Inga Hege*, Meike Hiedl, Karin Christine Huth and Jan Kiesewetter

Differences in clinical reasoning between female and male medical students

https://doi.org/10.1515/dx-2022-0081
Received July 14, 2022; accepted November 7, 2022;
published online November 18, 2022

Abstract

Objectives: In undergraduate medical education virtual patients (VPs) are a suitable method to teach clinical reasoning and support the visualization of this thinking process in a safe environment. The aim of our study was to investigate differences in the clinical reasoning process and diagnostic accuracy of female and male medical students.

Methods: During the summer term 2020, we provided access to 15 VPs for undergraduate students enrolled in a medical school in Bavaria, Germany. All interactions of the 179 learners within the VP system CASUS were recorded, exported, and analyzed.

Results: We found significant differences in the clinical reasoning of female and male learners. Female students documented more findings, differential diagnoses, tests, and treatment options and more often created a summary statement about the VP. Their overall performance was higher than those of their male peers, but we did not see any significant differences in diagnostic accuracy.

Conclusions: The significant differences between male and female medical students should be considered when planning teaching and research activities. A future study should investigate whether these differences can also be found in physicians.

Keywords: clinical reasoning; gender; virtual patients.

Introduction

Clinical reasoning is a core ability students have to learn during medical school and various dedicated teaching and assessment methods for clinical reasoning have been developed and explored. For example, script concordance tests or virtual patients (VPs) are used in teaching and assessing clinical reasoning skills [1]. Virtual patients provide a simulated and safe environment in which students can practice clinical reasoning and learn from errors without harming patients [2].

To structure the clinical reasoning process in virtual patients, a concept mapping approach has been integrated allowing learners to visualize components of illness scripts [3].

In a recent study we discovered different reasoning patterns among participants depending on whether they were able to solve the case on their own. Learners who could not solve the case on their own documented fewer problems, differential diagnoses, tests, treatment options, and connections [4]. However, we did not investigate other influencing factors on the process and diagnostic accuracy, such as learners´ gender.

In medical education students are often regarded as a homogenous group and a one-size-fits-all approach applied instead of focusing more on individual differences in learning, but also decision making. Such factors include for example, religion, experience, cognitive ability, personality, and gender [5]. Research has shown that female and male physicians differ in their clinical practice, that these differences impact clinical reasoning, and that it appears as if there are also differences in diagnostic accuracy [5]. Gender-related studies in healthcare education showed that male trainees tend to rate themselves as more competent than their actual training level and female students experience more gender biases by patients [6]. Females performed better and in less time in skills such as suturing [7], communication and empathy [8, 9], but performed equally in microsurgical skills [10]. In other learning settings female students showed a better performance than male learners, as shown by Das et al. for academic performance after a team-based training session [11]. In a study by Wahlquist et al. patient-centered attitudes in female medical students were significantly higher than in male students [12].

***Corresponding author: Inga Hege**, Medical Education Sciences, University of Augsburg, Universitätsstr. 2, 86159 Augsburg, Germany; and Institute for Medical Education, LMU Munich, Germany, E-mail: inga.hege@med.uni-augsburg.de
Meike Hiedl and Jan Kiesewetter, Institute for Medical Education, LMU Munich, Germany
Karin Christine Huth, Department of Conservative Dentistry and Periodontology, University Hospital, LMU Munich, Germany

Concerning clinical reasoning, a study by Groves et al. showed that female students had higher scores in solving paper-based clinical reasoning problems, i.e., identifying and interpreting clinical features and generating hypotheses, than their male peers [13]. However, this study did not investigate specific differences within the clinical reasoning process.

Therefore, our aim was to identify whether significant differences between female and male medical students in the clinical reasoning process exist in a virtual patient environment and if so which. For this purpose, we investigated aspects such as time on task, diagnostic accuracy, number and quality of identified symptoms and findings of the virtual patient, differential diagnoses, tests, treatment options, and number of composed summary statements.

Identifying gender differences in the clinical reasoning process will help to interpret future studies on clinical reasoning and determine whether a gender-specific teaching and assessment approach is necessary.

Materials and methods

Design of virtual patients

We developed 15 VPs in the CASUS system [14] covering a range of symptoms and diagnoses (Table 1).

The VPs are based on fiction and structured on 6–8 screen cards as follows: (1) Introduction with key symptom(s) and a picture of a fictional patient, (2) history taking in form of dialogs between the patient and the physician, (3) physical exam findings, (4) further data from additional tests and examinations, such as lab values or imaging, (5) final diagnosis made by the learners, and (6) treatment and management options. The VPs are combined with a concept mapping tool [3] in which learners are required to document and visualize the patient's problems and findings, differential diagnoses, examinations and tests, and treatment options (Figure 1). Additionally, they are prompted to compose a concise summary statement about the VP, which is analyzed and scored on a rubric developed by Smith et al. [15, 16]. To conclude a VP scenario, learners have to provide a final diagnosis and indicate their confidence with this decision on a scale from 0 to 100. If they are unsuccessful, they can request the solution from the system. All interactions with the system are recorded and stored in a database. Also, scores for each interaction with the map, such as adding items or submitting a final diagnosis and the summary statement are automatically rated by the system. Based on all these scores, an overall score for the concept map is calculated [3, 15].

The VPs were reviewed for content accuracy by a content matter expert and by a didactical expert for didactical issues and appropriateness of level of difficulty for 3rd/4th year medical students. The course was then offered during winter term 2019/2020 for the first time and we analyzed the session data obtained during that semester for any inconsistencies or needs for changes, focusing especially on the concept maps students created.

Table 1: Overview of VPs provided in the course, names and stories of the virtual patients are fictional and not based on real persons.

Virtual patient	Age and gender	Key symptom	Final diagnosis
Thomas Hechser	34, male	Back pain	Muscular tension
Annegret Huefner	40, female	Back pain	Cholecystolithiasis
Carmen Kuhnert	48, female	Incidental finding	White coat hypertension
Hannah Loewinger	32, trans female	Trauma	Anterior cruciate ligament tear
Martha Nebelhuber	86, female	Chest pain	Aortic valve stenosis
Khadija Okeke	5, female	Cough	Cystic fibrosis
Adeline Polignac	74, female	Reduced vigilance	Urinary tract infection
Viktor Prenzel	58, male	Micturition disorder	Adverse drug effect
Nathalie Roesler	20, female	Cough	Pulmonary embolism
Theo Schiller	20, male	Nausea	Intoxication
Greta Schilling	7, female	Joint pain	Borreliosis
Wiebke Sommer	26, female	Dyspnea	Pneumothorax
Mia Weindl	20, female	Trauma	Radius fracture
Harald Wenzel	68, male	Hematemesis	Ulcus bleeding
Peter Zanger	67, male	Chest pain	Pneumonia

Participants and data collection

For the summer term 2020 we provided a course with the 15 VPs via the virtual university in Bavaria (vhb) [17]. All students enrolled in one of the six Bavarian universities could get access to the course for free, but registration with their university credentials was necessary. We included all participants into our study who completed at least one VP, i.e., they submitted a final diagnosis. Consent for study participation with anonymized data was obtained from participants upon logging into CASUS. The integration of the VPs into the curriculum varied between the schools. A course certificate confirming course passing was automatically provided after the completion of 10 VPs.

Data analysis

All interactions of the learners with the VPs were recorded by the CASUS system and stored in a database. We exported the anonymized data from CASUS and imported it into SPSS 26 (IBM, USA) for further analysis. We compared male and female participants with the following dependent variables:

- Mean time on task (measured from opening a VP until closing it).
- Mean number of findings.
- Mean number of differential diagnoses.

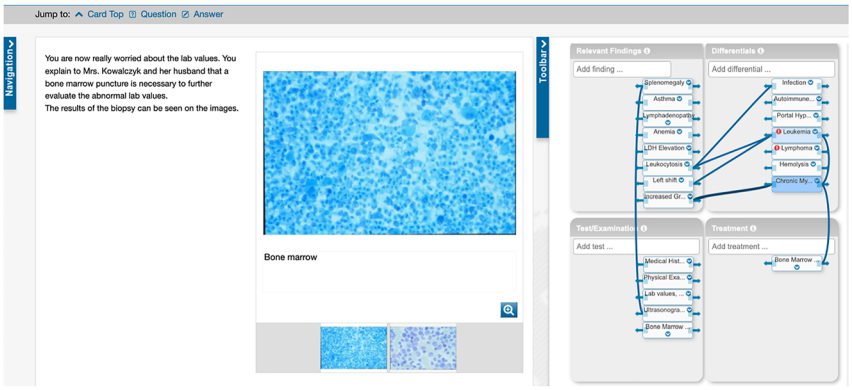

Figure 1: Screenshot of an exemplary VP in the CASUS system showing the case on the left side and the concept map created by learners on the right side.

- Mean number of tests.
- Mean number of treatment options.
- Mean number of connections.
- Mean total score for the concept map.
- Mean number of attempts until correct final diagnosis (diagnostic accuracy).
- Mean total number of composed summary statements (short summary of the case as if presenting to a senior physician).
- Mean total score for the composed summary statements.
- Mean number of requests for the correct final diagnosis to be revealed by the system, and
- Mean level of confidence with the final diagnosis decision.

For the analysis we used a MANOVA to compare all dependent variables at once. Prior to the analysis we tested the normal distribution of data using distance of the mahalanobis and Chi-squared test. Data was normally distributed, alpha error was set to $p<0.05$.

Results

Participants and sessions

Overall, 192 learners accessed the course and 179 learners (93.2%) completed at least one VP. From April 17th until July 31st, 2020 we recorded 1865 VP sessions, of which 1791 (96.0%) have been completed, i.e. a final diagnosis has been provided. On average, learners completed 10.0 VPs (min: 1, max: 15) without a significant difference between female and male students. All students but one (from the University of Erlangen) were enrolled at the University of Munich. Table 2 illustrates the non-significant differences between female and male students.

Concept map and summary statement

On average, female students reached a higher score and documented significantly more findings, differential diagnoses, tests, and treatment options, but the number of connections drawn was similar (see Table 3 for details).

Table 2: Differences between female and male learners concerning number and duration of VP sessions.

	Female	Male
Number of learners (n=192)	124 (64.6%)	68 (35.4%)
Learners completing at least one VP (n=179)	117 (65.4%)	62 (34.6%)
Mean age in years (n=179)	24.2	23.7
Mean number of VPs completed	10.2	9.7
Completed sessions (n=1791)	1,191	600
Mean time spent on one VP	16.0 min	15.1 min

Table 3: Number of items added to the concept maps and overall score.

	Female	Male
Number of findings	5.0[a]	3.9[a]
Number of differential diagnoses	4.4[a]	3.7[a]
Number of tests	3.3[a]	2.6[a]
Number of treatment options	1.7[a]	1.1[a]
Number of connections	0.4	0.4
Overall score for the concept map	0.32[a]	0.27[a]

[a]Statistically significant ($p<0.001$).

For 1120 VP sessions (62.5%) learners composed a summary statement. In 55.5% of VP sessions (n=333) male students and 74.5% (n=887) female students created a summary statement, this difference is significant ($p<0.001$). Additionally, female learners achieved significantly higher scores for their summary statements than male students (0.88 vs. 0.75, $p=0.012$).

Diagnostic accuracy

Overall, in 80.9% of the VP sessions (n=1,448) learners made the correct final diagnosis on the first attempt. In 80.6% of sessions (n=960) female students and in 81.3%

sessions (n=488) male students made the correct final diagnosis on the first attempt, but this difference was not significant. However, male learners significantly more often requested the final diagnosis from the system (19.0%, p=0.046) than female learners (15.3%).

We did not find a significant difference between male and female students concerning their confidence with their final diagnosis and errors (64.6% female, vs. 66.1% male). Male and female students both showed a lower confidence in VPs they gave up on finding the final diagnosis (57.0% female vs. 58.9% male).

Discussion

In our study we investigated differences between clinical reasoning processes of female and male medical students in a virtual patient learning environment.

We did not see any significant differences between female and male learners concerning the number of VPs they completed and the time they spent on average on these VPs. To receive the course certificate, the completion of 10 VPs was required and this was also the average completion rate for female and male learners. Thus, neither female nor male learners do more than they are required to for passing the course.

Female medical students showed a more thorough approach to creating the concept maps adding more findings, differential diagnoses, tests, and treatment options, which resulted in a higher score for the map. This result is comparable to other studies that showed a better performance of female students compared to their male peers [11]. It also confirms a study by Groves et al. that showed that female students identified more clinical features and generated more hypotheses when solving paper cases [13]. Another study showed that female physicians are more likely to order diagnostic tests [18], so our study results indicate that this pattern is already present during early clinical years in medical school.

Interestingly, despite the better performance on creating the concept map, female learners did not show a significantly higher diagnostic accuracy, i.e., coming up with the final correct diagnosis on the first attempt, than their male peers. A reason for this could be that the VPs were not difficult or complex enough to reward female students for their more thorough approach with a higher diagnostic accuracy than their male peers. However, male students gave up quicker in cases where they were not able to identify the correct final diagnosis on the first attempt. In an earlier study we showed that maps in which learners gave up and requested the final diagnosis from the system had a lower score, fewer summary statements, and contained fewer problems, differential diagnoses, tests, and treatment options [4].

Across all sessions we found a similar confidence of female and male learners with their decision for the final diagnosis. Other studies showed that male students tend to be more self-confident in their clinical skills performance [19].

Although female learners created more extensive maps and more often composed a summary statement, they did not spend significantly more time working on the VP. Additionally, male students gave up quicker on coming up with the correct diagnosis, which saved some time. Therefore, future research should look at the time needed for map creation, summary statement composition, and final diagnosis independently. Also, a study based on think-aloud with VPs could be beneficial to learn more about the rationale behind adding items to the concept map would be helpful to learn about why these differences exist in medical students.

Limitations

Our study has some limitations. Firstly, the VPs we used for our study are limited in their scope and can only give insights into gender differences of medical students for the symptoms and diagnosis covered by these 15 VPs. However, with our selection of diverse key symptoms we tried to minimize this effect. Secondly, the students received a course certificate, which motivated them to complete the required number of VPs, but not all VPs in the course. However, as we did not see a significant difference between male and female learners in the number of completed VPs, we believe this affects them in the same way.

Third, we do not have further information about the course participants, such as the exact semester, previous experience with virtual patients, and their socioeconomic status, so we cannot fully exclude that there are confounding factors not addressed or controlled for in our study. However, regarding the semester, we can assume that from the integration setting, students are mostly in their 3rd year and as they are all (but one) from the same university. As such, their previous experience with virtual patients is presumably the same as VPs with specific clinical reasoning focus are not introduced in the Munich curriculum in year 1 and 2. Further, the study was planned as a multi-center study with participants from medical schools across Bavaria, but the results showed that we had an overwhelming number of students from one school. While this makes differences between male and female

students less likely, it adds the limitation of a single-center study, so further research is needed with implementing a cross-institutional study.

Conclusions

Our study showed significant differences in the clinical reasoning process of female and male medical students when working on virtual patients. These differences should be considered when teaching and researching clinical reasoning. For example, male students could be prompted more often than female students to express and visualize their reasoning process. Future studies should also investigate whether these differences are limited to medical students or can also be seen in junior and senior physicians.

Acknowledgments: The authors would like to thank the Virtual University of Bavaria (vhb) for providing the funds for creating and sustaining the virtual patients. Moreover, we would like to thank all students working with the virtual patients and educators integrating them into their teaching.

Research funding: The creation of the VPs and the continuous support is funded by a grant from the Virtual University in Bavaria (vhb) (Project number: 18-I-05-16Heg1).

Author contributions: JK and IH planned the study, MH implemented the study, MH, JK, JCH, IH analyzed and interpreted the data, IH drafted the manuscript and all authors critically reviewed and revised the manuscript. All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: Authors state no conflict of interest.

Informed consent: Informed consent was obtained from all subjects and the study has been performed in accordance with the Declaration of Helsinki.

Ethical approval: Ethical approval for the study was granted by the Ethical Committee of the Ludwig Maximilians-University Munich, Germany (Project number: 19–571 KB).

References

1. Trowbridge RL, Rencic JJ, Durning SJ. Teaching clinical reasoning. Philadelphia, Pennsylvania, USA: American College of Physicians; 2015.
2. Kononowicz AA, Woodham LA, Edelbring S, Stathakarou N, Davies D, Saxena N, et al. Virtual patient simulations in health professions education: systematic review and meta-analysis by the digital health education collaboration. J Med Internet Res 2019;21:e14676.
3. Hege I, Kononowicz AA, Adler M. A clinical reasoning tool for virtual patients: design-based research study. JMIR Med Educ 2017;3:e21.
4. Hege I, Kononowicz AA, Kiesewetter J, Foster-Johnson L. Uncovering the relation between clinical reasoning and diagnostic accuracy – an analysis of learner's clinical reasoning processes in virtual patients. PLoS One 2018;13:e0204900.
5. Croskerry P. Individual variability in clinical decision making and diagnosis. In: Diagnosis: interpreting the shadows. London, UK: CRC Press; 2017.
6. Pearce G, Sidhu N, Cavadino A, Shrivathsa A, Seglenieks R. Gender effects in anaesthesia training in Australia and New Zealand. Br J Anaesth 2020;124:e70–6.
7. Chiu HY, Kang YN, Wang WL, Tong YS, Chang SW, Fong TH, et al. Gender differences in the acquisition of suturing skills with the da Vinci surgical system. J Formos Med Assoc 2020;119:462–70.
8. Graf J, Smolka R, Simoes E, Zipfel S, Junne F, Holderried F, et al. Communication skills of medical students during the OSCE: gender-specific differences in a longitudinal trend study. BMC Med Educ 2017;17:75.
9. Santos MA, Grosseman S, Morelli TC, Giuliano ICB, Erdmann TR. Empathy differences by gender and specialty preference in medical students: a study in Brazil. Int J Med Educ 2016;7:149–53.
10. Sudario-Lumague R, Chiang YC, Lin TS. Gender comparison of medical student microsurgical skills in a laboratory model. J Reconstr Microsurg 2018;34:359–62.
11. Das S, Nandi K, Baruah P, Sarkar SK, Goswami B, Koner BC. Is learning outcome after team based learning influenced by gender and academic standing? Biochem Mol Biol Educ 2019;47:58–66.
12. Wahlqvist M, Gunnarsson RK, Dahlgren G, Nordgren S. Patient-centred attitudes among medical students: gender and work experience in health care make a difference. Med Teach 2010;32:e191–8.
13. Groves M, O'rourke P, Alexander H. The association between student characteristics and the development of clinical reasoning in a graduate-entry, PBL medical programme. Med Teach 2003;25:626–31.
14. CASUS. Virtual patient system. 1999-2022; 2020. Available from: http://crt.casus.net [Accessed 6 Jul 2022].
15. Smith S, Kogan JR, Berman NB, Dell MS, Brock DM, Robins LS. The development and preliminary validation of a rubric to assess medical students' written summary statements in virtual patient cases. Acad Med 2016;91:94–100.
16. Hege I, Kiesewetter I, Adler M. Automatic analysis of summary statements in virtual patients - a pilot study evaluating a machine learning approach. BMC Med Educ 2020;20. https://doi.org/10.1186/s12909-020-02297-w.
17. Virtuelle hochschule bayern (vhb). Available from: https://www.vhb.org/en/ [Accessed 6 Jul 2022].
18. Freeborn DK, Levinson W, Mullooly JP. Medical malpractice and its consequences: does physician gender play a role? J Gend Cult Health 1999;4:201–14.
19. Gazibara T, Wiltshire-Fletcher M, Maric G, Kozic D, Kisic-Tepavcevic D, Pekmezovic T. Self-confidence and clinical skills: the case of students who study medicine in English in a non-English speaking setting. Ir J Med Sci 2019;188:1057–66.