Submitted: 29 December, 2022 Accepted: 14 March, 2023 Published: 08 September, 2023 DOI:10.22514/sv.2023.081

Open Access
Signa Vitae

ORIGINAL RESEARCH

Effects of the quality of medical history taking on diagnostic accuracy

Stefan Götz[1], Julie Fournier[1], Kai Tisjlar[2], Timur Sellmann[3,4], Stephan Marsch[2,*]

[1]University of Basel, 4031 Basel, Switzerland
[2]Intensive Care, University Hospital Basel, 4031 Basel, Switzerland
[3]Department of Anaesthesiology and Intensive Care Medicine, Bethesda Hospital, 47053 Duisburg, Germany
[4]University of Witten/Herdecke, 58455 Witten, Germany

***Correspondence**
stephan.marsch@usb.ch
(Stephan Marsch)

Abstract

Diagnostic errors are a relevant health-care problem. Although medical history taking is usually the first step in patients' assessment there are only limited data on the association of its quality and diagnostic accuracy. Accordingly, this prospective randomized simulator-based single-blind trial aimed to investigate the effects of initial cues and history taking skills on diagnostic accuracy. 198 medical students (135 females) were given the task to assess a patient presenting with simulated acute pulmonary embolism. Participants were randomized to six versions of the scenario differing only in the initial cues, *i.e.*, in the reply of the patient to the initial question about the reason for his visit. In three of six versions, initial cues were restricted to thoracic symptoms (chest pain, dyspnoea, or combination of both). In the remaining three versions, initial cues consisted of thoracic and extra-thoracic (leg pain, immobilization) symptoms. The primary outcome was diagnostic accuracy. The number of initial cues was unrelated to diagnostic accuracy. However, the combination of extra-thoracic and thoracic cues resulted in a higher diagnostic accuracy than thoracic cues only (52/96 *vs.* 35/102, $p = 0.006$). In multivariate regression, the number of questions asked from the categories "risk factors of pulmonary embolism" (regression coefficient 0.15, $p < 0.001$) and "dyspnea" (regression coefficient 0.12, $p < 0.001$) predicted diagnostic accuracy. Moreover, questions relating to "immobilization" (regression coefficient 0.42, $p < 0.001$), "onset of dyspnea" (regression coefficient 0.23, $p = 0.003$), and "modifying factors of chest pain" (regression coefficient 0.20, $p = 0.04$) independently predicted diagnostic accuracy. Interestingly, more systematic history taking was associated with lower diagnostic accuracy (regression coefficient −0.27, $p < 0.001$). The present trial demonstrates that during history taking cues initially revealed by the patient, kind and category of questions asked during the interview, and the interview's structural systematics affect diagnostic accuracy.

Keywords

Randomized trial; Medical simulation; Diagnostic error; Initial cue; Emergency medicine; Medical education

1. Introduction

Diagnostic errors are a relevant health-care problem [1, 2]. As medical history taking is usually the first step in a patient's assessment its quality is most likely a first and important barrier against diagnostic errors. There are, however, only limited data on the association of the quality of medical history taking and diagnostic accuracy.

Medical history taking starts with asking the patient to reveal the reason for his/her presentation to the health-system. The following expression of the presenting symptoms can be regarded as initial complaints or initial cues [3]. Most symptoms of any distinct disease can vary between affected patients. In addition, the expressions of patients expiring identical symptoms may also vary. Thus, physicians have to deal with a potentially high variance in initial cues, *i.e.*, symptoms revealed to them upon their opening question of history taking. Previous work has demonstrated that initial cues can lead to diagnostic errors [4, 5]. There are, however, currently no data on the relation of number, kind, and combination of initial cues with diagnostic accuracy.

After the revelation of the presenting complaints medical history taking continues with a physician-led interview exploring, among others, the presenting complaints, past illnesses, risk factors for the presumed and/or other diseases, and social history. History taking skills can be assessed by the categories of breadth and depth [6]. While breadth considers covering the relevant problem areas a patient might have, depth relates to covering the relevant details of each problem area. A further tool to assess the quality of history taking is

the logical sequence or appropriate flow of questions asked [6]. Whether and how breadth, depth, and logical sequence of medical history taking relates to diagnostic accuracy is largely unknown. Extent and time frame of medical history taking depends on the medical context so that history taking is shorter and more focused in emergency situation. Rapid evolution of their disease and limited opportunities to detect and correct errors makes patients in acute care particularly vulnerable to harm related to a delayed or wrong diagnosis [7, 8]. There are however only limited data on the value of medical history taking for diagnostic accuracy in emergent cases. Acute pulmonary embolism (PE) is commonly encountered in emergency settings. Though timely diagnosis and treatment is paramount to prevent morbidity and mortality, PE is prone to diagnostic errors and delayed diagnosis [9–12].

The goal of this prospective randomized trial was to investigate the effects of initial cues and the quality of medical history on diagnostic accuracy. Specific aims were to test the hypotheses that diagnostic accuracy is influenced (1) by kind and number of symptoms revealed to the question of the reason for the patient's presentation (*i.e.*, initial cues), and (2) by breadth, depth and logical sequence of history taking. The trial was conducted with medical students assessing a simulated case of acute pulmonary embolism.

2. Methods

2.1 Participants

From 2014 till 2019, 4th years medical students from the University of Basel were offered to participate in voluntary simulator-based workshops conducted in the Simulation Centre of the Intensive Care Unit of the University Hospital Basel, Switzerland. Participants had no previous experience with high-fidelity simulation. The topic PE had been covered by lectures in their curriculum. The present prospective, randomized single-blind trial (participants prior to their performance in the simulated scenario not being aware of the purpose of the study) was conducted as part of the workshops. After their performance in the simulated scenario, the purpose of the study was revealed to the participants. The study is registered (ClinicalTrials.gov: NCT05760859) and reported herein according to the extensions to the Consolidated Standards of Reporting Trials (CONSORT) statement of the Reporting Guidelines for Health Care Simulation Research [13].

2.2 Simulated Scenario

Participants received a standardized instruction of the patient simulator (human patient simulator, Laerdal Medical SimMan). The mannequin was equipped with a speaker in its head that broadcasts the voice of a remote operator allowing verbal interaction of the participants with the "patient". For the present trial, the operator producing the voice of the "patient" was an experienced physician (K.T., S.M.). The "patient" did not speak unless asked and responded to all participants' questions during history taking according to a pre-defined checklist (Table 1).

Participants performed the trial scenario on their own, were given the role of a resident-on-call in the emergency department, and were instructed by a senior physician (confederate) to triage a newly arrived patient. Upon entering the simulator room, the participants encountered the "patient" lying on a hospital bed and dressed in a hospital gown. Participants were assisted by a confederate nurse which was instructed to display a helpful manner, but to act only on command and not to interfere in any way with the assessment and history taking. However, the nurse was instructed to specify the exact point of pain localization on the mannequin if participants asked the patient about the localization of his chest pain; in addition, the nurse was instructed to describe the left calf as slightly swollen in case the participant chose to examine the leg. At the end of the study time of five min, the senior physician entered the room, announced that he required the participant to temporarily leave his/her patient to help with another emergent case, and asked the participants on their presumptive diagnosis.

2.3 Randomization

The scenario was a simulated pulmonary embolism. Participants were randomized (computer-generated numbers) to six versions of the scenario. The six versions differed only in the initial cues, *i.e.*, in the reply of the patient to the initial question about the reason for his visit. In three out of the six versions, initial cues were restricted to thoracic symptoms: chest pain (version 1); dyspnoea (version 2); and chest pain and dyspnoea (version 3). In the remaining three versions, initial cues consisted of a combination of thoracic and extra-thoracic symptoms: chest pain, dyspnoea, and leg pain (version 4); chest pain, dyspnoea, and immobilisation (version 5); chest pain, dyspnoea, leg pain, and immobilisation (version 6). Apart from these differences in initial cues the scenarios were identical and in all versions all questions during history taking were answered using the same checklist.

2.4 Data analysis

Data analysis was performed using video-recordings performed during scenarios. Trained observers noted kind and timing of all questions and measures. Questions were categorized in six pre-defined categories (table). Depth of information gathering was assessed for each category. The three problem areas chest pain, dyspnea, and leg pain or leg swelling were defined as relevant and, accordingly, breadth of information gathering was defined as number of relevant problem areas identified (possible range 0–3). Systematics of history taking was assessed as follows: "logical sequence questions" were defined as questions that were either immediately preceded or immediately followed by a question in the same category; "jump-back questions" were defined as questions separated from a previous question from the same category, by one or more questions from a different category (*e.g.*, in a timely sequence of questions relating to the onset of chest pain, diabetes, dyspnea, and quality of chest pain, the latter would be rated as "jump-back question" in the category "chest pain").

TABLE 1. Checklist of pre-defined patient's answers to question asked during history taking.

Questions of the Participants	Answers of the Patient
Category "Chest pain"	
Presence of pain	Present and main problem
Quality of pain	Stabbing, like a knife
Pain intensity	Intense; if a pain score is asked: 7–8 out of 10
Time of Onset	30 minutes prior to arrival in the ER
Description of onset	Sudden
Pain Localization	Left side of the chest
Pain Radiation	No radiation
Circumstances	Sudden begin while sitting at a desk, working
Trigger, cause or provocation	No trigger factors or provocation
Modifying (aggravating or alleviating) factors	Breathing dependent: Lessened by shallow breaths, intensified by deep breaths
Previous experience	None; first time experience
Category "Dyspnea"	
Time of onset	30 minutes prior to arrival in the ER
Trigger, cause or provocation	None
Modifying (aggravating or alleviating) factors	Lessened by bed rest; intensified by activity
Intensity	Intense
Cough	Not present
Emission	Not present
Tachypnea	Yes
Category "Associated symptoms"	
Nausea or Vomiting	Not present
Sweating	Not present
Fever	Not present
Fear	Worried
Category "Risk factors for pulmonary embolism"	
Presence of leg pain	Slight pain present in left calf
Leg swelling	Slight swelling of left calf
Pain intensity	Mild; if a pain score is asked: 3 out of 10
Previous Trauma	Ankle trauma during sport two weeks ago
Thrombosis prophylaxis	None
Immobility	Yes, cast immobilization over 11 days
Presence of cancer	Not present
B-Symptoms	No
Long journey	No
Previous venous thrombosis	No
Previous pulmonary embolism	No
Category "Cardiovascular risk factors"	
Smoking	One pack of cigarettes a day for 30 years
Hypertension	No, normal according to GP
Diabetes	No
Dyslipidemia	No, according to GP
Family history	Mother died of stroke at 70, father 82 with light dementia, 2 healthy brothers
Category "Personal and social history"	
Alcohol consumption	Not excessive
Other illnesses	None
Allergies	None
Current medication	None
Drugs	None
Last GP Visit	Yearly checkup almost a year ago
Age	49 years old
Next of kin	Married, no children
Profession	Accountant

ER: Emergency room; GP: General practitioner.

2.5 Outcomes

The primary endpoint was diagnostic accuracy, defined as stating the correct presumptive diagnosis of pulmonary embolism when asked at the end of the scenario by the senior physician or previously revealing that diagnosis to the patient and/or the nurse. Mentioning the correct diagnosis of pulmonary embolism as possible differential diagnosis to the senior physician, the patient, or the nurse was also rated as correct diagnosis. Secondary outcomes included the effects of cues, breadth of information gathering, questions' category, questions' type and systematic history taking on diagnostic accuracy.

2.6 Statistical analysis

Data is presented as means ± SD unless otherwise stated. Statistical analyses were performed using SPSS (version 28.0.1; IBM, Armonk, NY, USA). ANOVA (analysis of variance), chi-squared test and *t*-test were applied as appropriate. Univariate linear regression was used to test the predictive effects of questions' type, questions' category, and systematic history taking on diagnostic accuracy. Variables with a $p < 0.05$ were explored using forward stepwise multivariate regression analysis. A p-value < 0.05 was considered to represent statistical significance.

3. Results

198 medical students (135 females, 63 males) were included in the trial and completed the scenario. No protocol violations occurred so that data from all 198 participants were included in the final analysis. A correct presumptive diagnosis was made by 87 (44%) participants. Alternative presumptive diagnoses were acute coronary syndrome (42%), other diagnoses (10%), and "don't know" (4%) respectively.

3.1 Effect of initial cues

Table 2 displays the results according to the allocated cues. Participants allocated to both extra-thoracic and thoracic cues were significantly more likely to provide a correct presumptive diagnosis than participants allocated to thoracic cues only (52/96 *vs.* 35/102, $p = 0.006$). Moreover, extra-thoracic cues increased the percentage contribution of question in the categories "risk factors PE" (6 ± 10% *vs.* 13 ± 13%, $p < 0.001$) and decreased the percentage contribution of question in the categories "cardiovascular risk factors" (11 ± 13% *vs.* 7 ± 9%, $p < 0.001$).

Compared to thoracic cues the combination of thoracic and extra-thoracic cues significantly ($p < 0.001$) increased the breadth of information gathering (Table 2). In univariate analysis, the breadth of information gathering was positively predictive of diagnostic accuracy (regression coefficient 0.29, $p < 0.001$). Moreover, participants identifying all three relevant problem areas (chest pain, dyspnea, leg pain or swelling) had a significantly higher diagnostic accuracy (41/57 *vs.* 46/141, $p < 0.001$) than participants identifying two or less relevant problem areas.

3.2 Effect of questions' category

All question asked were within the frame foreseen by our checklist. The number and category of questions asked are displayed in Table 2. The total number of questions asked did not predict a correct presumptive diagnosis (regression coefficient 0.11, $p = 0.11$).

In univariate regression, the depth of information gathering, *i.e.*, the number of questions asked in the categories "risk factors for pulmonary embolism" (regression coefficient 0.35, $p < 0.001$) and "dyspnea" (regression coefficient 0.36, $p < 0.001$) were positively predictive; the number of questions asked in the categories "cardiovascular risk factors" (regression coefficient −0.22, $p = 0.02$) was negatively predictive; the number of questions asked in the category "chest pain" (regression coefficient −0.13, $p = 0.06$) was negatively predictive with a borderline significance; and the number of questions in the category "personal/social (regression coefficient −0.06, $p = 0.42$) had no predictive value. In multivariate regression both "risk factors for pulmonary embolism" (regression coefficient 0.15, $p < 0.001$) and "dyspnea" (regression coefficient 0.12, $p < 0.001$) independently predicted a correct presumptive diagnosis.

3.3 Effect of questions' type

In univariate regression including all types of questions asked during history taking, six question significantly predicted diagnostic accuracy: "chest pain" (regression coefficient 0.15, $p = 0.04$), "chest pain intensity" (regression coefficient −0.22, $p = 0.002$), "previous experience of chest pain" (regression coefficient −0.16, $p = 0.02$), "modifying factors of chest pain" revealing respiratory dependency (regression coefficient 0.21, $p = 0.003$), "onset of dyspnea" revealing acute onset of dyspnea (regression coefficient 0.26, $p < 0.001$), and immobilization (regression coefficient 0.33, $p < 0.001$). In multivariate regression "immobilization" (regression coefficient 0.42, $p < 0.001$), "onset of dyspnea" (regression coefficient 0.23, $p = 0.003$), and "modifying factors of chest pain" (regression coefficient 0.20, $p = 0.04$) independently predicted a correct presumptive diagnosis.

3.4 Effect of systematic history taking

There was no difference between groups with thoracic cues only and groups with thoracic and extra-thoracic cues with regard to number of "logical sequence questions" and "jump-back" questions. In univariate analysis the percent contribution of "logical sequence questions" over all categories of questions (regression coefficient −0.27, $p < 0.001$) and the percent contribution of "logical sequence questions" in the category "chest pain" (regression coefficient −0.14, $p = 0.047$) negatively predicted diagnostic accuracy while the percent contribution of "jump-back" questions (regression coefficient −0.23, $p < 0.001$) positively predicted diagnostic accuracy. In multivariate analysis only "logical sequence questions" over all categories of questions persisted as predictor, albeit negative, of diagnostic accuracy.

TABLE 2. Results according to allocated cues.

	Thoracic cues			Thoracic and extra-thoracic cues		
	Chest pain (n = 36)	Dyspnea (n = 30)	Chest pain & Dyspnea (n = 36)	Chest pain, dyspnea & leg pain (n = 36)	Chest pain, dyspnea, leg pain & immobilization (n = 30)	Chest pain, dyspnea, & immobilization (n = 30)
Female:Male	27:9	23:7	27:9	21:15	20:10	17:13
Correct diagnosis (n)	11	8	11	16	14	27*
Number of questions asked (n)	9 ± 3	8 ± 3	9 ± 3	8 ± 4	8 ± 4	9 ± 3
Category chest pain (%)	58 ± 16	45 ± 21	59 ± 20	53 ± 17	51 ± 23	45 ± 24*
Category dyspnea (%)	5 ± 7	14 ± 14	4 ± 7	14 ± 9	13 ± 12	18 ± 16
Category risk factors for PE (%)	4 ± 9	4 ± 9	5 ± 10	10 ± 11	17 ± 11†	19 ± 14†
Category cardiovascular risk factors (%)	10 ± 12	15 ± 15	12 ± 12	7 ± 11	4 ± 6	2 ± 4
Category personal and social (%)	22 ± 17	20 ± 14	20 ± 15	14 ± 13	13 ± 13	13 ± 14
Systematic questions (%)	83 ± 16	68 ± 18	82 ± 19	58 ± 22	63 ± 20	57 ± 20
Jump-back questions (%)	7 ± 8	19 ± 14	9 ± 10	18 ± 15	16 ± 14	21 ± 14
Breadth of history taking (range 0–3)	1.5 ± 0.6	1.8 ± 0.6	1.5 ± 0.7	2.4 ± 0.6†	2.5 ± 0.7†	2.5 ± 0.5†

**: $p < 0.05$ vs. all other groups; †: $p < 0.05$ vs. groups with thoracic cues only. PE: pulmonary embolism.*

4. Discussion

The present trial demonstrates that during medical history taking cues initially revealed by the patient as well as the physician-led interview's breadth of information gathering, kind and category of questions asked, and structural systematics affect diagnostic accuracy.

Despite the presence of validated diagnostic algorithms [14–16], the diagnosis of pulmonary embolism is frequently missed or delayed [9, 10, 12]. Symptoms of PE are heterogeneous and patients may vary in their selection and expression of presenting complaints. In the present trial, the symptoms revealed by the patient in answering the physician's opening question of why he is seeking medical help were experimentally controlled and termed initial cues [3]. There are only limited data on the role of initial cues on diagnostic accuracy in PE or other diseases. Previous work demonstrated that wrong initial cues can induce diagnostic errors [5]. Moreover, the initial cue of a history of mental disorder wrongly affects physicians' estimates of the probability of a potentially fatal disease and the likelihood of further diagnostic evaluation [4]. The present trial is the first to report that the number of initial cues did not affect diagnostic accuracy. However, the combination of thoracic and extra-thoracic cues resulted in an increased breadth of history taking and a higher diagnostic accuracy than thoracic cues alone. Positive cue validity and cue consensus [3, 17] of the combination of thoracic and extra-thoracic cues may thus be regarded as high. However, the diagnostic accuracy was not uniformly high in all groups confronted with thoracic and extra-thoracic cues. While the cue "leg pain" or "leg pain" in combination with "immobilization" did not significantly improve diagnostic accuracy, the cue immobilization itself apparently triggered suspicion of pulmonary embolism. This may suggest that the cue "leg pain" blurred the value of the cue "immobilization" or served as a distraction. Thus, the positive or negative validity of cues might depend on the context of concurrent cues. In keeping with the present results, previous studies reported that physicians' awareness of immobilization is associated with less delayed and missed diagnoses of PE [12, 18]. Moreover, in a recent meta-analysis, the overall diagnostic value of signs and symptoms of deep venous thrombosis was found to be ambiguous and associated with both increased and decreased delays in diagnosing PE [12].

After the patient has revealed his/her reason for seeking medical help (*i.e.*, the initial cues) every piece of information gathered during the following assessment can be regarded as further cue. In a case vignette study involving physicians with differing experience and medical students, lung scan results and leg examination were weighted most heavily in making the diagnosis of pulmonary embolism [17]. In a computer case simulation medical student's information-gathering pattern, including history, physical examination, and ancillary testing, distinct patterns of information-gathering were identified that

differed in the rates of diagnostic error [19]. In the present trial focusing on history taking only, the total number of questions asked, equaling the total number of cues obtained, did not affect diagnostic accuracy. Breadth of information gathering [6], *i.e.*, identifying the relevant problem areas significantly predicted diagnostic accuracy. Moreover, the depth of information gathering [6] *i.e.*, the number of questions asked in the two categories "risk factors for pulmonary embolism" and "dyspnea" independently predicted diagnostic accuracy. In addition, three distinct questions related to immobilization, the (acute) onset of dyspnea, and the modifying factors of the chest pain (revealing respiratory dependency) independently predicted the correct presumptive diagnosis of pulmonary embolism. These findings are in keeping with previous studies reporting a beneficial effect of the symptoms immobilization [11, 12, 18] and pleuritic chest pain [9, 11] on diagnostic accuracy and timeliness in PE. However, contradictory findings were reported for the diagnostic value of the symptom dyspnea [9, 11, 12]. Our results suggest that rather than performing a complete and exhaustive interview, identifying the relevant problem areas, sufficiently exploring the relevant thematic categories and asking the very few decisive questions appears to be key for diagnostic accuracy. Most likely, both relevant thematic categories and decisive questions are specific for any given disease.

Previous studies have shown the value and high diagnostic yield of expert history taking in other entities, such as syncope, where similarly decisive questions and focus on select thematic categories is associated with higher diagnostic accuracy [20]. Accordingly, we expected that a complete and structured medical history would have a positive correlation with diagnostic accuracy, presupposing that a greater yield of information would increase the likelihood of suspecting PE or conversely decrease the likelihood of suspecting differential diagnoses. However, in the present trial more systematic history taking resulted in a lower diagnostic accuracy. A possible explanation for this finding may be that participants hypothesizing an acute coronary syndrome as presumptive diagnosis performed a thorough and structured cardiac interview. However, as our trial focused on one disease only, it remains unclear whether this finding is generalizable to other conditions or specific for pulmonary embolism only. Further research is therefore necessary to investigated the effects of the systematics of history taking on diagnostic accuracy in different diseases.

Limitations of the present trial are the single-centre design and limitations inherent in simulator-based studies like the absence of real patients and real environment. However, simulation allowed for highly standardised conditions for a large number of participants, which would be very difficult to achieve in real life. Moreover, the effects of initial cues are very difficult to assess in real cases. As symptoms of our scenario like leg swelling are difficult to role-play, using standardized patients as alternative to a simulator mannequin would have offered only little added value. Simulations with manikins enhanced by augmented reality may provide a further degree of realism in depiction of symptoms, though their value may be in simulations where the focus lies more on clinical presentation than in medical history taking, with select symptoms on the manikin itself as an adjunct [21, 22]. Our participants were medical students which may be regarded as further limitation. The present trial assessed history taking during the initial assessment of a patient presenting with a single disease and the results can therefore not necessarily be generalized to constellations of other diseases. Current knowledge on diagnostic errors results largely from retrospective error analysis and observational studies [23]. This research depends on errors being identified as such and is thus prone to hindsight bias and under-detection. The present trial is one of the very few, where the development of errors or the lack thereof is prospectively followed. This is a particular strength and demonstrates the feasibility of prospective randomized trials on medical errors in the simulator settings.

Our trail has several implications: first, the observation that initial complaints affect diagnostic accuracy is a novel finding. All information available in a patient encounter prior to the start of the physician-led history taking can be regarded as an initial cue. Prior work demonstrated that an initial cue of a wrong à priori diagnosis [5] or of a pre-existing psychiatric history [4] can induce a diagnostic error. The present trial appends to this knowledge by demonstrating that initially revealed patients' complaints are cues that may lead to diagnostic errors too. Second, initial cues have so far received little attention in both research and medical education. Further research is necessary to investigate the role of initial cues in different diseases and develop and teach countermeasures to mitigate their effects on diagnostic errors. Third, our findings highlight that the quality of medical history taking is related to diagnostic errors. Though history taking is usually the first step in patients' assessment there is little evidence on how to best conduct an interview to ensure patients' safety. This knowledge gap should be addressed by future research, followed by implementation in medical education.

5. Conclusions

This prospective randomized trial assessed diagnostic accuracy in a simulated scenario of acute pulmonary embolism. Initial cues, *i.e.*, the patient's initial complaints revealed in answering the physician's opening question of why he is seeking medical help, affected diagnostic accuracy. Regarding the quality of medical history taking, breadth and depth of information gathering independently predicted diagnostic accuracy as did asking three distinct questions related to immobilization, the onset of dyspnea, and the modifying factors of the chest pain. Counterintuitively, a more systematic pattern of history taking was associated with a lower diagnostic accuracy. As history taking plays a dominant role in patients' care and prevention of diagnostic errors, further research should focus on how to deal with initial cues and on interview patterns.

AVAILABILITY OF DATA AND MATERIALS

Video recordings are strictly confidential and cannot be shared. Anonymised data can be shared on reasonable request from the corresponding author.

AUTHOR CONTRIBUTIONS

SM—designed the research study. SG, JF and KT—performed the research. SG, SM and TS—analyzed the data. SG and SM—wrote the manuscript. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The trial was approved by the regional ethical committee (Ethikkommission Nordwestschweiz; EKNZ 85/04) and all participants gave written informed consent.

ACKNOWLEDGMENT

Not applicable.

FUNDING

This research received no external funding.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] Institute of Medicine and National Academies of Sciences, Engineering and Medicine. Improving diagnosis in health care. The National Academies Press: Washington DC. 2015.

[2] Newman-Toker DE, Wang Z, Zhu Y, Nassery N, Saber Tehrani AS, Schaffer AC, *et al*. Rate of diagnostic errors and serious misdiagnosis-related harms for major vascular events, infections, and cancers: toward a national incidence estimate using the “Big Three”. Diagnosis. 2021; 8: 67–84.

[3] Hertwig R, Meier N, Nickel C, Zimmermann P, Ackermann S, Woike JK, *et al*. Correlates of diagnostic accuracy in patients with nonspecific complaints. Medical Decision Making. 2013; 33: 533–543.

[4] Solmi M, Fiedorowicz J, Poddighe L, Delogu M, Miola A, Høye A, *et al*. Disparities in screening and treatment of cardiovascular diseases in patients with mental disorders across the world: systematic review and meta-analysis of 47 observational studies. American Journal of Psychiatry. 2021; 178: 793–803.

[5] Meyer FML, Filipovic MG, Balestra GM, Tisljar K, Sellmann T, Marsch S. Diagnostic errors induced by a wrong a priori diagnosis: a prospective randomized simulator-based trial. Journal of Clinical Medicine. 2021; 10: 826–839.

[6] Setrakian J, Gauthier Gv, Bergeron L, Chamberland M, St-Onge C. Comparison of assessment by a virtual patient and by clinician-educators of medical students’ history-taking skills: exploratory descriptive study. JMIR Medical Education. 2020; 6: e14428.

[7] Hautz WE, Kämmer JE, Hautz SC, Sauter TC, Zwaan L, Exadaktylos AK, *et al*. Diagnostic error increases mortality and length of hospital stay in patients presenting through the emergency room. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine. 2019; 27: 54.

[8] Kunitomo K, Harada T, Watari T. Cognitive biases encountered by physicians in the emergency room. BMC Emergency Medicine. 2022; 22: 148–155.

[9] Hendriksen JMT, Koster-van Ree M, Morgenstern MJ, Oudega R, Schutgens REG, Moons KGM, *et al*. Clinical characteristics associated with diagnostic delay of pulmonary embolism in primary care: a retrospective observational study. BMJ Open. 2017; 7: e012789.

[10] Mansella G, Keil C, Nickel CH, Eken C, Wirth C, Tzankov A, *et al*. Delayed diagnosis in pulmonary embolism: frequency, patient characteristics, and outcome. Respiration. 2020; 99: 589–597.

[11] Swan D, Hitchen S, Klok FA, Thachil J. The problem of under-diagnosis and over-diagnosis of pulmonary embolism. Thrombosis Research. 2019; 177: 122–129.

[12] van Maanen R, Trinks-Roerdink EM, Rutten FH, Geersing GJ. A systematic review and meta-analysis of diagnostic delay in pulmonary embolism. European Journal of General Practice. 2022; 28: 165–172.

[13] Cheng A, Kessler D, Mackinnon R, Chang TP, Nadkarni VM, Hunt EA, *et al*. Reporting guidelines for health care simulation research. Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare. 2016; 11: 238–248.

[14] Le Gal G, Righini M, Roy P, Sanchez O, Aujesky D, Bounameaux H, *et al*. Prediction of pulmonary embolism in the emergency department: the revised geneva score. Annals of Internal Medicine. 2006; 144: 165.

[15] Anderson D, Rodger M, Ginsberg J, Kearon C, Gent M, Turpie A, *et al*. Derivation of a simple clinical model to categorize patients probability of pulmonary embolism: increasing the models utility with the SimpliRED D-dimer. Thrombosis and Haemostasis. 2000; 83: 416–420.

[16] van der Hulle T, Cheung WY, Kooij S, Beenen LFM, van Bemmel T, van Es J, *et al*. Simplified diagnostic management of suspected pulmonary embolism (the YEARS study): a prospective, multicentre, cohort study. The Lancet. 2017; 390: 289–297.

[17] Wigton RS, Hoellerich VL, Patil KD. How physicians use clinical information in diagnosing pulmonary embolism. Medical Decision Making. 1986; 6: 2–11.

[18] Smith SB, Geske JB, Morgenthaler TI. Risk factors associated with delayed diagnosis of acute pulmonary embolism. The Journal of Emergency Medicine. 2012; 42: 1–6.

[19] Delzell JE, Chumley H, Webb R, Chakrabarti S, Relan A. Information-gathering patterns associated with higher rates of diagnostic error. Advances in Health Sciences Education. 2009; 14: 697.

[20] Sutton R, van Dijk N, Wieling W. Clinical history in management of suspected syncope: a powerful diagnostic tool. Cardiology Journal. 2014; 21: 651–657.

[21] Alcazar Artero PM, Pardo Rios M, Greif R, Ocampo Cervantes AB, Gijon-Nogueron G, Barcala-Furelos R, *et al*. Efficiency of virtual reality for cardiopulmonary resuscitation training of adult laypersons: a systematic review. Medicine. 2023; 102: e32736.

[22] Zackoff MW, Cruse B, Sahay RD, Fei L, Saupe J, Schwartz J, *et al*. Development and implementation of augmented reality enhanced high-fidelity simulation for recognition of patient decompensation. Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare. 2021; 16: 221–230.

[23] Zwaan L, Monteiro S, Sherbino J, Ilgen J, Howey B, Norman G. Is bias in the eye of the beholder? A vignette study to assess recognition of cognitive biases in clinical case workups. BMJ Quality & Safety. 2017; 26: 104–110.

How to cite this article: Stefan Götz, Julie Fournier, Kai Tisjlar, Timur Sellmann, Stephan Marsch. Effects of the quality of medical history taking on diagnostic accuracy. Signa Vitae. 2023; 19(5): 68-74. doi: 10.22514/sv.2023.081.