compared to female doctors; and are more likely to be sanctioned. It is unclear why.

Aims

To examine sex differences in the assessment scores of General Practitioners (GP) under investigation by the General Medical Council (GMC), compared to GPs not under investigation; and whether scores mediate any relationship between sex and sanction likelihood.

Design and setting

Retrospective analysis of administrative Tests of Competence (ToC) GP data. ToC are written and clinical assessments taken by doctors under investigation by the GMC who have significant performance concerns, and a comparator group of volunteer doctors.

Methods

Analysis of variance to compare written and clinical ToC performance by sex and GP group (under investigation vs volunteers). Path analysis to explore the relationship between sex, written and clinical ToC performance, and investigation outcome.

Results

On the written test, women GPs under investigation outperformed men (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, women outperformed men in both GP groups (Cohen's d=0.61, *P*<0.0001). Higher clinical score predicted

remaining on the register without a warning or sanction; with no independent effect of sex controlling for assessment performance.

Conclusion

Women outperform men on clinical assessments, even among GPs with generally very poor performance. Male GPs under investigation may have particularly poor knowledge. Further work is required to understand potential sex differences in who takes the ToC and how this impacts on sex differences in investigation outcomes.

Keywords

Female, General Practice, General Practitioners, Male, Performance, Primary Health Care.

How this fits in

Men are more likely to face disciplinary action and have their medical registration acted upon than women. Doctors who are referred to the GMC with 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, women perform better than men at both 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 women perform better at the clinical assessment than men.

Introduction

Erasure or suspension from the United Kingdom (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 compared to female doctors (3). The reasons underlying this difference are not clear.

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

We hypothesised men on average have greater gaps in their medical knowledge and clinical skills, which contributes to sex differences in disciplinary action.

Study aims

We aimed to explore:

- The presence and magnitude of sex differences in performance on the written and clinical components of the General Medical Council's (GMC) tests of competence (ToC).
- Whether ToC 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 ToC.
- Whether any relationship between sex and disciplinary action is mediated by ToC performance.

Method

Study design, setting and source of data

Retrospective cohort study of General Practitioners (GPs) in the UK: GPs undergoing a ToC 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 a ToC. Data were obtained from the GMC and the Research Department of Medical Education at University College London.

The GMC Fitness to Practise investigation process

The GMC's role 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 - see (14, 15).

The GMC triage information recieved, identifying 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 the ToC (for details see (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 they are fit to practise.

The Test of Competence (ToC)

The ToC are designed to identify gaps in the knowledge and clinical skills of doctors under investigation for poor performance.(13) It comprises a written knowledge test (KT) and an Objective Structured Clinical Exam (OSCE). In General Practice there is also the Simulated Surgery (SS), which simulates a typical GP surgery, and assesses the doctor's clinical, management and communication skills.(19) There were two versions of the SS in circulation during the timeframe of this study. Most doctors completed version 1; version 2 was primarily used to reassess doctors. We did not use the OSCE for this study because the OSCE and Simulated Surgery have been found to have significant overlap.(19)

To ensure the ToC are fair and fit for purpose, they are 'piloted' with doctors with no known fitness to practise concerns and who volunteer and are paid to complete the ToC. There is no ToC 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) For further information about the ToC see (13, 21, 22).

Outcome of the GMC FtP investigation

We examined whether the outcome of the GMC investigation was related to GP sex, written assessment score and SS score, and whether ToC performance mediated the relationship between sex and outcome.

We collapsed investigation outcome into a nominal variable: no sanction imposed, warning or sanction imposed, and no longer registered on the List of Registered Medical Practitioners (LRMP) (see Table 1). We combined warnings and sanctions since warnings are recorded on the LRMP for a period of time and can affect career progression.(1)

Population

All GPs between 2008 and 2013 under investigation by the GMC and required to complete the ToC between 2008 and 2014. We focused on GP because it is the largest specialty and attracts the most complaints.(23)

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

Statistical methods

We z-transformed the KT scores for both GP groups (under investigation; volunteers). We performed analysis of variance to test the performance at the KT and SS by doctors' sex and GP group, checking for interaction effects. We built multinomial and binomial logistic regression models, and then a path analysis model 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 V.12/SE.

The STROBE statement (24) guided our reporting.

Results

Descriptive results

Doctors under investigation

120 GPs completed a KT and SS; 24 (20%) were female and 96 (80%) were male. 3/120 (2.5%) GPs' KT scores were missing, and a further 5/120 (4%) GPs' had completed version 2 of the SS; their data were excluded.

Volunteer doctors

482 GPs (43% men) completed a KT between 2008 and 2014. Complete data were available for 325/349 GPs who completed a SS between 1997 and 2006. 22/349 were missing sex data, and a further 2/349 were missing complete SS data, and were excluded.

The ToC results

Means and standard deviations for the KT and SS in GPs under investigation and for volunteer GPs are shown in Table 2, with distributions of scores shown in Figure 1.

Written assessment (KT)

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(1, 595)=5.16, *P*=0.02]. In GPs under investigation, women obtained higher scores than men (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)

440 GPs (115 under investigation and 325 volunteers) completed the SS. The mean score was 59.7 (SD=18.7); the range 4.2 to 94.3. There was no interaction between sex and GP group: F(1,436)=0.10, *P*=0.75. There was a main effect of sex, with women 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 scores and the GMC investigation outcome

We then focused our study on the 112/120 investigated GPs with complete data. A Fisher's exact test showed no sex differences in sanctions (p=0.61; of 22

females, 12 were sanctioned, 6 were not, and 4 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 women significantly outperformed men on the SS (P<0.001).

Figure 2 shows the KT and SS scores of GPs under investigation by outcome type (no sanction, warning/sanction imposed, 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, we collapsed outcome type into a binary variable (no sanction vs. 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 (see 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.

Discussion

Summary

Among GPs under investigation by the GMC, women outperformed men on the written and clinical (Simulated Surgery, SS) components of the Tests of Competence. Among GPs not under investigation, women outperformed men on the SS only. Among GPs under investigation, low SS score increased the likelihood of having a warning or sanction imposed, but men were no more likely than women 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 ToC is one component; and the outcome 'warning or sanction imposed' does not demonstrate the severity, which may differ by ToC score levels.

Strengths and limitations

A strength was the inclusion of nearly all GPs investigated by the GMC for fitness to practise concerns who completed a ToC. The comparator group of doctors not under investigation enabled us to examine sex differences at different levels of performance. A weakness is that the volunteer GPs are a self-selecting group that may not be representative of the overall GP population. It is also not possible to generalise our findings to other specialties.

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

Within the group of investigated GPs it would have been interesting to explore whether ethnicity was associated with performance or outcome, but unfortunately ethnicity data was missing for nearly one third of these GPs. It would also have been of interest to explore the relationship between number of years since PMQ (or age), sex and performance to see if the recent changes in the sex distribution of GPs (28, 29) and the role and format of the examinations required to gain MRCGP (30) have influenced the performance of GPs, however due to the small number of female GPs and GPs who had been qualified for 10 years or less, we were not able to explore this further. We also did not include data on the ToC OSCE, although the SS is a better predictor of investigation outcome. (19)

Comparison with existing literature

Women generally outperform men in postgraduate medical exams in GP (8, 10, 25) (31) and other specialties.(7, 9, 32) We have shown 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.

Implications for research

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

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

Women doctors in general have a more patient-centred approach and ask more psychosocial questions, which stimulates more patient disclosure.(35) Women doctors, including those under investigation, may therefore learn more information from patients, and perform better at the clinical assessment. Women 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. Women 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 women's 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 (SpLD). Doctors with a SpLD 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 SpLDs 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 SpLD may affect learning and performance, but also how to better identify and support those doctors with SpLD.

Organisational factors can also affect performance;(37) for example, professional isolation from peers and colleagues can limit opportunities for feedback and development,(33) and male GPs may be more at risk of professional isolation, whether in a single-handed or group practice.

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

Conclusions

Women GPs under investigation performed better at written and clinical assessments than their male counterparts, replicating the finding of poorer male academic performance within a highly selected population. Among GPs not under investigation, women outperformed men in the clinical but not the written assessment. Further exploration into possible reasons for these sex differences is required.

Unlike previous research, there was no evidence to suggest a sex difference in warning/sanction rates among this highly selected group of GPs who completed a ToC 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. Previous research demonstrated male doctors receive more sanctions after controlling for time since PMQ, non-domestic PMQ and specialty.(3) Future work should explore whether reasons for referral to the GMC (allegation type) differs between the sexes and whether certain allegations types are associated with 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.

Additional information

Funding

KW is funded by a National Institute for Health Research Career Development Fellowship. The views expressed are those of the author(s) 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 UCL Research Ethics Committee (Project ID: 5025/001).

Competing interests

EU is supported by a University College London Impact studentship

KW is supported by a National Institute for Health Research Fellowship and is a HEFCE supported staff member at UCL Medical School. She received funding from the General Medical Council 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 (United Kingdom).

JD is the President of the Royal College of Physicians of England, 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 Practice contract with the General Medical Council.

HP is a Higher Education Funding Council for England (HEFCE) supported staff member at University College London. He has received research funding from the General Medical Council.

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 RDME, UCL for assisting access to the RDME data and for assisting with the queries we had about the dataset.

References

1. Community Research. The effects of having restrictions on practice or warnings. 2015 November 2015.
2. Rimmer A. GMC warnings should be scrapped, says MDU. BMJ Careers. 2014.
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, Holmboe ES, Lipner RS. Performance during internal medicine residency training and subsequent disciplinary action by state licensing boards. Ann Intern Med. 2008;148(11):869-76.
6. Ferguson E, James D, Madeley L. Factors associated with success in medical school: systematic review of the literature. BMJ. 2002;324(7343):952-7.
7. Dewhurst NG, McManus C, Mollon J, Dacre JE, Vale AJ. 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):186-93.
9. Richens D, Graham TR, James J, Till H, Turner PG, Featherstone C. Racial and gender influences on pass rates for the UK and Ireland specialty board examinations. J Surg Educ. 2016;73(1):143-50.
10. Shaw B, Fox J, Brown J, Hart A, Mamelok J. An investigation of factors affecting outcome of the clinical skills assessment (CSA) in general practice specialty training. Educ Prim Care. 2014;25(2):91-5.
11. Menzies L, Minson S, Brightwell A, Davies-Muir A, Long A, Fertleman C. An evaluation of demographic factors affecting performance in a paediatric membership multiple-choice examination. Postgrad Med J. 2015;91(1072):72-6.
12. Rushd S, Landau AB, Khan JA, Allgar V, Lindow SW. 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):249-54.
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. Medical Act 1983 2018 [Available from: https://www.gmc-uk.org/about/legislation/medical_act.asp.
15. General Medical Council, Medical Schools Council. Professional behaviour and fitness to practise: guidance for medical schools and their students. General Medical Council and Medical Schools Council; 2016.

16. General Medical Council. The investigation process 2018 [Available from: https://www.gmc-uk.org/concerns/the_investigation_process.asp.
17. General Medical Council. Doctors under investigation - performance assessments 2018 [Available from: https://www.gmc-uk.org/concerns/doctors_under_investigation/performance_assessments.asp.
18. General Medical Council. Guidance on making decisions on voluntary erasure applications: GMC; 2014 [Available from: http://www.gmc-uk.org/voluntary_erasure_guidance.pdf_25416412.pdf.
19. Jayaweera HK, Potts HWW, Keshwani K, Valerio C, Baker M, Mehdizadeh L, et al. The GP tests of competence assessment: which part best predicts fitness to practise decisions? BMC Med Educ. 2018;18(1):2.
20. Mehdizadeh L, Sturrock A, Myers G, Khatib Y, Dacre J. Doctors who pilot the GMC's Tests of Competence: who volunteers and why? Postgrad Med J. 2014;90(1070):675-9.
21. General Medical Council. Undergoing Tests of Competence as part of a Performance Assessment 2018 [Available from: https://www.gmc-uk.org/concerns/doctors_under_investigation/undergoing_a_test_of_competence.asp.
22. Dacre J, Potts HW, Sales D, Spencer H, Sturrock A. The development of a new method of knowledge assessment: tailoring a test to a doctor's area of practice. Acad Med. 2009;84(8):1003-7.
23. General Medical Council. The state of medical education and practice in the UK. London; 2014.
24. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. PLoS Med. 2007;4: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, Higham J, Dacre J. Exploring the underperformance of male and minority ethnic medical students in first year clinical examinations. Adv Health Sci Educ Theory Pract. 2008;13(5):607-16.
28. Bostock N. The rise of women in general practice. GPonline. 2018 8 March 2018.
29. General Medical Council. The state of medical education and practice in the UK. London, UK; 2017.
30. Hutt P. The new MRCGP exam. Br Med J. 2006;332(gp199).
31. Siriwardena AN, Irish B, Asghar ZB, Dixon H, Milne P, Neden C, et al. Comparing performance among male and female candidates in sex-specific clinical knowledge in the MRCGP. Br J Gen Pract. 2012;62(599):e446-50.

32. Unwin E, Potts HWW, Dacre J, Elder A, Woolf K. 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 Careers. 2012.
34. Mehdizadeh L, Sturrock A, Myers G, Khatib Y, Dacre J. 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):756-64.
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):322-31.
37. Cohen D, Rhydderch M. Measuring a doctor's performance: personality, health and well-being. Occup Med (Lond). 2006;56(7):438-40.
38. Chamorro-Premuzic T, Furnham A. Personality predicts academic performance: Evidence from two longitudinal university samples. Journal of Research in Personality. 2003;37(4):319-38.
39. Ferguson E, James D, O'Hehir F, Sanders A, McManus IC. 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):429-32.
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):1050-6.
41. Clack GB, Head JO. Gender differences in medical graduates' assessment of their personal attributes. Med Educ. 1999;33(2):101-5.
42. Inglehart MR, Brown DR. Gender differences in values and their impact on academic achievement. 10th Annual Meeting of the International Society of Political Psychology; 4-7th July 1987; San Francisco, California1987.
43. Woolf K. The academic underperformance of medical students from ethnic minorities. 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):394-8.
45. Shrewsbury D. Trainee doctors with learning difficulties: recognizing need and providing support. Br J Hosp Med (Lond). 2012;73(6):345-9.

Table 1 Definition of outcome types

Outcome type (following GMC FtP investigation)	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 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 LMRP, and erasure from the LRMP. (See (3, 23) for further details.)
No longer registered on the LRMP	Administrative erasure from LRMP, voluntary erasure from LRMP or doctor deceased. (See (23) for further details.)

LRMP: List of Registered Medical Practitioners

Table 2 Mean and standard deviation of the standardised KT scores (z-score) and standardised SS scores (%) by General Practitioner (GP) group and sex

	GP under investigation		**GP volunteers**	
Knowledge test	Men (N=93)	Women (N=23)	Men (N=206)	Women (N=276)
Mean KT z-score (SD)	-5.16 (4.70)	-3.93 (3.20)	-0.01 (0.98)	0.01 (0.98)
Simulated Surgery	Men (N=92)	Women (N=23)	Men (N=268)	Women (N=57)
Mean SS % score (SD)	52.0 (18.2)	64.4 (18.7)	59.7 (18.4)	70.5 (15.1)

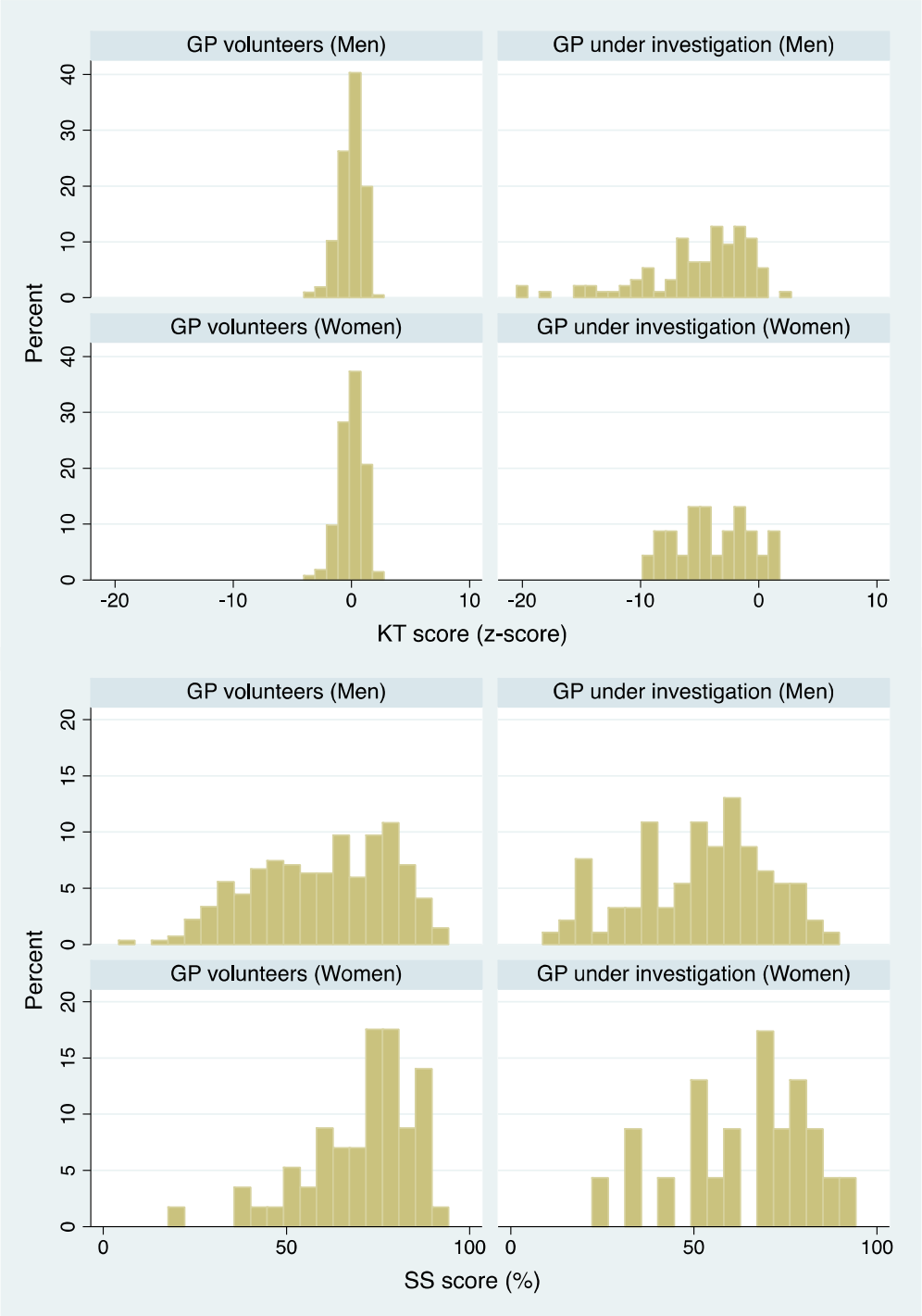

Figure 1 Distribution of standardised Knowledge Test (KT: written assessment) and Simulated Surgery (SS: clinical assessment) scores by General Practitioner (GP) group and sex.

Figure 2 Written and clinical assessment scores of General Practitioners (GPs) under investigation and General Medical Council (GMC) investigation outcome, by GP sex.

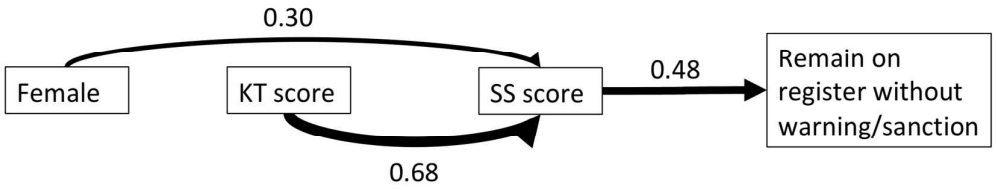

Figure 3 Path diagram showing the relationship between sex, Knowledge Test (KT) score, Simulated Surgery (SS) score and outcome. Path coefficients are standardised beta coefficients from the regression analysis. Solid black arrows indicate positive effects, with the width of the arrow lines being proportional to the effect sizes. Only path significant at P<.05 are shown. A high SS score made it more likely that a General Practitioner would end up on the register without a warning/sanction following an investigation, which itself was predicted by being female, and a by a high KT score.