Uncorrected final version

Please cite the published version:

Woolf Katherine, Haq Inam, McManus I. Chris, Higham Jenny, Dacre Jane. Exploring the underperformance of male and minority ethnic medical students in first year clinical examinations. *Advances in Health Sciences Education Theory and Practice*. 2008 Dec;13(5):607-16. Epub 2007 May 9. DOI 10.1007/s10459-007-9067-1

Running head

Exploring the underperformance of male and minority ethnic medical students in first year clinical examinations

Title

Exploring the underperformance of male and minority ethnic medical students in first year clinical examinations

Author(s) and affiliation(s)

Katherine Woolf[1] Inam Haq[2] I C McManus[3] Jenny Higham[4] Jane Dacre[1]

1. Academic Centre for Medical Education, Division of Medical Education, Royal Free & University College Medical School, Archway Campus, 4th Floor, Holborn Union Building, Highgate Hill, London N19 5LW

2. Brighton & Sussex Medical School, BSMS Teaching Building, University of Sussex Brighton, East Sussex, BN1 9PX

3. Department of Psychology, University College London, 26 Bedford Way, London WC1H 0AP

4. Imperial College School of Medical, Academic Department of Obstetrics and Gynaecology, Imperial College London, St Mary's Campus, Mint Wing, South Wharf Road, London W2 1NY.

Corresponding author:

Katherine Woolf

Address: ACME, 4th Floor, Holborn Union Building, UCL Archway Campus, Highgate Hill, London N19 5LW, UK.

Tel: 020 7288 3546.
Fax: 020 7288 3322.
Email: k.woolf@medsch.ucl.ac.uk

Exploring the underperformance of male and Minority Ethnic medical students in undergraduate first year clinical examinations

Abstract (193 words)

Evidence shows that medical students from Minority Ethnic (ME) backgrounds and male medical students underperform in undergraduate examinations. Our study confirmed these findings in 1st year clinical (Year 3) medical students, and further explored this disparity in performance.

We conducted a series of meta-analyses to measure the effects of sex and ethnic group on the written examination and Objective Structured Clinical Examination (OSCE) scores of three groups of Year 3 medical students at two London UK medical schools (n=1051; 46.0% male; 48.7% White).

Male and ME students scored lower on written and OSCE assessments. Both assessments were statistically significantly correlated (mean r=0.45) and therefore the effects of sex and ethnic group were measured on each exam after being adjusted for the effect of the other. Although sex and ethnic differences remained on the OSCE when adjusted for written performance, these differences disappeared on the written when it was adjusted for OSCE performance.

These findings may reflect a relative deficit in practical clinical knowledge in male and Minority Ethnic Year 3 students. Results were unlikely to be due to examiner bias, as the machine-marked unadjusted written exam results showed significant sex and ethnic differences.

Key words: assessment, clinical, ethnic, gender, medical education, medical students, sex, undergraduate.

Introduction

Over the last decade evidence has been accumulating that male students and students from Minority Ethnic (ME) backgrounds tend to underperform in a variety of medical school assessments relative to female students and students from White ethnic backgrounds. For example, ME and male students have been found to score lower in final clinical and written examinations (Liddell & Koritas, 2004; Koenig et al., 1998; Lumb & Vail, 2004; Xu et al., 1993) be less likely to achieve honours in a medical degree (James & Chilvers, 2001), and be more likely to fail final examinations compared to their female and White counterparts (Dillner, 1995; McManus et al., 1996; Wass et al., 2003). In Australia, students of Aboriginal and Torres Island origin were more likely to be rated "unsatisfactory" in first year medical examinations compared to students of European origin (Kay-Lambkin et al., 2002). Female medical students have been found to score higher and be awarded more distinctions in obstetrics and gynaecology assessments (Bienstock et al., 2002; Higham & Steer, 2004), as well as performing better overall in final examinations (McDonough et al., 2000).

The reasons for male and ME students' underperformance are unclear. Haq, Higham, Morris and Dacre (2005) found that Year 3 London medical students of Asian origin who were educated in the UK and used English as their first language, performed less well in Objective Structured Clinical Examinations (OSCEs) and written assessments compared to their White peers. This implies the Asian students were not underperforming because of difficulties using English or because of large differences in educational background. Wass and her colleagues (2003) suggested that poor communication skills in a subset of male ME students may have accounted for some of the differences in performance in a final year clinical communication skills examination. Neither Wass *et al.* (2003) nor McManus and his colleagues (1996) found any evidence of overt racial discrimination in the marking of final year medical assessments.

Worldwide standards of medical education state that medical students need to a) have an understanding of the scientific knowledge fundamental to acquiring and applying clinical science; and b) have sufficient clinical knowledge and skills to assume appropriate clinical responsibility upon graduation (World Federation for Medical Education, 2003). In Year 3 at the two UK medical schools which took part in this study, scientific theoretical medical knowledge (which tends to be learned mainly from books and traditional teaching methods such as lectures) is assessed using written machine-marked multiple choice questionnaires and extended matching questions; whereas practical clinical knowledge and skills (which tend to be learned from patient contact in outpatients clinics, on hospital wards and in clinical skills centres) are assessed using an OSCE.

Although written examinations and OSCEs are designed to measure different knowledge and skills, there is some overlap. A student's theoretical medical knowledge may influence how they perform in an OSCE (Wilkinson & Frampton, 2004) and as such, an OSCE score cannot be said to be purely a measure of practical clinical knowledge, but probably also incorporates a measure of theoretical medical knowledge. Therefore, although previous studies have reported that male and ME students have lower OSCE (and written exam) scores, it is uncertain whether or not this was related to those students' practical clinical knowledge and/or their theoretical medical knowledge. If the ethnic and sex inequities in student examination performance are to be addressed, it is helpful to clarify whether students who score relatively lower in written examinations and OSCEs are particularly lacking in practical clinical, or theoretical medical knowledge, or both.

This study aims to replicate previous findings of sex and ethnic differences in OSCE and written examination performance in three groups of Year 3 students from two London medical schools. Furthermore, as we have seen that it is likely that there is some overlap in the constructs that both tests measure, we wish to explore the relationship between sex,

ethnic group and performance on the part of the OSCEs which is independent of performance on the written examinations, and on the part of the written examinations which is independent of performance on the OSCEs. We hope this will help us to gain greater insight into whether male and minority ethnic underperformance is particularly in the domain of aspects of learning which are related to written examination (e.g. theoretical medical knowledge) or in those related to OSCE performance (e.g. practical clinical knowledge and skills).

Methods

We collected demographic and examination data from medical school record and self report measures for two groups of Year 3 students from the Royal Free & University College Medical School [RFUCMS1 (n=367); RFUCMS2 (n=382)] and one group from Imperial College School of Medicine (ICSM, n=337), both in London, UK.

We obtained complete written and OSCE examination data for 97.4% (1059/1087) of students from medical school records. Students without complete examination data (n=28) were excluded from the analyses. Also see Haq *et al*. (2006).

Of the students with complete examination data, self-reported sex data were available for all but one student, who was excluded. Self reported ethnic group data (using the UK 2001 Census categories) were obtained for 671 students. Students' ethnic group was categorised into White (including White, White British, White Irish, White Other) and Minority Ethnic (all other ethnic group codes, including mixed White and Other ethnic group). 380 students did not self report their ethnic group and were assigned to either White or Minority Ethnic group by two authors (KW and JH) on the basis of their name and photograph. Inter-rater agreement was good, with only three differences in decisions, all of which were resolved by discussion. Seven students did not self report ethnic group (including one for whom sex was

not known) and also did not have a photograph and were therefore excluded from the analyses. See Table 1.

Table 1 about here

Examinations

All students took an end of Year 3 OSCE and written assessments. Similar examinations are held in the two schools, with written assessments comprising four examinations every 10 weeks at RFUCMS and one end of year examination at ICSM. These examinations are machine-marked and consist of true/false multiple choice questions and/or extended matching questions. OSCEs at both schools had acceptable reliability (Cronbach's alpha: RFUCMS 2002-2003= 0.67, ICSM=0.89, RFUCMS 2003-2004 unavailable). All examination results are expressed as percentages.

Theoretical medical knowledge and practical clinical knowledge

Hereafter we use the term “theoretical medical knowledge” to indicate knowledge which tends to be learned mainly from books and traditional teaching methods such as lectures; we use the term “practical clinical knowledge” to indicate knowledge and skills which tend to be learned from patient contact in outpatients clinics, on hospital wards and in clinical skills centres.

Statistical analyses

SPSS for Windows v11 was used for conventional statistical analyses. Independent effects of sex and ethnic group were calculated using Analyses of Variance (ANOVA) and Analyses of Covariance (ANCOVA). Results from the three separate groups could not be combined into a single ANOVA for an overall effect since the examinations differed in their marking schemes and specific content, and therefore could not be equated formally, and neither

could it be assumed that their distributions were equivalent. Instead therefore, a meta-analytic approach was taken, to combine these otherwise separate studies of equivalent effects. A meta-analytic approach also had the advantage of allowing formal tests for heterogeneity of the studies. Effect sizes were calculated using the spreadsheets freely available from http://www.work-learning.com/effect_size_download.htm. Reference groups for effect sizes were females, and those in the white ethnic group. Meta-analyses we performed using MIX, a meta-analysis software programme freely available from http://www.mix-for-meta-analysis.info/.

All p values smaller than 0.05 were considered statistically significant.

Results

Sex, ethnic group and performance in written examination and OSCE per group

ANOVAs were used to calculate the mean difference between White and ME, and male and female students' scores on the written examination and on the OSCE, for each group of students, and to check for sex by ethnic group interactions. All interactions were non-significant. The *t* statistics from these ANOVAs were then used to calculate the effect sizes (Cohen's d) shown in table 2. The negative direction of the effect sizes indicates that male students scored lower than female students and ME students scored lower than White students. Cohen considers 0.15 to 0.4 to be a small effect, and 0.4 to 0.75 to be a medium effect (Thalheimer & Cook, 2002), so that these effects are small to medium.

Table 2 about here

Meta analyses a), b), c) and iv

Meta-analyses using a fixed effect model of effect sizes derived from continuous variables were used to compare and summarise findings from the three groups of students. The reference groups were female and White.

We carried out four meta-analyses to measure the effect of (i) sex and (ii) ethnic group on written exam performance; and the effect of (iii) sex and (iv) ethnic group on OCSE performance (total n=1051 for each meta-analysis). Each meta-analysis was highly statistically significant indicating a clear overall effect with males and ME students performing less well on both written and OSCE assessments (see table 3).

Due to the fact that the data were from three groups of Year 3 London medical students, two of which were consecutive years at the same medical school; and all of which were demographically similar, we did not expect the groups to be heterogeneous in their effects of sex or ethnic group. Non-significant Q tests of heterogeneity of variance confirmed this. See Table 3.

Table 3 about here

The forest plots for each meta-analysis are shown below in Figure 1.

Figure 1 about here

Correlations between written examination and OSCE results

The written examination and OSCE results were positively correlated for all groups [RFUCMS1 r=0.44; RFUCMS 2 r=0.58; ICSM r=0.37; p<0.001 in all cases]. The mean correlation across groups was r=0.45, thus the mean $r^2 = 0.22$, indicating that, across groups, just over a fifth of the total variance in the OSCE results was related to the written examination results (Howell, 2002). Thus, despite the tests measuring different constructs overall, there was a significant overlap.

Analyses of covariance

We used ANCOVA to measure the relationship between ethnic group, sex and performance on the OSCE whilst taking into account performance on the written test of medical knowledge, as we felt this may provide a clearer picture of sex and ethnic differences in practical clinical ability. Similarly, we measured the relationship between ethnic group, sex and performance on the written test whilst taking into account performance on the OSCE, as we felt this may provide a clearer picture of sex and ethnic differences in theoretical medical knowledge.

Effect sizes were calculated as before, with negative values indicate lower male and ME scores (see Table 4).

Table 4 about here

Meta-analyses e), f), g) and h)

We carried out a further four meta-analyses, using effect sizes for continuous variables, to measure the effect of (v) sex and (vi) ethnic group on written exam performance adjusted for OCSE performance; and on the effect of (vii) sex and (viii) ethnic group on OSCE performance adjusted for written exam performance (total n=1051 for each meta-analysis).

The results of meta-analyses e) and f) showed that there were no statistically significant sex ($p=0.11$) or ethnic group ($p=0.06$) differences on the written examination performance once it was adjusted for OSCE performance. By contrast, the results of meta-analyses g) and h) showed that male and Minority Ethnic students still scored statistically significantly lower on the OSCE, even once it was adjusted for written examination performance. As before, all Q tests for heterogeneity of variance were non-significant, as expected. See Table 5.

Table 5 about here

Forest plots for meta-analyses e) to h) are shown below in figure 2.

Figure 2 about here

Conclusions and Discussion

As in previous studies, this study finds that males and Minority Ethnic group members perform less well in both written examinations and OSCEs in year 3. However, although the sex and ethnic differences remained when the OSCE was adjusted to take into account its correlation with the written examination, when written results were adjusted to take into account the correlation with the OSCE, the sex and ethnic differences were no longer significant.

The moderate yet highly statistically significant correlation between the OSCE and written exam scores suggests an overlap between the constructs being measured by both assessments. This correlation is not unexpected as it is known that OSCE performance is predicted by written examination performance, probably because "good knowledge leads to good OSCE performance" (Wilkinson & Frampton, 2004; p1114). It may be equally true however that this transfer of learning can occur the other way round, with good practical clinical learning improving written performance. Indeed, medical students typically say that studying basic sciences after clinical attachments makes the science easier to understand, that they have more intuitive insights into what they are studying, and that they are more likely to elaborate on their knowledge (Coles, 1998).

The OSCE is designed to measure primarily practical clinical knowledge and skills, whereas the written exam is designed to measure primarily theoretical medical knowledge. Bearing in

mind the correlation between the two exams, it seems likely that the OSCE also measures some amount of theoretical medical knowledge and similarly, the written examination also measures some amount of practical clinical learning and skills. The remainder of the variance in each exam (i.e. the unshared systematic variance not accounted for by measurement error) is probably due to differences in practical clinical knowledge and skills for the OSCE, and to differences in theoretical medical knowledge for the written.

The fact that adjusting the written performance to take into account performance on the OSCE *does* remove the effects of sex and ethnic group, whereas the sex and ethnic differences on the OSCE do *not* disappear when the OSCE is adjusted for the written assessment, suggests that there are sex and ethnic differences in the shared variance between the OSCE and the written, as well as in the unshared variance in the OSCE, but *not* in the unshared variance in the written exam. As such, we propose that there are sex and ethnic differences on practical clinical knowledge and skills (i.e. on the shared and unshared variance in the OSCE) but not on theoretical medical knowledge (i.e. not on the unshared variance in the written). Thus our results provide some evidence that male and ME students may have a relative deficit in practical clinical knowledge compared to female and White students. Thus our results suggest that the poorer performance of male students and of ME students may result primarily from a relative deficit in their practical clinical knowledge, rather than in the more theoretical knowledge acquired from textbooks. If correct, that has some practical importance.

While it might be tempting to assume that this difference is bias on the part of OSCE examiners, this is unlikely to be the case because the *unadjusted* written examination results, which are marked by machine, also show these sex and ethnic differences. Instead, it seems likely that these results reflect genuine differences in ability.

In conclusion, this study has shown that ME students and male students underperformed compared to White students and female students in both written and clinical Year 3 assessments, which may have been due to a relative deficit in practical clinical knowledge. The amount of knowledge gained from clinical experience is related to learning style, which is in turn related to examination performance (Martin, Stark & Jolly, 2000; McManus, Richards, Winder & Sproston, 1998) and we are currently undertaking further studies to investigate whether learning styles vary by ethnic group and sex, and whether this has an impact on clinical and/or theoretical knowledge. This will be carried out in conjunction with a qualitative and quantitative research project to explore the barriers to practical clinical learning experienced by medical students.

Ethical approval

Ethical approval for this study was granted by the UCL Graduate School Ethics Board.

Acknowledgements

Thank you to Henry W.W. Potts and to anonymous reviewers for their helpful comments**.**

References

Bienstock, J.L., Martin, S., Tzou, W. and Fox, H.E. (2002). Medical students' gender is a predictor of success in the obstetrics and gynecology basic clerkship. Teaching and Learning in Medicine 14(4):240-243.

Coles, C. (1998). How students learn: the process of learning. (In Jolly, B. & Rees, L. (Eds.), Medical education in the millennium. (pp. 63-82) (Oxford: Oxford University Press).

Dillner, L. (1995) Manchester tackles failure rate of Asian students. British Medical Journal 310:209

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

Higham, J. and Steer, P.J. (2004). Gender gap in undergraduate experience and performance in obstetrics and gynaecology: analysis of clinical experience logs. British Medical Journal 328:142-143

Howell, D. C. (2002). Statistical methods for psychology. 5th Ed. (Belmont, CA: Thomson Wadsworth).

James, D. and Chilvers, C. (2001). Academic and non-academic predictors of success on the Nottingham undergraduate medical course 1970-1995. Medical Education 35:1056-1064.

Kay-Lambkin, F., Pearson, S.A., and Rolfe, I. (2002). The influence of admissions variables on first year medical school performance: a study from Newcastle University, Australia. Medical Education. 36(2):154-9.

Koenig, J.A., Sireci, S.G., and Wiley, A. (1998) Evaluating the predictive validity of MCAT scores across diverse applicant groups. Academic Medicine 73(10):1095-106.

Liddell, M.J., and Koritas, S. (2004). Effect of medical students' ethnicity on their attitudes towards consultation skills and final year examination performance. Medical Education 38:187-198.

Lumb, A., and Vail, A. (2004) Comparison of academic, application form and social factors

in predicting early performance on the medical course. Medical Education ;38:1002-1005.

Martin, I.G., Stark, P. and Jolly, B. (2000). Benefiting from clinical experience: the influence of learning style and clinical experience on performance in an undergraduate objective structured clinical examination. Medical Education 34:530-534.

McDonough, C.M., Horgan, A., Codd, M.B. and Casey, P.R. (2000). Gender differences in the results of the final medical examination at University College Dublin. Medical Education 34:30-34.

McManus, I.C., Richards, P., Winder, B. and Sprotson, K. (1996). Final Examination Performance of medical students from ethnic minorities. Medical Education 30(3):195-200.

McManus, I.C., Richards, P., Winder, B. and Sprotson, K. (1998). Clinical experience, performance in final examinations, and learning style in medical students: prospective study. British Medical Journal 316:345-350.

Thalheimer, W., and Cook, S. (2002, August). How to calculate effect sizes from published research articles: A simplified methodology. Retrieved March 12, 2007 from http://work-learning.com/effect_sizes.htm.

Wass, V., Roberts, C., Hoogenboom, R. (2003). Effect of ethnicity on performance in a final OSCE: qualitative and quantitative study. British Medical Journal 326:800-803.

Wilkinson, T.J. and Frampton, C.M. (2004). Comprehensive undergraduate medical assessments improve prediction of clinical performance. Medical Education 38: 1111-6.

World Federation for Medical Education. (2003) Basic Medical Education. WFME Global Standards for Quality Improvement.(WFME, Copenhagen).

Xu, G., Veloski, J.J., Hojat, M., Gonnella, J.S. and Bacharach, B. (1993). Longitudinal comparison of the academic performances of Asian-American and White medical students. Academic Medicine 68(1):82

Table 1: Sex and ethnic group of three groups of Year 3 students from two London medical schools.

Group	**Male**	**Female**	**Assigned ethnic group**		**Self reported ethnic group**		**Total**		**Total**
			White	**ME**	**White**	**ME**	**White**	**ME**	
RFUCMS 2002-2003	158	199	57	82	126	92	183	174	357
RFUCMS 2003-2004	163	200	108	7	62	186	170	193	363
ICSM 2003-2004	162	169	59	67	100	105	159	172	331
Total	483	568	224	156	288	383	512	539	1051

Table 2: Size of the effects (Cohen's d) of sex and ethnic group on written examination and OSCE scores for three groups of medical students (RFUCMS0203, RFUCMS0304, ICSM0304)

	Sex (male-female)		Ethnic group (Minority Ethnic-White)	
Effect sizes (Cohen's d)	Written	OSCE	Written	OSCE
RFUCMS0203	-0.17	-0.27	-0.13	-0.29
RFUCMS0304	-0.31	-0.43	-0.29	-0.35
ICSM0304	-0.30	-0.41	-0.29	-0.17

Table 3. Results of four meta-analyses assessing sex (male vs female) and ethnic (Minority Ethnic vs White) differences in performance on Year 3 written and OSCE assessments

Results of Q tests of heterogeneity of variance also shown.

	Predictor variable	Outcome variable	Q (p value)	Mean *d* (95% CI)	p value for meta-analysis
Meta-analysis a)	Sex	Written score	1.0 (0.61)	-0.26 (-0.38 to -0.14)	<0.0001
Meta-analysis b)	Ethnic group	Written score	1.5 (0.45)	-0.24 (-0.36 to -0.11)	0.0001
Meta-analysis c)	Sex	OSCE score	1.3 (0.43)	-0.37 (-0.49 to -0.25)	<0.0001
Meta-analysis d)	Ethnic group	OSCE score	1.3 (0.52)	-0.27 (-0.39 to -0.15)	<0.0001

Table 4: Effect sizes (Cohen's *d*) for sex and ethnic group on written examination performance adjusted for OSCE performance, and OSCE performance adjusted for written exam performance, for three groups of medical students (RFUCMS0203, RFUCMS0304, ICSM).

	Sex (male-female)		Ethnic group (Minority Ethnic-White)	
Effect sizes (Cohen's d)	Written adjusted for OSCE	OSCE adjusted for written	Written adjusted for OSCE	OSCE adjusted for written
RFUCMS0203	-0.06	-0.22	-0.01	-0.26
RFUCMS0304	-0.09	-0.30	-0.12	-0.22
ICSM	-0.16	-0.32	-0.25	-0.07

Table 5: Results of four meta-analyses measuring sex (male vs. female) and ethnic (Minority ethnic vs. White) differences in performance on Year 3 written examination adjusted for OSCE performance; and on Year 3 OSCE performance adjusted for written examination performance. Results of Q tests of heterogeneity of variance also shown.

	Predictor variable	Outcome variable	Covariate	Q (p value)	Mean *d* (95% CI)	p value for meta-analysis
Meta-analysis e)	Sex	Written score	OSCE score	0.46 (0.80)	-0.10 (-0.22 to 0.02)	0.11
Meta-analysis f)	Ethnic group	Written score	OSCE score	2.55 (0.28)	-0.12 (-0.24 to 0.00)	0.06
Meta-analysis g)	Sex	OSCE score	Written score	0.54 (0.76)	-0.28 (-0.40 to -0.16)	<0.0001
Meta-analysis h)	Ethnic group	OSCE score	Written score	1.63 (0.44)	-0.19 (-0.31 to -0.07)	0.002

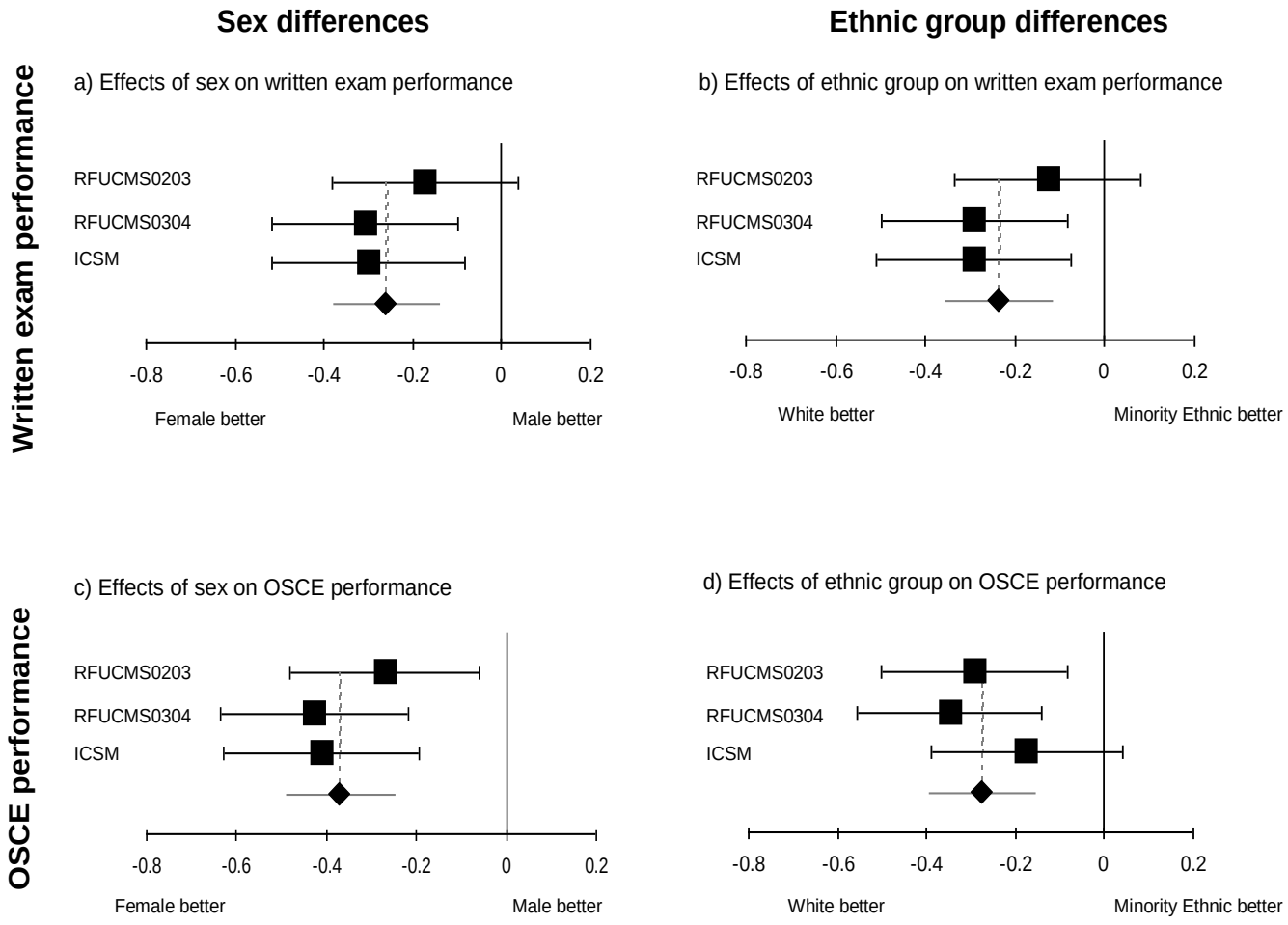

Figure 1: Forest plots for meta-analyses of a) effects of sex on written exam performance, b) effects of ethnic group on written exam performance, c) effects of sex on OSCE performance, and d) effects of ethnic group on OSCE performance. The result for each meta-analysis is shown as a black diamond, with 95% confidence intervals shown in grey. The black squares with black confidence intervals represent each group's effect.

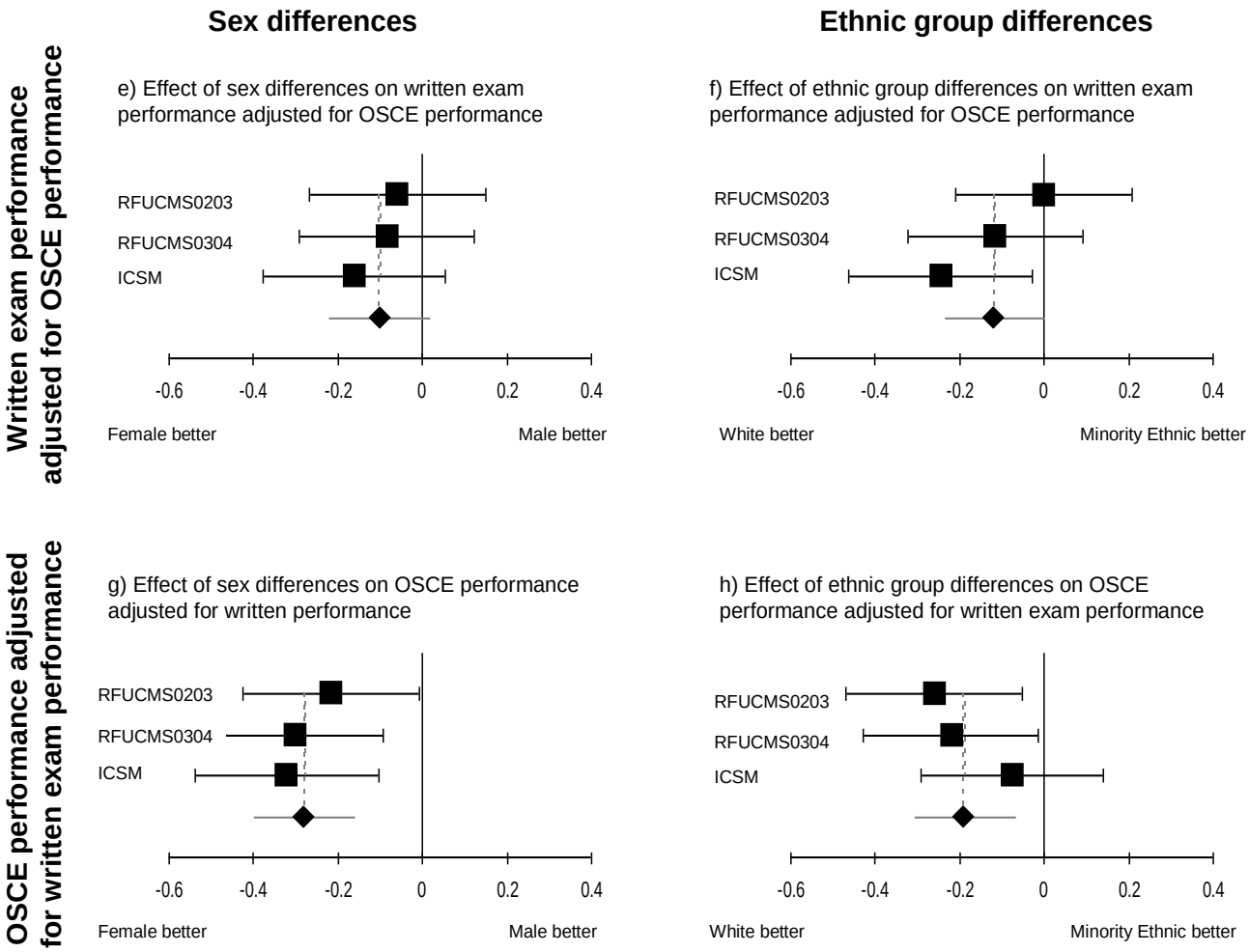

Figure 2: Forest plots for meta-analyses of e) effects of sex on written exam performance adjusted for OSCE performance, f) effects of ethnic group on written exam performance adjusted for OSCE performance g) effects of sex on OSCE performance adjusted for written exam performance, h) effects of ethnic group on OSCE performance adjusted for written exam performance. The result for each meta-analysis is shown as a black diamond, with 95% confidence intervals shown in grey. The black squares with black confidence intervals shown each individual group's effect.