TITLE: *Sex and gender biases in health: using clinical simulations to uncover them*

ABSTRACT

Sex and gender inequalities in health have traditionally been approached by biological, methodological and social explanations. However, less attention has been given to the role of health systems in contributing to and generating these differences. Clearly, to the extent that the professionals and researchers working in health systems hold gender and sex biases, part of the documented sex and gender health inequalities might be attributable to them. Ultimately, the channels through which these biases might be translated into observable inequalities take the form of, among others, delays in diagnoses, tests, screening and treatment, misdiagnoses, failed continuity of care, no follow-up procedures, etc., all of which will have an important and lasting impact on people's health outcomes and long-term care needs.

Our study uses clinical simulations performed by medicine students and residency doctors to identify and quantify gender and sex biases in the different tasks involved in the clinical management of a patient: assessment of symptoms; monitoring; tests and diagnostic procedures; initial and main diagnosis; treatment administration; assessment of patient progress, response, and final status; and communication with the patient. The stage of the academic and professional trajectory in which these biases originate and their evolution over the years of studying and training are also be analysed and identified, filling an important gap in the literature.

Author: Cirstina Pujol Martínez

Universitat de Barcelona

1. INTRODUCTION

1.1. Theoretical framework

Gender and sex inequalities in health have traditionally been addressed through biological, methodological, and social explanations (1). While biological factors explain health differences between men and women through the role of sex hormones and other differences in physiological systems, methodological explanations focus on differences in access to and use of health services: women generally tend to seek more medical advice, report health problems when appropriate, and seek more differentiated and specialized help than men. Social explanations, on the other hand, emphasize the social determinants of health, namely the social and material conditions and circumstances such as education, socioeconomic status, the labour market, domestic work and caregiving, or healthy habits, which impact people's health conditions and access to health services (1).

These three approaches to gender and sex inequalities in health have produced evidence focusing on how gender differences in health stem from gender inequalities in access to social protection resources or significant differences in exposure to important health risk factors, such as unhealthy behaviours, occupational risks, or inequalities in the distribution of domestic and caregiving burdens (1). These explanations are important, relevant, and pertinent for identifying sources and causes of health inequalities. However, there are significant gaps in the literature regarding the role that health systems and the professionals working within them may play in contributing to increasing or reducing these differences. It is logical to assume that part of the inequalities and differences in health conditions between men and women could be attributable to biases present in health professionals, their education and their training. These biases embedded in health systems can translate into observable inequalities in the form of errors and inaccuracies in the different phases of the clinical management of patients who access health systems: delays or errors in diagnosis, medical tests and treatment administration; discontinuities in the care process; lack and errors in follow-up procedures, etc. All of this might have a significant and lasting impact on people's health conditions and long-term care needs.

Sex and gender differences in health conditions are natural and common: many biological, behavioural, and exposure factors are naturally responsible for the distinct susceptibility and consequences that various diseases have for men and women(2). Although we could consider these differences inevitable, in the sense that they may occur naturally due to biological disparities, understanding how these differences interact and affect diseases is essential for health systems and the professionals working in them to provide optimal and equitable healthcare through appropriate clinical management. Nonetheless, we are still far from this differentiated and specialised form of health care to take place.

In fact, a 2019 study by Nature Communications provides clear evidence of how gender and sex inequalities in health are sustained over time, become chronic, and affect a wide range of diseases and conditions (more than 700), not just those traditionally more vulnerable and exposed to gender biases (3). A systematic review by Piccardi et al. (2018) on patient safety in primary care found that women “are more likely to receive inappropriate diagnosis, treatment, or referrals” and “are not offered the same diagnostic and therapeutic treatment” in comparison to men (4). This is also true for other levels of care: different studies on diverse health conditions show that women are more likely to receive suboptimal care regarding diagnosis, screening, continuity of care, treatment or suffer more drug-adverse reactions and adverse patient safety events (5–10).This logic

also applies to the phenomenon observed in conditions already known, for decades now, to present significant different presentations and prognosis depending on whether they affect men or women, and which are also the leading cause of mortality in women worldwide: cardiovascular diseases, despite having concentrated much of the research resources in this field, continue to be understudied, underrecognized, misdiagnosed, and mistreated, especially when women suffer from them (11). It is logical to think, therefore, that there must be some factors that explain these evident differences beyond possible biological explanations for the different presentations and effects that certain diseases have on patients depending on their sex or gender.

Some important research during the last decade points to the importance of implicit biases held by healthcare providers. In the context of healthcare, implicit biases refer to those unconscious believes, stereotypes, prejudices or perceived feelings that influence a health provider's judgement, and sometimes, even behaviour, based upon a patient's personal characteristics or belonging to a certain population group (12,13). Implicit biases can have an impact on anyone's judgement and, therefore decision-making, regardless of that person explicit rejection of negative ideas, connotations, or believes of disadvantaged groups. As with "aversive racism" (14), an open and explicit support of egalitarian practices or rejection of sexism does not necessarily mean an absence of implicit bias on gender and sex issues. Hence, implicit biases may lead to impartial care and are of special concern in healthcare because they are hard to detect and may imply a double burden on already vulnerable groups. This latter phenomenon was coined by the work of political philosophers De-Shalit and Wolff as "corrosive disadvantage": vulnerable subgroups of population might suffer sever grudges because of implicit biases, which, in turn, may lead to further disadvantages (15).

A systematic review by FitzGerald and Hurst (2017) shows that healthcare professionals hold "the same levels of implicit bias as the wider population", likely influencing "diagnostic and treatment decisions and levels of care in some circumstances" (12). Another more recent review by Champagne-Langabeer & Hedges (2021) gathers evidence on how the gender of health providers impacts different areas of clinical practice and manifests in a variety of contexts. The results show that a large majority of research in this field focuses on cardiovascular diseases and mental health diseases in the form of medical negligence and diagnostic errors (16–18); uncertainty in making diagnoses (19); symptom interpretation (20); risk perception and need for prevention (21); and the search for risk factors when creating clinical histories (18,22). Their review also indicates that implicit gender biases in healthcare providers are likely acquired through culture and education and are one of the origins of the perpetuation of health inequalities because they affect medical decision-making processes(13). Additionally, it is relevant to acknowledge the likeliness of health professionals to acquire new biases during their academic training and through clinical experience via cognitive processes. In fact, a systematic review by Saposnik, G. et al. (2016) indicates that cognitive factors such as overconfidence, anchoring biases, information biases, availability biases, and risk tolerance may be associated with diagnostic errors and other inaccuracies in patient clinical management. All evidence available highlights the need to carry out more studies to produce evidence that would allow to determine the prevalence of these biases and understand their impact on medical decisions made by healthcare personnel, errors in patient clinical management, and patient health (23).Our study aims to contribute precisely to this aspect of the literature by breaking down the following unanswered questions in this relevant field.

None of the studies or articles summoned in the reviews above mentioned collects evidence on the impact of biases on clinical management as it is the case of our study. First, the purpose of our research is to identify and quantify the impact of biases in the form of inequalities throughout the main set of tasks involved in the primary stage of the clinical management of a patient: anamnesis, initial diagnosis, assessment of the patient evolution and first response, request of medical tests based on the first diagnosis orientation, main diagnosis, treatment administration, assessment of the patient response and communication with the patient.

Second, there is still insufficient evidence on the potential effects these biases may have on a broad spectrum of diseases, not just the health conditions that are already known to be vulnerable to sex and gender differences in the form of symptomatic presentation and prognosis, such as cardiovascular diseases or mental health.

Third, it is still unknown how and when these biases are constructed and produced throughout the academic and professional development of healthcare professionals, how they interact in healthcare settings and, if they do, the way in which they translate into different health outcomes for both men and women.

In the research we propose, these biases and differences in healthcare delivery are observed and identified through simulated scenarios that will produce an opportunity for the participants of our study (medicine students as well as already active junior physicians) to operate and develop their work in an environment similar to a real-life event.

1.2. The use of clinical simulations

Clinical simulations offer an extraordinary opportunity to identify and understand the extent to which errors involved in disease and clinical management happen more often in in women than men patients as well as the type of health conditions that present more confusion or are most vulnerable to the lack of gender-sensitive approaches by health professionals. In this sense the use of clinical simulations and standardized patients to uncover and disclose sex and gender biases in medical students and healthcare professionals and their effects on health outcome variables will be a novelty in the literature.

Clinical simulations and the use of standardized patients are common tools that have long been used in medical settings to enhance the education and training level of students and have often been used to improve both technical and non-technical skills in healthcare professionals. The ultimate and general goal of clinical simulations is to develop better clinical practice skills or procedures to learn about patient safety issues in order to avoid them in real-life scenarios (24). Evidence suggest that the experiential learning opportunity offered in simulations has potential benefits in the form of outcomes such as improved teamwork, procedural skills, leadership, communication and clinical outcomes (25–28). It has also been shown that specialized learning of specific skills based on simulation-action can help improve patient outcomes (26). Medicine students use simulations to get acquainted with the clinical duties they will perform in the future, such as history taking, physical examination, diagnosis, and clinical management. Healthcare professionals also use medical simulation, although to a lesser extent, to learn about patient safety, the use of new techniques, interventions, protocols and methods, to observe

particular and specific cases, or simply to improve their teamwork and cooperation abilities and other soft skills.

As for research, in some cases, medical simulation has also been used to address sex and gender issues, but the focus is usually placed on gender and sex differences in the assessment of the performance of those being evaluated (29–31) or on biases hidden behind the methods used during simulation practices, which might be unfairly affecting students' or professionals' performances, assessment and, consequently, career development (32–35). However, simulations have been rarely used to evaluate sex and gender-based healthcare disparities, sex and gender biases in healthcare professionals, or to address sex and gender-based medical care, even if they represent an extraordinary platform to do so(36,37). Some of the few examples of research in this particular field focus on the use of standardized patient simulation to identify discrimination and poor service delivery for transgender and gender-diverse minorities (38–41) or for people who live with HIV/AIDS(42). However, gar from intending to downplay the importance and relevance of the research done on discrimination against gender minorities or people who live with HIV/AIDS, the fact that research on women's health is still not getting enough attention might just be proving the point of how implicit sex and gender biases can be in healthcare or how some people might believe that discrimination against women is a thing of the past, even if the scarcity of research and evidence produced in this field points to the opposite direction.

1.3.Main objectives of the research

Our study fills this research gap in the literature and will help address the above-mentioned challenges by achieving two main objectives:

a) Identifying how sex and gender biases operate in healthcare settings and the determination of the causal relationship between these biases and resulting sex and gender differences in the healthcare delivery process in the form of failed clinical management throughout the continuum of care or adverse patient safety events, such as an incorrect anamnesis of the patient, diagnostic errors, mistakes in clinical procedures (medical tests, treatment administration) or the communication with the patient.
b) Evaluating the stage of the medical career in which these biases originate. The focus will be on different stages of the medical career; our sample will include undergraduate medicine students and residency interns (known as *Médico Interno Residente*, MIR, in Spanish).

Special attention will be given to the policy implications of the study by contributing to improve clinical simulation practices developed in universities, medical faculties and healthcare settings to identify and revert gender and sex biases among their students, residents and senior healthcare providers. At the same time, the results will also be informative for healthcare managers that are concerned about gender biases in their institutions, as it will provide valuable information on the stage of the medical career in which it is most cost-effective to expose practitioners to simulation-based training on gender biases.

It is important to note that we refer to both *sex* and *gender* biases and differences in healthcare. Although both sex and gender biases are intrinsically linked, mostly because

knowledge and educational gaps in sex differences between men and women are partially caused by gender norms (43,44), emphasizing the difference between the two of them, in this case, is relevant. Sex biases and differences refer to those produced by knowledge and educational gaps based on the biological sex of the patient and the impact this one has on the health conditions of a patient, such as symptomatology and presentation of the disease, risk factors related to genetics, or treatment administration, which, in turn, are or can be independent to the gender of the patient. Gender biases and differences, instead, are based on gender or sexist assumptions and behaviours on women's health or the lack of educational awareness on gender diversity, among other elements. Given that we explore both types of biases and differences, we prefer to capture the full complexity of this matter by explicitly mentioning both terms.

2. DESIGN, COLLECTION OF DATA AND METHODS

2.1.Design of the study

This article presents an observational study that explores sex and gender biases and differences in healthcare delivery. The data used for analysis has been collected between January and May 2024 and gathers information about the performance of medicine students or medical interns (known as *Médico Interno Residente*, MIR, in Spanish) in clinical simulations. Simulation exercises are part of the academic or specialized training undertaken by medicine students or MIRs at the simulation laboratories located at the facilities of the universities collaborating in our research project.

Clinical simulations are a common practice for medicine students and MIRs at different stages of their academic and medical training. Each simulation practice or exercise has its own form of evaluation, based on a student's performance in analysing the presented clinical case and their ability to correctly perform the diagnosis and the rest of the clinical procedure. The evaluation criteria are always established by each professor or evaluator and is adapted to specific learning needs. However, all simulations include a series of steps that are essential for safeguarding patient safety by avoiding any incorrect diagnoses or clinical procedure errors, such as inducing side effects from medications or any system errors during the rest of the healthcare process, which are considered the minimum prerequisites for optimal and high-quality care.

The simulation cases that we used for the collection of data all followed the same procedure and evaluation. All simulation exercises were divided into three different parts: a pre-briefing, a simulated scenario with a standardized patient and a debriefing. In the pre-briefing, the participants were introduced to the simulation exercise and got a brief period of time to get situated and understand the main objectives of the exercise. Participants were also presented with a brief introduction about the area or field of medicine which the clinical case involved (i.e., cardiology, urology, psychiatry, traumatology, general medicine, etc.), as well as what was expected from them as doctors. However, participants never knew what specific clinical case (i.e., a certain disease, syndrome, medical emergency, injury, etc.) the patient was going to present with. None of the participants were allowed to perform or take a case alone: participants always had to perform in groups and were expected to perform as a real medical team in a real-life scenario by making joint decisions, interacting and helping each other and comply with

the objectives of the simulation exercise by executing the main tasks involved in the clinical management of patient. These were the following:

a) Conduct a complete interview and get all the indispensable and necessary information about the patient, their health status, their clinical and familiar history, as well as the symptomatology.
b) Make an initial diagnosis orientation.
c) Request or perform the necessary medical tests (diagnostic tests, laboratory work or screenings) to enable the differential diagnosis.
d) Act upon the patient's response and evolution.
e) Make a definitive diagnosis.
f) Administrate treatment accordingly.
g) Communicate and inform the patient about the process, as well as ask for their informed consent in order to perform the necessary tasks to solve the clinical case.

In the debriefing, participants went through the case and the scenario they encountered, shared their experience with the rest of the group that observed while performing the simulation exercise, learned and reflected about the possible mistakes or errors made, understood how and what they could improve about clinical management and had the opportunity to ask questions, learn more about the patient's case or any specific issue related to its clinical management or the simulation they just performed.

The simulations that we used for our study were mandatory simulation exercises included in the academic and training program of medicine students and residenct interns in 4 different Catalan universities: Universitat de Barcelona (UB), Universitat Pompeu Fabra (UPF), Universitat de Lleida (UdL) and Universitat Rovira i Virgili (URV). Medicine students and MIRs are not academically or professionally evaluated based on their performance during a clinical simulation practice. Instead, simulation exercises are designed to provide a safe space where participants can learn and deliver healthcare without the pressure of an evaluation, a real-life case scenario or the fear or making a fatal mistake with a real-life patient. In this sense, the main objective of the participants was to strengthen their clinical abilities and consolidate their medical knowledge, as well as become aware of what is indispensable to deliver an excellent health service and what could be improved about their performance.

Given the aim of the simulation exercises and the environment in which they are undertaken, it is important to mention that clinical simulations offer an extraordinary opportunity to isolate the effect of gender and sex biases among students and healthcare professionals and their impact on disease management and the health outcomes of patients, since other factors that may affect clinical management, especially those influencing the interaction between doctor and patient, are excluded. Simulations allow participants to act as they would in a healthcare setting, where they can access the necessary information about the patient's clinical history and act according to the needs of the clinical case, while avoiding the potential hurdles and obstacles that might hinder any process involved in clinical management, such as resistances from other professionals or the patients themselves, lack of time or resources, delays along the process, requesting medical tests, administrating treatment, etc. Simulations offer the opportunity to evaluate the decisions made along the clinical management process based on the healthcare provider's judgement, opinion, knowledge and actions.

2.2.Data collection

Because of ethical reasons, it was mandatory to inform simulation attendees about our study before the simulated case took place. However, with the aim to avoid bias, medicine students and MIRs were told about the possibility to voluntarily participate in a study before the simulation started, but the information about the study was carefully delivered. The information provided made them aware about the fact that their performance would be analysed in order to assess the use of clinical simulation to improve disease management abilities in the field of medical training, but it was made clear that they would never be academically or professionally evaluated for that purpose. Participants were never told about gender or sex biases either, as that could have altered the participant's performance in regards to this specific issue, which happens to be a fundamental asset and object of our study.

The data was collected through structured observations: the participants performance was assessed using a standardized procedure that allowed coding the results of the simulation evaluation based on a previously agreed criteria and using an evaluation sheet. (Table 1). Participants were randomly assigned to the treatment group (treatment variable: simulated female patients) or the control group (control variable: simulated male patients). While participants performed in the simulated scenarios, the evaluators in charge of the simulation exercise assessed their performance for each and every indicator listed in the evaluation form shown in Table 1. The items considered for the evaluation of the clinical management included the assessment of the patient's symptoms; the initial diagnosis based on the symptomatic presentation and the patient's medical and family history; the assessment of the patient's first response; the monitorization of the patient's vital signs and clinical status; the request of diagnostic tests, screenings and lab work according to the initial diagnostic orientation; the main diagnosis based on the symptomatic presentation, the patient's medical and family history and the results of the requested diagnostic tests; the administration of treatment; the assessment of the patient's response to the treatment; the communication with the patient; and the final assessment of the patient's status.

For every item or task involved in the clinical management process, the evaluator assessed and graded the medical team according to their performance, given three different options: if the participants developed more than 70% of the task, the execution was considered to be performed correctly without major errors or mistakes; if the participants executed only between 70% and 50% of the task, the task was considered to be partially wrong because of the omission of some of the processes involved or because an important error or mistake was made; if the participants executed less than 50% of a task, the execution was considered to be wrong, either because it involved major errors, severe mistakes or simply because it was not executed at all. In the case of a partially or totally wrong "Initial diagnosis" or "Main diagnosis", evaluators had to explicitly state what part of the process drove the participants to a wrong or poor execution.

Table 1. Evaluation sheet to assess the performance of the medical team during the clinical management of a patient

Items included in the process of clinical management		**Performance assessment: did the participant executed the following tasks?**			
		Yes (70% and more)	**Partially (70%-50%)**	**No (less than 50%)**	**Task not included in process**
1	Symptoms assessment				
2	Initial diagnosis (symptoms assessments, medical and family history)				
			What was **NOT** assessed correctly if graded as "partially" or "no" ☐ Symptoms assessment ☐ Medical and family history		
3	Assessment of patient's first response and evolution				
4	Monitorization of patient; vital signs and assessment of clinical status				
5	Request of medical tests				
6	Main diagnosis (symptoms, medical and family history, tests and screenings)				
			What was **NOT** assessed correctly if graded as "partially" or "no ☐ Symptoms assessment ☐ Medical and family history ☐ Tests and screenings		
7	Treatment administration				
8	Assessment of patient's response				
9	Communication to the patient of the diagnosis, procedures performed and request for informed consent if needed				
10	Final assessment of patient's status				

2.3. Data sample and information on the patient's characteristics, clinical cases, simulation participants

The data collected during the simulations does not only provide the results of the participants' performance on the clinical case, but also offers information about key characteristics of the standardized patient: mainly, their biological sex and their health issue. The number of participants and supervisors in each simulation exercise was also collected, as well as their gender. In the case of participants, the academic grade year was collected for medicine students (3rd, 4th, 5th or 6th). In the case of MIRs, the year of residency was also provided.

Hence, the data allows to distinguish between two separate groups: medicine students at the university level and residency interns. This distinction is important because, although the simulations always took place within the university setting, residency interns, while still training, are physicians who are already working and have more clinical experience than undergraduate level medicine students. The sample includes all 3rd, 4th, 5th, and 6th-year medical students from the universities collaborating in the study (UB, UPF, UdL, and URV) and residency doctors that train and practice using clinical simulations at these same faculties. First and second-year students are excluded because they typically do not

perform clinical simulations due to the lack of academic knowledge required for such exercises. There are no other exclusion criteria.

A total number of 96 simulation exercises have been used to collect data. 268 individuals agreed to participate in the study. 30.60% of the participants in simulation exercises were men, while 69,03% were women (women represent 73.44% of all medicine students in the Catalonia). One participant self-identified as non-binary (0.37%). A similar distribution holds for the supervisors (29.47% men; 70.53% women)(45). As for the standardized patients, 40.00% of them were male and 60.00%. were female.

Simulations exercises covered a wide range of diseases and health issues. Based on the International Classification Disease (ICD) code for diseases, we have identified 33 different diseases, belonging to thirteen different disease groups, among which we find diseases of the nervous system, the circulatory or cardiovascular system, the ear or mastoid process, the genitourinary system, the immune system, the respiratory system, the musculoskeletal system or connective tissue, neoplasms, developmental anomalies, infectious or parasitic diseases, mental, behavioural or neurodevelopmental disorders and traumatic injuries, poisoning or certain other consequences of external causes.

2.4.Data analysis and methods

The data collected in each simulation exercise is analysed by comparing the performance in each of the tasks involved in the clinical management based on the main variables of interest of our research: the patient's biological sex, the patient's health condition, and the participant's experience gradient. In order to facilitate the analysis, we transform each of the evaluated tasks into a binary variable: we assigned the value of 1 if the participant was graded with a good performance (70% of the task or more executed) and a 0 if the participant was graded either with a partial or wrong execution (70% of the task or less executed), given that a partial or wrong execution of the task indicate that important or severe errors or omissions have impeded a correct development of the task which, in some cases, could pose a risk for the safety and health of the patient. Hence, for each task, a mean score close to 1 indicates an overall good performance. In contrast, a mean score between 0 and 0.5 indicates an overall bad performance. This system also allow us to compare the difference in the total number of tasks developed incorrectly between male and female patients.

This straightforward approach aims to detect the causal effect between treating a simulated female patient (compared to a male patient) and the participants relative ability to complete all clinical tasks listed in the evaluation criteria for that specific case, which, ultimately, is the main goal of our study. However, because our collection of data also gathers information about the participant's experience gradient (year of undergraduate program or residency program), this methodology also allows to capture the degree of success in the development of clinical management and the healthcare delivery process depending on the participant's level of experience and medical knowledge. In addition, we have also been able to collect the simulation supervisor's gender. In this case, this information allows to capture differences in the evaluation of the simulation participants between men and women supervisors in terms of strictness or by helping simulation participants get back on track if mistakes are made along the process.

Data analysis is conducted using the OLS (Ordinary Least Squares) method, an econometric estimation model that will establish whether there is a statistically significant relationship between our research variables of interest and the participants' performance in simulation exercises. This will allow us to observe whether there are gender and sex biases among participants and whether they affect the clinical management of patients, measured by the participants' ability to perform the main tasks listed in Table 1.

3. (PRELIMINARY) RESULTS [1]

3.1. Biases and differences in the clinical management process

The role of the patient's biological sex

We begin with a simple regression analysis in which we compare and estimate the difference in the performance of the different tasks involved in the clinical management of a patient based on the patient's biological sex. We find statistically significant differences in the two first tasks: the assessment of symptoms and the initial diagnosis orientation based on the symptoms of the patient and their medical and familiar history. We find that, while for male patients the mean score of the symptoms assessment is 0.79 points, the mean score for female patients is 0.22 points less (0,57; $p<0.05$). This means that while male patients get their symptoms assessed correctly almost 80% of the time, female patients only do so 57% of the time. As for the initial diagnosis, this difference is even more pronounced: the mean score for male patients is 0.90 points, while the mean score for female patients is 0.655 points. In this case, the probability of female patients to receive an optimal initial diagnosis decreases by 0.24 points ($p<0.01$). (Table 2).

We observe that females also have a lower probability of receiving optimal care for the tasks that involve the monitorization of the patient and the main diagnosis when compared to male patients. However, these differences are not statistically significant and, especially in the case of the main diagnosis, are not as pronounced as the initial two.

In the processes that involve the request of medical tests, the administration of treatment and the communication with the patients, male patients obtain slightly worse care, but the differences are minimal and statistically non-significant.

[1] [Note to reviewers: only the main and basic results are presented because data collection finished just some days ago. Only 8 tasks from the evaluation form have been considered for the preliminary results in order to ease the analysis: assessment of symptoms, initial diagnosis, monitorization, request of medical tests, main diagnosis, treatment administration and communication with the patient. "Assessment of the patient's first response" and "final assessment of patient status" have been excluded because of lack of relevance].

Table 2. Mean scores obtained by the medical team while performing the main tasks involved in clinical management, based on the patient's biological sex

Clinical managament performance based on the patient sex

	Symptoms (1)	Initial dx (2)	Monitorization (3)	Tests (4)	Main diagnosis (5)	Treatment (6)	Comms (7)
Patient sex: female	-0.218**	-0.240***	-0.131	-0.052	0.018	0.005	0.062
	(0.098)	(0.088)	(0.109)	(0.084)	(0.114)	(0.113)	(0.105)
Constant	0.789***	0.895***	0.611***	0.838***	0.613***	0.515***	0.618***
	(0.076)	(0.068)	(0.083)	(0.066)	(0.088)	(0.088)	(0.082)
N	94	93	86	93	77	83	87
R^2	0.051	0.075	0.017	0.004	0.0003	0.00002	0.004
Adjusted R^2	0.041	0.065	0.005	-0.007	-0.013	-0.012	-0.008
Residual Std. Error	0.467 (df = 92)	0.420 (df = 91)	0.500 (df = 84)	0.399 (df = 91)	0.491 (df = 75)	0.506 (df = 81)	0.480 (df = 85)
F Statistic	4.944** (df = 1; 92)	7.367*** (df = 1; 91)	1.437 (df = 1; 84)	0.381 (df = 1; 91)	0.024 (df = 1; 75)	0.002 (df = 1; 81)	0.341 (df = 1; 85)

*p < .1; **p < .05; ***p < .01

In general, when accounting for the total number of errors or mistakes made during the whole clinical management of a patient in a simulation exercise, female patients have a lower probability of receiving optimal care, as more errors are captured when they are being attended. In our results, we estimate that the mean number of errors made during the whole healthcare delivery process for male patients is 2.37, while for female patients is 0.61 points higher. However, the difference is not statistically significant.

Table 3. Errors made in the clinical management of the patient

Mean number of total errors made throughout the whole management of the patient

	Errors in clinical management (1)	(2)	(3)
Patient sex: female	0.614		
	(0.554)		
Supervisor gender: woman		0.690	
		(0.594)	
Presence of women in the medical team			-0.792
			(0.947)
Constant	2.368***	2.250***	3.294***
	(0.429)	(0.499)	(0.720)
N	95	95	95
R^2	0.013	0.014	0.007
Adjusted R^2	0.002	0.004	-0.003
Residual Std. Error (df = 93)	2.643	2.642	2.651
F Statistic (df = 1; 93)	1.230	1.348	0.700

*p < .1; **p < .05; ***p < .01

The role of the supervisor

We observe that, in general, women supervisors are stricter that men supervisors when evaluating the group of participants that act as the medical team. Except for the assessment of symptoms, women supervisors grade participants with lower scores for the rest of the tasks involved in the clinical management when compared to men supervisors.

However, this difference is statistically significant only in the mean score of the treatment administration.

Table 4. Mean scores obtained by the medical team while performing the main tasks involved in clinical management, based on the supervisor's gender

Clinical managament evaluation based on supervisor gender

	Symptoms (1)	Initial dx (2)	Monitorization (3)	Tests (4)	Main diagnosis (5)	Treatment (6)	Comms (7)
Supervisor gender: woman	0.075	-0.098	-0.145	-0.124	-0.145	-0.223*	-0.140
	(0.108)	(0.098)	(0.123)	(0.089)	(0.123)	(0.123)	(0.109)
Constant	0.607***	0.821***	0.727***	0.893***	0.727***	0.682***	0.750***
	(0.090)	(0.082)	(0.104)	(0.075)	(0.104)	(0.106)	(0.090)
N	94	93	77	93	77	83	87
R^2	0.005	0.011	0.018	0.021	0.018	0.039	0.019
Adjusted R^2	-0.006	0.0001	0.005	0.010	0.005	0.027	0.007
Residual Std. Error	0.478 (df = 92)	0.434 (df = 91)	0.486 (df = 75)	0.395 (df = 91)	0.486 (df = 75)	0.496 (df = 81)	0.476 (df = 85)
F Statistic	0.480 (df = 1; 92)	1.006 (df = 1; 91)	1.405 (df = 1; 75)	1.914 (df = 1; 91)	1.405 (df = 1; 75)	3.264* (df = 1; 81)	1.637 (df = 1; 85)

*p < .1; **p < .05; ***p < .01

Results also show that women supervisors tend to assign more errors (0.719 points) in the whole process of healthcare delivery than men supervisors (mean number of errors when the supervisor is a men: 2.25; vs women supervisors: 2.96. p>0.05). However, the difference is not statistically significant (Table 3).

The role of the gender distribution in the medical team

The results show that the gender distribution of the medical team, measured by the percentage of women composing the medical team, barely plays any role in the performance of tasks involved in the clinical management process. Only in the case of the communication with the patient, a higher presence of women doctors in the medical team increases the mean score of the task, with a statistically significant difference. As for the rest of the processes involved, no statistically significant differences are appreciated.

Table 5. Mean scores obtained by the medical team while performing the main tasks involved in clinical management, by the gender distribution of the medical team (percentage of women in the group)

Clinical managament performance based on the gender distribution of the medical group

	Symptoms (1)	Initial dx (2)	Monitorization (3)	Tests (4)	Main diagnosis (5)	Treatment (6)	Comms (7)
Supervisor gender: woman	-0.138	0.039	0.194	0.083	0.194	-0.028	0.356**
	(0.171)	(0.157)	(0.196)	(0.144)	(0.196)	(0.188)	(0.174)
Constant	0.757***	0.725***	0.488***	0.749***	0.488***	0.538***	0.402***
	(0.130)	(0.119)	(0.147)	(0.109)	(0.147)	(0.144)	(0.134)
N	94	93	77	93	77	83	87
R^2	0.007	0.001	0.013	0.004	0.013	0.0003	0.047
Adjusted R^2	-0.004	-0.010	-0.0003	-0.007	-0.0003	-0.012	0.036
Residual Std. Error	0.477 (df = 92)	0.436 (df = 91)	0.488 (df = 75)	0.399 (df = 91)	0.488 (df = 75)	0.506 (df = 81)	0.469 (df = 85)
F Statistic	0.656 (df = 1; 92)	0.063 (df = 1; 91)	0.979 (df = 1; 75)	0.331 (df = 1; 91)	0.979 (df = 1; 75)	0.022 (df = 1; 81)	4.169** (df = 1; 85)

*p < .1; **p < .05; ***p < .01

Results also show that the presence of women in medical teams reduces the number of errors made during the clinical management of the patient compared to those group without women doctors (-0.792 points; p>0.05). However, the difference is not statistically significant (Table 3).

4. DISCUSSION

The two first items of the clinical management process, the assessment of symptoms and the initial diagnosis, are extremely important, since they set the first orientation for healthcare providers and guides the rest of the service delivery. In the case of a simulation exercise, participants have many strategies and opportunities to set the clinical management of the patient back to track along the process, but in a real-life scenario, where doctors might not be as supervised or might not have other resources to get to a definitive diagnosis, failing to understand the symptomatology of the patient, the disease presentation and setting a wrong initial diagnosis might have fatal consequences for the rest of the clinical course and the patient's prognosis.

We observe that there are no significant differences for any of the rest of tasks involved in the clinical management process. We believe this happens for two non-excluding possible reasons. The first responds to the fact that the clinical management process has its own checks and balances, working in favor of mending errors based on the more subjective tasks of the process. This is especially true if we consider that, right after the first initial diagnosis, the medical team introduces clinical and diagnostic tools, such as the monitorization and the medical tests, that provide objective information of the patient and its clinical status that allow doctors to change their first revise their first approach and switch the diagnosis if needed. The second reason we believe this happens is that in simulations practices, supervisors help participants (the medical team) along the process by asking questions, giving hints, or simply by putting more emphasis on important clinical information about the patient that is provided to the doctors. This might explain why in most cases, after a wrong initial assessment of symptoms and initial diagnosis, most participants are able to get a definitive and correct main diagnosis.

As for the role of the health provider's gender in the clinical management of the patient, we do not appreciate important differences in any of the tasks involved, except for the communication with the patient. This might implies that women doctors might have better communication skills, provide more information to the patient and are more attentive about getting the informed consent of the patient than men doctors.

5. CONCLUSIONS

This study represents a clear advancement both academically and socially, with a special emphasis on the implications our results may have for the implementation of public health policies. The design of an effective simulation tool capable of efficiently and directly detecting gender and sex biases will enable academic institutions, health managers, and policymakers to choose the most cost-effective and appropriate strategies to reverse them, exposing medical students and professionals to simulation-based training on gender biases.

On all these central issues mentioned above and about which there is still insufficient scientific evidence, the study attempts to provide rigorous answers, as they are fundamental factors for designing interventions capable of reducing and eliminating sex and gender inequalities in health settings. The systematic review by Alcalde-Rubio et al. in 2020 evidenced that most interventions aimed at reducing inequalities usually succeed in "reducing gender inequalities in at least one key health factor, even when they were not explicitly intended or aimed at reducing gender differences"(46). This is very

important for the proposal of our research, as, following this reasoning, more effective and efficient interventions could be conceived in the future based on the results of our research, allowing a more precise quantification of gender and sex biases and identifying their origins.

Additionally, diagnostic errors or delays, when conducting tests, screenings, treatment, or inadequate clinical follow-up can cause significant healthcare and economic inefficiencies in health systems and have a wide range of harmful consequences: diagnostic errors can be fatal if they result in the death of a patient or severely affect morbidity, can cause avoidable suffering to patients, can delay treatment administration or lead to unnecessary or harmful treatment, and can even have psychological or economic repercussions for patients. Ultimately, they can also cause harm to the professionals involved in the diagnostic process and undermine trust in the public health system(47). In reference to all this, the study identifies when and why diagnostic errors occur, at what stage of the academic and/or professional career it is more likely for biases responsible for these errors to appear, and if, ultimately, they are the cause of the gender disparities in health observed in clinical practice. Finally, given that women are more affected by multimorbidity than men(48–52), identifying gender and sex biases in healthcare professionals and their impact on disease management will help avoid this double cause of diagnostic errors.

In conclusion, our study contributes significantly to the understanding and reduction of gender and sex inequalities in healthcare. By using clinical simulation to identify and quantify gender biases in medical students and healthcare professionals, we provide valuable evidence for the development of more effective and efficient interventions. This research has the potential to inform and transform clinical practice, leading to improved health outcomes and reduced health disparities.