Does Patient Gender Influence Medical Care Among Students? Evaluation During an Objective Structured Clinical Examination

Joana Le Boudec (joana.le-boudec@unisante.ch)
Unisanté https://orcid.org/0000-0001-7363-1821

Sylvie Félix
Universite de Lausanne Faculte de biologie et medecine

David Gachoud
Universite de Lausanne Faculte de biologie et medecine

Matteo Monti
Universite de Lausanne Faculte de biologie et medecine

Mirjam Schuler Barazzoni
Universite de Lausanne Faculte de biologie et medecine

Carole Clair
Universite de Lausanne Faculte de biologie et medecine

Research article

Keywords: gender, students, objective structured clinical examination, diagnosis, management, stereotypes, bias

Posted Date: October 14th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-61600/v1

Abstract

Background: Gender bias induces gender inequality in health. In this study, we evaluated gender bias during a local objective structured clinical examination (OSCE).

Methods: We assessed gender bias by using two clinical cases—generalized anxiety disorder (GAD) and ascending aorta dissection (AAD)—during an OSCE performed among fifth-year medical students. For each situation, half of the students encountered male and half encountered female standardized patients (SPs). Except for gender, variables were identical in each clinical case. Patients, students, and examiners were blinded to the purpose of the study. Medical history, clinical examination, diagnosis, and management were compared between male and female SPs. The interaction between student and SP gender was analyzed.

Results: A total of 110 medical students were observed (55% women). For GAD, students arrived at the correct diagnosis more often for female SPs than for male SPs (diagnosis completed, partially completed, and not completed in 47%, 16%, and 36%, respectively, of female SPs vs. 22%, 20%, and 58%, respectively, of male SPs, $p = 0.02$). The nature of their symptoms was more often asked of male SPs (completed, partially completed, and not completed in 51%, 4%, and 0% of male SPs, respectively, vs. 38%, 17%, and 0% of female SPs, respectively, $p = 0.002$), and associated physical symptoms were more often explored in female SPs (completed and not completed in 84% and 16% of female SPs, respectively, vs. 65% and 35% of male SPs, respectively, $p = 0.03$). For AAD, an emergency was better identified in female SPs (95% identified in female SPs vs. 76% in male SPs, $p = 0.005$) and examination of femoral pulses was more often performed in female SPs (88% completed in female SPs vs. 54% in male SPs, $p < 0.0001$). The interaction between SP and student gender was not significant.

Conclusion: The gender bias observed supports the need to address unconscious biases and to raise student awareness of gender stereotypes likely to lead to underrecognition or subtreatment of disease in patients of both genders.

Background

One of the principal responsibilities of medical schools is to train physicians to become proficient in clinical reasoning (1) and clinical decision making. Clinical reasoning is the cognitive process that ultimately leads to an action (diagnosis and therapy). This complex process involves medical knowledge, as well as emotional, cultural, institutional, and human influences, and can lead to possible personal partialities and bias in clinical decision making. Biases of the cognitive type have been well described (2), and the propensity to run into such bias can be reduced (3, 4). Different strategies have been proposed to reduce cognitive biases in clinical reasoning that apply at the individual level, the level of the work organization, or the level of medical schools (5). Bias first has to be identified. Among additional biases influencing medical decision making, two types (6) have been conceptualized and identified as contributing to disparities in health between women and men: not taking into account medical gender

specificities when present (e.g., ignoring differences in the presentation of ischemic heart disease between men and women (7)) and gender stereotypes, that is, treating or considering men and women differently when it is not justified (e.g., neglecting to identify depression as a possible diagnosis in men (8)). Although gender bias is more and more recognized in medical research (9, 10), little is yet known about its influence in medical practice and education. Training in gender medicine aims to improve gender knowledge and skills of medical students. There is a need to identify gender bias in clinical practice to better address the challenge of moving from gender bias attitudes toward gender awareness among young doctors (4). Previous studies have examined the interaction of the medical candidate and the patient's gender from the perspective of validating the fairness of clinical examinations. But not yet clearly shown is whether being a male or female patient can influence the student's evaluation (11).

At Lausanne University Medical School, clinical and communication competencies are regularly assessed through objective structured clinical examinations (OSCEs). OSCEs have already been used to assess gender bias (12). They offer a way to investigate bias among advanced students by having them face identical clinical situations with identical clinical signs and symptoms in standardized patients (SPs).

Methods

Aim

In this study, we investigated whether patient gender influences student management in terms of communication, clinical assessment, investigations, diagnosis, and proposed care.

Study design and setting

Medical school in Switzerland consists of 6 years divided into 3 years of a bachelor's program and 3 years of a master's program. At Lausanne University Medical School, students undergo a summative assessment of clinical competencies through the summative OSCE at the end of their fifth study year, just before entering 10 months of internship. During the OSCE, history taking, physical examination, and diagnostic and therapeutic management are assessed through station-specific checklists. Communication competencies are assessed at every station by using the same global rating scale derived from the Analytic Global OSCE Rating developed by Hodges and McIlroy (13), which consists of four items that explore the ability of students to respond to the patient's needs, the quality of the structure of the interview, and verbal and nonverbal expression. In the spring of 2017, fifth-year medical students underwent an OSCE that included five stations, each lasting 13 minutes. The OSCE took place over 2 days: half of the students were assessed on day 1, the remainder on day 2. On day 1, students were exposed to two different gender-sensitive OSCE stations. Half of the students encountered a female standardized patient (SP), and the other half a male SP. Apart from gender, all other patient characteristics (habits; medical, social, and family history) were similar.

Students, SPs, and examiners were all blinded to the purpose of the study.

Study population

Half of the fifth-year medical students (n = 110) at Lausanne University Medical School, who were taking their OSCE on day 1 of the OSCE session in March 2017, took part in the study. Evaluation was made by eight examiners for the first vignette, two women and six men, and by four examiners for the second vignette, all men. According to the standard setting of our OSCE, all examiners directly observed the student-SP encounter and filled out the evaluation form during the encounter.

Case vignettes and their evaluation

Vignette 1: Generalized anxiety disorder. The case was summarized as follows: *"A middle-aged patient presents to her or his family doctor's practice with generalized anxiety. The manifestations are an anxious feeling, trouble sleeping, agitation, and brooding over a hypothetical drama occurring in her or his family. The patient also mentions thoracic oppression when asked, and spontaneously presents anxiety about his/her financial situation without objective reasons. She or he seems tense, is talkative, and is somewhat focalized on her or his concerns, playing with the pen and often changing position."* Students were asked to take a medical history, propose investigations, arrive at a diagnosis, refer the patient, and/or reschedule the patient for a follow-up appointment (students were not asked to perform a clinical examination). Students were evaluated on the following: medical history (13 items: asking about the nature [characteristics] of symptoms; beginning, trigger, and evolution of symptoms; first episode or not; associated symptoms; symptoms of depression; suicidal thoughts; hallucinations; quality of sleep; substance abuse; helping strategies; previous medical history), arriving at a diagnosis, proposing further investigations and follow-up, and communication. Items of communication were rated on a 5-point scale from "completely" (5 points) to "not at all" (1 point). All other items were considered to be "completed" (2 points), "partially completed" (1 point), or "not completed" (0 points), with certain items of medical history only being "completed" (2 points) or "not completed" (0 points).

Vignette 2: Ascending aorta dissection: The case was described as follows: *"An older patient of tall and slim build presents with acute chest pain. The pain is described as violent, extending to the shoulder blades, having started 2 hours earlier. The patient has a cold feeling in his/her left foot. The patient is known for untreated hypertension. On examination, the patient is ill looking, has higher blood pressure on the right arm than on the left, and a diastolic murmur over the aortic area. The pulse is obliterated on the left leg."* Students are expected to recognize the vascular origin of pain and arrive at a diagnosis of type A aortic dissection. A secondary objective was to fulfill an appropriate clinical examination. Students were evaluated on the following: medical history (nine items: trigger, nature of pain, evolution of pain, dyspnea, accompanying symptoms, cardiovascular risk factors, previous medical history, medication, family medical history), clinical examination (five items: measurement of blood pressure, blood pressure on both arms, cardiac auscultation, femoral pulses, radial pulses), arriving at a diagnosis, proposing a computed tomography (CT) scan to confirm the diagnosis, identifying the emergency, proposing analgesia, and communication skills. Points were attributed in the same manner as for Vignette 1.

The vignettes are not available for publication due to confidentiality issues.

Statistical analysis

Single checklist item scores of the students who evaluated female versus male SPs were compared by using chi-squared tests for each item.

We also created scores to summarize the numerous points of medical history and physical examination (only in Vignette 2). Scores were created as the sum of all points through all items of one category of medical history or physical examination (only in Vignette 2). For Vignette 1, the global score for medical history summed to a maximum of 30 points per student (minimum 0 points). For Vignette 2, the score for medical history resulted in a maximum of 20 points, and the score for physical examination resulted in a maximum of 11 points. The scores between students encountering female SPs and those encountering male SPs were summed and the total results compared by using the Wilcoxon-Mann-Whitney test.

In addition, results were analyzed by using linear regressions, with scores as the dependent variable and student gender, SP gender, and the interaction between these two variables as independent variables in order to assess whether there was an interaction between SP and student gender. The interaction with the examiners' gender was not analyzed, in consideration that there were few examiners (four per vignette) and only male examiners evaluated the second vignette.

The STATA 14 statistical package was used for statistical analysis. A P value of < 0.05 was considered statistically significant.

Results

A total of 110 students participated in the OSCE on day 1 and were included in the analyses; 60 (55%) were women and 50 (45%) were men (Table 1). They were all exposed to the two clinical scenarios played by eight different SPs, four women and four men.

Table 1
Study population

	Total	Women	Men
Clinical case 1: Anxiety			
Students, n (%)	110	60	40
Standardized patients	8	4	4
Examiners	8	2	6
Clinical case 2: Ascending aorta dissection			
Students	110	60	40
Standardized patients	8	4	4
Examiners	4	0	4

Vignette 1: Generalized anxiety disorder

The main finding was that students were more likely to arrive at the correct diagnosis when encountering female SPs than they were when encountering male SPs. The diagnosis was completed, partially completed, and not completed in 47%, 16%, and 36% of the cases, respectively, when the SP was a woman versus 22%, 20%, and 58%, respectively, when the SP was a man ($p = 0.02$) (Fig. 1). The total score for medical history was not statistically different between female and male SPs (Fig. 2). Regarding specific items of medical history, students better characterized the nature of symptoms when encountering male SPs (completed, partially completed, and not completed in 51%, 4%, and 0% of men, respectively, vs. 38%, 17%, and 0% of women, respectively, $p = 0.002$), whereas they more often asked female SPs about associated physical symptoms (completed and not completed in 84% and 16% of women, respectively, vs. 65% and 35% of men, $p = 0.03$) (Additional file 1). Diagnostic workup and the decision for a referral or a follow-up appointment were similar for both men and women (Fig. 1). Assessment of communication did not differ between the evaluation of men and women (Additional file 2).

The interaction between student and SP gender regarding diagnosis, nature of the symptoms, and associated physical symptoms was not significant (data not shown).

Vignette 2: Ascending aorta dissection

We found that the emergent character of the situation was better identified by students in contact with female SPs than it was by those in contact with male SPs, with 95% of students identifying the emergency of the situation in women versus 76% in men ($p = 0.005$) (Fig. 3). The total score for medical history was not statistically different when we compared students in contact with male and with female SPs (Fig. 4), but pain characteristics and family medical history were better explored in female SPs (pain

characteristics completed, partially completed, and not completed in 75%, 25%, and 0% of women, respectively, vs. 22%, 29%, and 3%, respectively, of men, $p = 0.001$; family medical history completed in 71% of women vs. 52% of men, $p = 0.04$) (Additional file 3). The total score for physical examination was not significantly different (Fig. 4), but femoral pulses were more often taken in female SPs (88% completed in female SPs vs. 54% in male SPs, $p < 0.0001$) (Additional file 4). Frequency of correct diagnosis, CT scan proposal, and analgesia were not different (Fig. 3). Assessment of communication also did not differ (Additional file 5).

The interaction between student and SP gender as a function of the understanding of the emergency of the situation, total status score, femoral pulses, pain characterization, family medical history, and communication score was not significant (data not shown).

Discussion

This study shows the presence of gender bias to some extent among senior medical students encountering female and male SPs during the OSCE.

Identified bias

Students were more likely to diagnose GAD in female than in male SPs. This occurred even though the nature of the symptoms was better investigated in male SPs and the physical symptoms were more discussed for female SPs. This is a well-described bias with psychiatric illness in general, and anxiety in particular, being more prevalent and more easily accepted in women and thus more likely to be attributed to women (14, 15). Students in this study were not an exception to this observation. Other studies have similarly shown that diagnoses were more accurate in the population groups with a higher prevalence of disease. For example, one study, using a clinical vignette, demonstrated that family physicians were more likely to diagnose chronic obstructive pulmonary disease in male patients than in female patients (16).

AAD was recognized as an emergency more easily in female SPs, family history taking and pain characterization was performed better in them, and female SPs were better examined; they were especially more likely to have their femoral pulses tested. These findings were surprising, as we expected aortic dissection to be considered more of a masculine disease. Our findings might have different explanations, but they remain hypothetical, as no causal factors were investigated in this study. These results can be considered congruent with the proven reality that an aortic pathological condition is associated with higher mortality in women, despite the fact that it is more common in men (17). In this case, therefore, this condition can be an appropriate consideration and not gender bias.

Regarding the difference in the examination of inguinal folds, this might be explained by the fact that this anatomical region is considered more intimate in men compared with that in women. Indeed, gender bias in physical examination has already been described, for example, by colleagues investigating cardiac examination, which was less precise in women because of the location of the breasts in the mitral area

(18). In our study, the inguinal region could have been considered more intimate in men than in women because of their external genitalia.

This study showed objective differences in the clinical management of male and female SPs with the same diagnoses. Understanding why students had different attitudes toward women and men is beyond the scope of this study and the variables collected. It would be interesting, however, to explore these reasons in future studies, for example, by using qualitative research.

Absence of bias

Students did not present gender bias in the management of patients (decision making to perform a complementary examination or to refer the patient to a specialist for follow-up), unlike what has been shown in various medical conditions, for example, in a real-life setting that measured the investigation of stable angina in women versus men (19). Indeed, bias tends to appear when decisions are taken under pressure, stress, or emergency, as shown, for example, by colleagues making a cardiovascular diagnosis (20). Such stress was minimized in our artificial and standardized setting. Nevertheless, another vignette-based study conducted among 503 practicing cardiologists in the United States (2) showed a bias toward ordering more angiographies in men than in women, despite the fact that it was not a real clinical situation. The authors conducted an online survey that measured implicit bias (Implicit Association Test) regarding strength and risk taking and also presented a clinical situation of a patient with intermediate risk of coronary artery disease. Cardiologists rated the likelihood of coronary artery disease as being similar in women and men, but the usefulness of stress testing was more often rated as high in women and the usefulness of angiography more often as high in men. Interestingly, cardiologists with a higher implicit bias in risk taking rated the necessity for angiography in men as higher even when it was adjusted for perceived likelihood of coronary artery disease.

Our students were probably low in bias overall. The literature shows little stereotyping in students entering medical school and gender awareness seems to improve with education and age (21, 22).

Strength and limitations

Our study has some limitations. The clinical vignette setting with SPs is a good approximation of the reality experienced by patients and doctors in many ways, but it is still an exam and thus an artificial situation. Medical history and examination were probably performed more systematically and thoroughly compared with what would be done in a real context of emergency or primary care settings. Bias appears in situations in which patients are under stress (minimized in this trial by simulation), as shown, for example, by colleagues in cardiovascular situations (20). This lack of stress might have led to an underestimation of real bias in our study. Furthermore, we examined single OSCE stations. There may be unmeasured confounders interfering with gender, such as examiner idiosyncrasies, simple memory lapses etc. The results cannot be generalized to OSCEs in general.

The small number of participants is another limitation and might explain why we did not find that student gender modified the effect of patient gender on the management of clinical situations. A study with more participants would be needed to assess not only the interaction between student and patient gender, but also the effect of the examiners' gender.

Another limitation is the OSCE setting. Although OSCE scenarios are standardized, we cannot completely exclude the possibility of some gender-specific differences in role-playing. Nevertheless, this problem is minimized because training of SPs is standardized and performed by the same SP trainer; the quality of role-playing is also verified during the examination by the SP trainer. The performance ability of the SP may also influence the results, even though the fact that there were multiple SPs helps preventing from interpreting a specific personality trait as an effect of gender.

Our study has several strengths. It was set in standardized conditions, allowing bias to be measured in relation to gender because of the standardization of social determinants such as age, socio-economic status, and ethnicity. Despite the minimization of bias in the vignette setting, having SPs act as clinical cases is also closer to real interaction than is the use of written clinical vignettes. Furthermore, students, SPs, and examiners were all blinded to the purpose of the study, which is particularly important when studying unconscious bias.

Conclusion

As previously demonstrated for race, ethnicity, and social class (23), clinical vignettes, especially with SPs, are a useful way to assess bias in health care. Our study showed new evidence of gender bias among medical students in the evaluation of clinical situations of GAD and AAD, but this biased evaluation did not seem to influence the final management (treatment and referrals) of patients. Surprisingly, biases were in favor of women in both cases. The gender bias observed is likely to lead to underrecognition (e.g., of generalized anxiety disorder in men) or subtreatment of disease (in both women and men). Our study supports the necessity of addressing unconscious biases to raise student awareness of gender stereotypes in medicine.

Abbreviations

AAD	ascending aorta dissection
CT	computed tomography
GAD	generalized anxiety disorder
OSCE	objective structured clinical examination
SP	standardized patients

Declarations

Ethics approval and consent to participate

The study was submitted to the Cantonal Ethics Committee for Research on Human Beings (CER-VD) for ethical approval (Req-2017-00335). Ethical approval was not required (decision on May 23, 2017).

Consent for publication

Not applicable

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request. Restrictions apply to the case vignettes described thoroughly in the text, which are only available after permission of the Faculty of Biology and Medicine of the University of Lausanne.

Competing interests

The authors have no conflict of interest to declare.

Funding

The authors are remunerated by their departments (Department of Training, Research and Innovation, Center for Primary Care and Public Health University of Lausanne, as well as the Clinical Skills Unit and Medical Education Unit, Faculty of Biology and Medicine, University of Lausanne, Switzerland) and did not receive any industry funding for this study.

Authors' contributions

Conception and design: CC, DG, MM, SF, MSB; acquisition of data: CC; analysis and interpretation of data: JLB, CC; drafting and critical revision: JLB, CC, DG, MM, SF, MSB; final approval: all authors.

Acknowledgements

We would like to thank Ariane Christen, Francine Viret, Josiane Boegli, Emilie Fasel, and Raphael Bonvin for their contribution to this study.

References

1. Holmboe ES, Durning SJ. Assessing clinical reasoning: moving from in vitro to in vivo. Diagnosis (Berl). 2014;1(1):111–7.
2. Daugherty SL, Blair IV, Havranek EP, Furniss A, Dickinson LM, Karimkhani E, et al. Implicit Gender Bias and the Use of Cardiovascular Tests Among Cardiologists. J Am Heart Assoc. 2017;6(12).
3. Hernandez R. Medical Students' Implicit Bias and the Communication of Norms in Medical Education. Teach Learn Med. 2018;30(1):112–7.

4. Verdonk P, Benschop YW, de Haes HC, Lagro-Janssen TL. From gender bias to gender awareness in medical education. Adv Health Sci Educ Theory Pract. 2009;14(1):135–52.
5. Singh H, Graber ML. Improving Diagnosis in Health Care–The Next Imperative for Patient Safety. N Engl J Med. 2015;373(26):2493–5.
6. Ruiz MT, Verbrugge LM. A two way view of gender bias in medicine. J Epidemiol Community Health. 1997;51(2):106–9.
7. Aggarwal NR, Patel HN, Mehta LS, Sanghani RM, Lundberg GP, Lewis SJ, et al. Sex Differences in Ischemic Heart Disease: Advances, Obstacles, and Next Steps. Circ Cardiovasc Qual Outcomes. 2018;11(2):e004437.
8. Gender differences in. depression detection: A comparison of clinician diagnosis and standardized assessment [press release]. US: American Psychological Association; 1991.
9. Holdcroft A. Gender bias in research: how does it affect evidence based medicine? J R Soc Med. 2007;100(1):2–3.
10. Regitz-Zagrosek V. Sex and gender differences in health. Science & Society Series on Sex and Science. EMBO Rep. 2012;13(7):596–603.
11. Chambers KA, Boulet JR, Furman GE. Are interpersonal skills ratings influenced by gender in a clinical skills assessment using standardized patients? Adv Health Sci Educ Theory Pract. 2001;6(3):231–41.
12. Schleicher I, Leitner K, Juenger J, Moeltner A, Ruesseler M, Bender B, et al. Examiner effect on the objective structured clinical exam - a study at five medical schools. BMC Med Educ. 2017;17(1):71.
13. Hodges B. OSCE! Variations on a theme by Harden. Med Educ. 2003;37(12):1134–40.
14. Seedat S, Scott KM, Angermeyer MC, Berglund P, Bromet EJ, Brugha TS, et al. Cross-national associations between gender and mental disorders in the World Health Organization World Mental Health Surveys. Arch Gen Psychiatry. 2009;66(7):785–95.
15. Pigott TA. Gender differences in the epidemiology and treatment of anxiety disorders. J Clin Psychiatry. 1999;60(Suppl 18):4–15.
16. Delgado A, Saletti-Cuesta L, Lopez-Fernandez LA, Gil-Garrido N. Luna Del Castillo Jde D. Gender inequalities in COPD decision-making in primary care. Respir Med. 2016;114:91–6.
17. Hultgren R. Abdominal aortic aneurysms-gender aspects on prevalence, treatment, and concurrent aneurysms. Thorac Cardiovasc Surg. 2013;61(1):15–21.
18. Chakkalakal RJ, Higgins SM, Bernstein LB, Lundberg KL, Wu V, Green J, et al. Does patient gender impact resident physicians' approach to the cardiac exam? J Gen Intern Med. 2013;28(4):561–6.
19. Daly C, Clemens F, Lopez Sendon JL, Tavazzi L, Boersma E, Danchin N, et al. Gender differences in the management and clinical outcome of stable angina. Circulation. 2006;113(4):490–8.
20. Chiaramonte GR, Friend R. Medical students' and residents' gender bias in the diagnosis, treatment, and interpretation of coronary heart disease symptoms. Health Psychol. 2006;25(3):255–66.

21. Andersson J, Verdonk P, Johansson EE, Lagro-Janssen T, Hamberg K. Comparing gender awareness in Dutch and Swedish first-year medical students–results from a questionaire. BMC medical education. 2012;12:3.
22. Rrustemi I, Locatelli I, Schwarz J, Lagro-Janssen T, Fauvel A, Clair C. Gender awareness among medical students in a Swiss University. BMC medical education. 2020;20(1):156.
23. Haider AH, Sexton J, Sriram N, Cooper LA, Efron DT, Swoboda S, et al. Association of unconscious race and social class bias with vignette-based clinical assessments by medical students. JAMA. 2011;306(9):942–51.

Figures

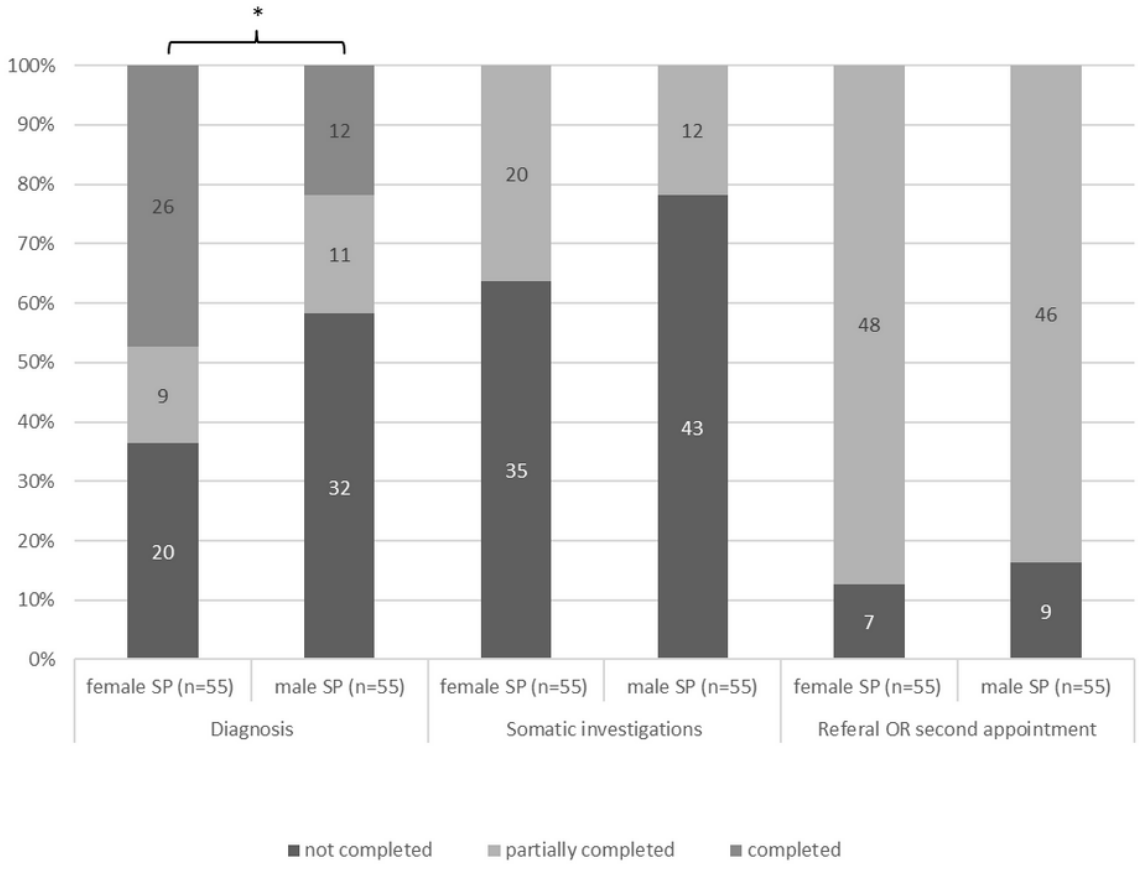

Figure 1

Vignette 1: Management (diagnosis, physical investigations, referral or second appointment) by standardized patient (SP) gender. *p < 0.05.

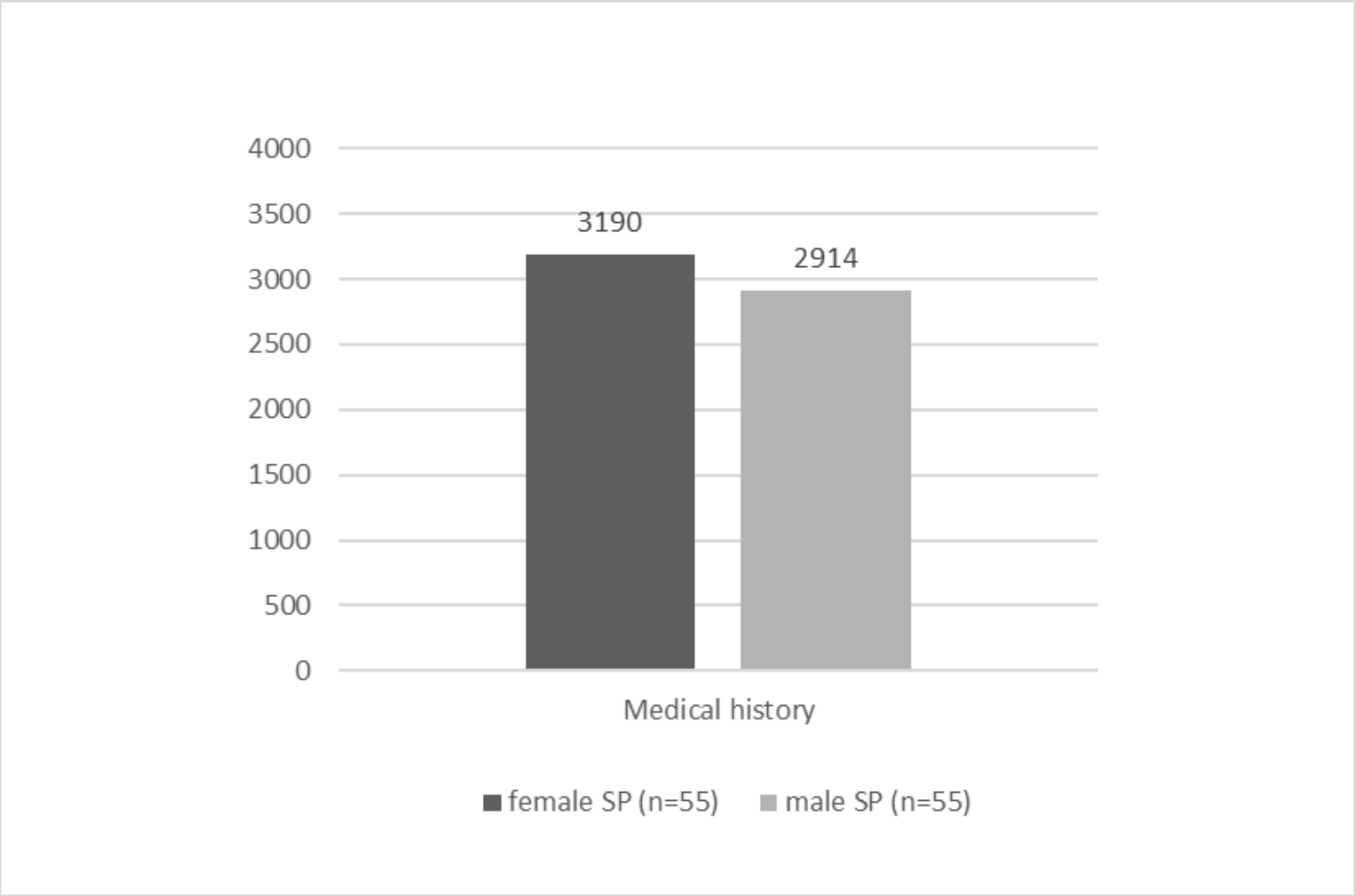

Figure 2

Vignette 1: Total score for medical history by standardized patient (SP) gender.

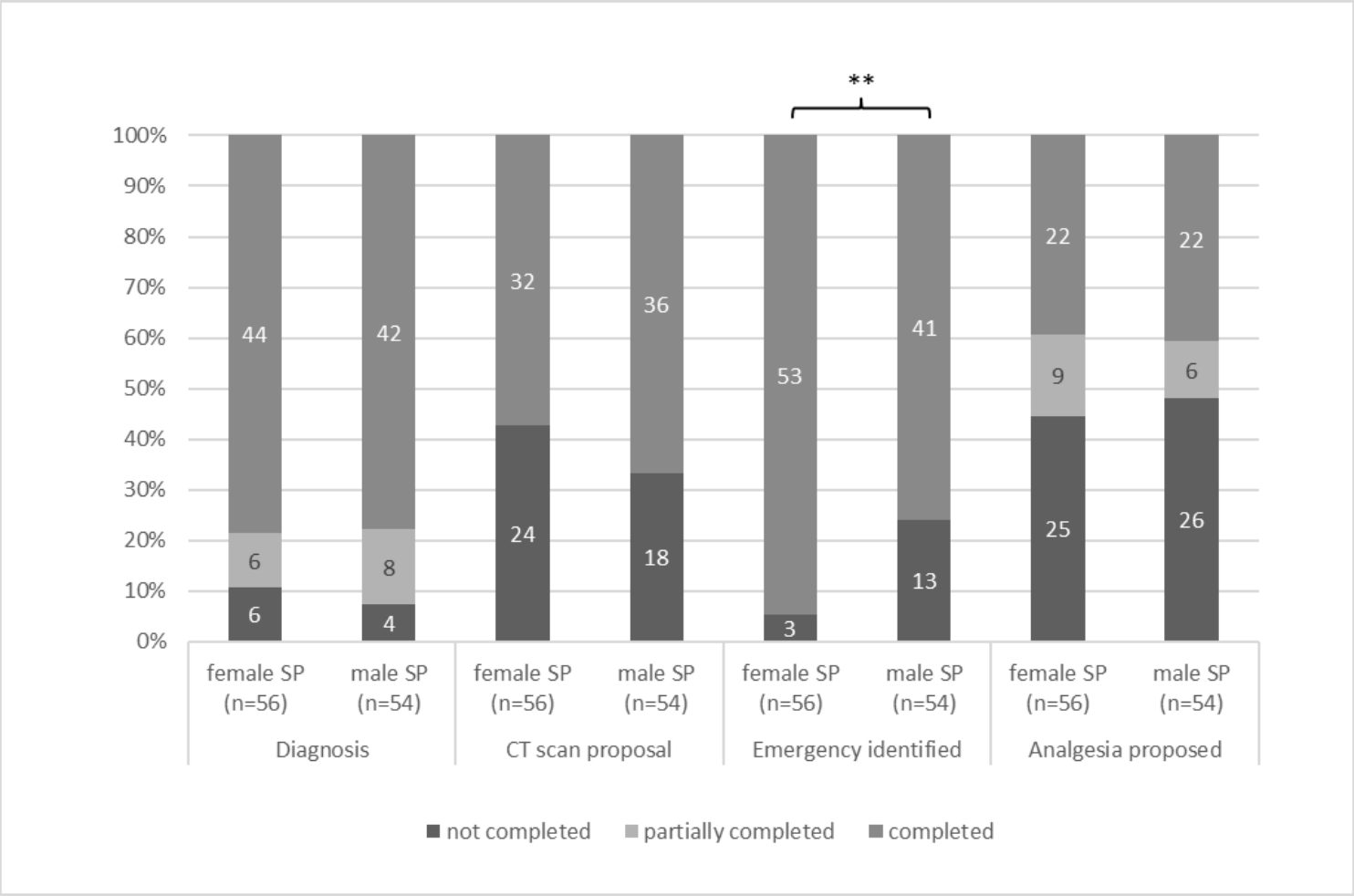

Figure 3

Vignette 2: Management (diagnosis, CT scan proposal, emergency identification, analgesia) by standardized patient (SP) gender. CT: computed tomography. **p < 0.01.

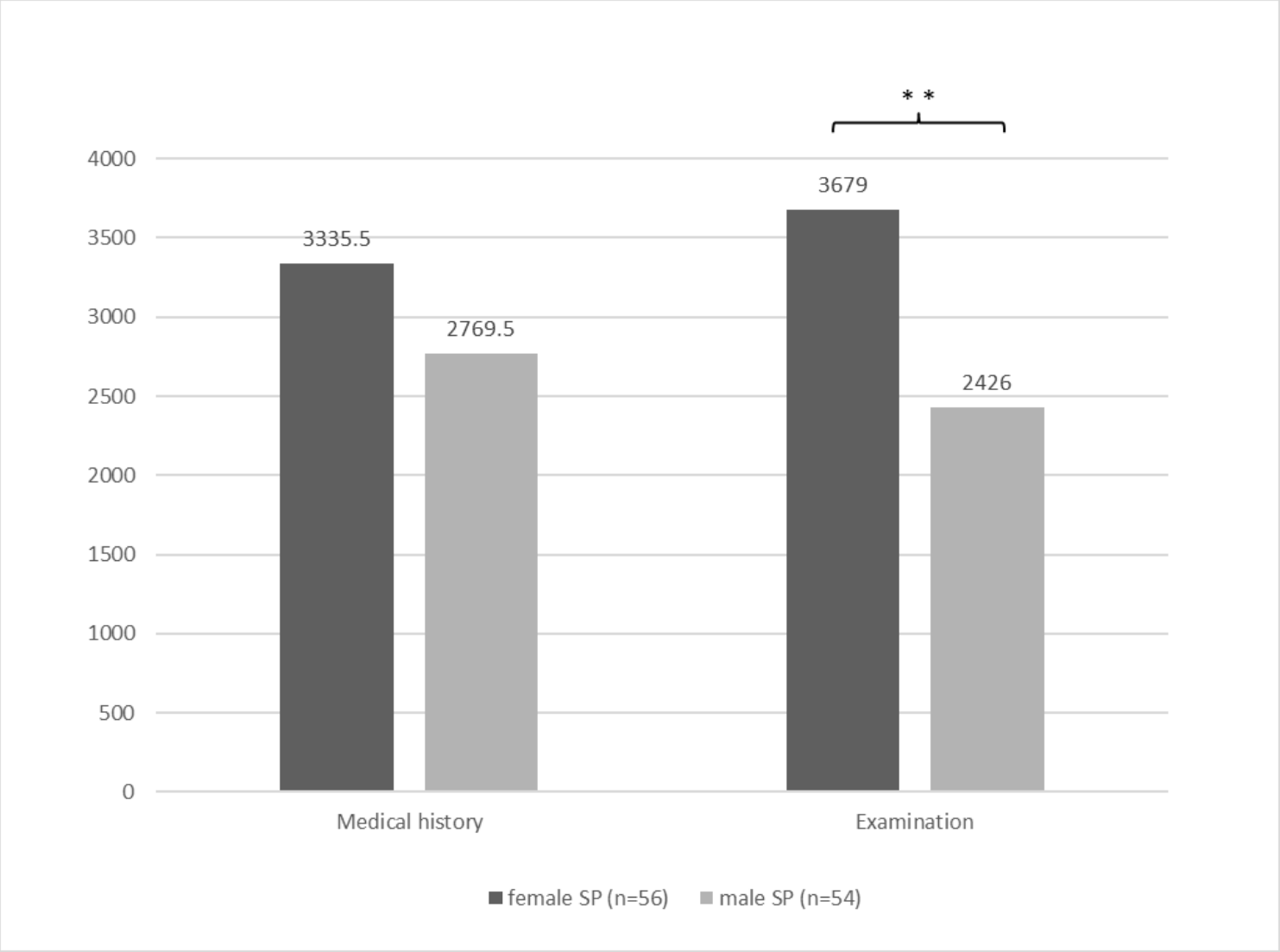

Figure 4

Vignette 2: Total scores for medical history and physical examination by standardized patient (SP) gender. **p < 0.01.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additional.docx