University of Groningen

What Can We Learn from In-Depth Analysis of Human Errors Resulting in Diagnostic Errors in the Emergency Department

Baartmans, Mees C.; Hooftman, Jacky; Zwaan, Laura; Van Schoten, Steffie M.; Erwich, Jan Jaap H.M.; Wagner, Cordula

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in University of Groningen/UMCG research database](#)

What Can We Learn From In-Depth Analysis of Human Errors Resulting in Diagnostic Errors in the Emergency Department: An Analysis of Serious Adverse Event Reports

Mees C. Baartmans, MSc,* Jacky Hooftman, MSc,†‡ Laura Zwaan, PhD,‡ Steffie M. van Schoten, PhD,†
Jan Jaap H.M. Erwich, MD, PhD,§ and Cordula Wagner, PhD*†

Introduction: Human error plays a vital role in diagnostic errors in the emergency department. A thorough analysis of these human errors, using information-rich reports of serious adverse events (SAEs), could help to better study and understand the causes of these errors and formulate more specific recommendations.

Methods: We studied 23 SAE reports of diagnostic events in emergency departments of Dutch general hospitals and identified human errors. Two researchers independently applied the Safer Dx Instrument, Diagnostic Error Evaluation and Research Taxonomy, and the Model of Unsafe acts to analyze reports.

Results: Twenty-one reports contained a diagnostic error, in which we identified 73 human errors, which were mainly based on intended actions (n = 69) and could be classified as mistakes (n = 56) or violations (n = 13). Most human errors occurred during the assessment and testing phase of the diagnostic process.

Discussion: The combination of different instruments and information-rich SAE reports allowed for a deeper understanding of the mechanisms underlying diagnostic error. Results indicated that errors occurred most often during the assessment and the testing phase of the diagnostic process. Most often, the errors could be classified as mistakes and violations, both intended actions. These types of errors are in need of different recommendations for improvement, as mistakes are often knowledge based, whereas violations often happen because of work and time pressure. These analyses provided valuable insights for more overarching recommendations to improve diagnostic safety and would be recommended to use in future research and analysis of (serious) adverse events.

Key Words: diagnostic error, diagnostic safety, human error, adverse event, root cause analysis, emergency department

From the *Nivel, Netherlands Institute for Health Services Research, Utrecht; †Department of Public and Occupational Health, Amsterdam Public Health Research Institute, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam; ‡Institute of Medical Education Research Rotterdam (iMERR), Erasmus Medical Centre, Rotterdam; and §Department of Obstetrics and Gynecology, University Medical Centre Groningen, University of Groningen, Groningen, the Netherlands.
Correspondence: Mees Baartmans, MSc, Otterstraat 118-124, 3513 CR Utrecht, the Netherlands; P/O Box 1568, 3500 BN Utrecht, the Netherlands (e-mail: m.baartmans@nivel.nl).

Research on diagnostic error has gained interest since the report "Improving diagnosis in health care" of the National Academies of Science, Engineering, and Medicine was published in 2015.[1] The increase in attention for these errors is not without reason. Diagnostic errors imply a great risk for patient safety[2]; it is estimated that most people will experience a diagnostic error in their lifetime.[1] Diagnostic errors are defined by Graber et al[3] (2005) as a diagnosis that was unintentionally delayed, wrong, or missed, as judged from the eventual appreciation of more definitive information. Consequences of diagnostic errors are often more severe (i.e., higher mortality rates) and more often considered preventable than other types of errors.[4,5] Studies in the Netherlands show that diagnostic errors are causative for a large proportion of adverse events in healthcare.[5,6]

The emergency department (ED) is especially prone to diagnostic errors. It is a decision-dense department where many diagnostic decisions have to be made, often under time pressure and with high levels of uncertainty.[7–9] A study into the causes of diagnostic errors in the ED, using closed malpractice claims, identified multiple different factors that contributed.[10] In almost all cases (96%), at least one cognitive factor was present, such as mistakes in judgment, lack of knowledge, and lapses in vigilance or memory. Given this large contribution of cognitive factors, it is important to gain more knowledge about their nature.

Most studies that are focused on diagnostic error make use of retrospective data from patient records. Although these records can be useful for finding diagnostic errors, they lack information on the thought processes of the clinician. Interpretation of the cognitive processes based on the available information in patient records is difficult.[11] For example, a situation in which a physician decides not to order a certain laboratory test because they thought it was irrelevant is very different from a situation in which the laboratory test was not ordered because the physician forgot to do so and calls for a different improvement measure. These types of nuances are difficult to extract from patient record review. To gain insight in those cognitive processes and thus reasons for certain diagnostic decisions or diagnostic errors, more detailed information from the physicians themselves should be gathered, for example, through interviews. A combination of clinical data and interviews with involved physicians has proven to be useful to gather more information on clinical thought processes.[12] These insights could help with suggesting recommendations for improvement.

In the current study, we have used data from serious adverse event (SAE) reports, which are based on extensive event investigation and include details of decisions made by the involved clinicians gathered through interviews. These reports are supposed to be a rich source of information and more suitable for subtracting information on the cognitive thought processes. For the current study, we focused on human errors—in which cognitive factors play an essential role—because these have a large contribution

in diagnostic errors,[10] although we do recognize that these errors are commonly facilitated by other factors (e.g., system factors).

We have used 3 well-known and established methods for analyzing the reports. For identification of reports that contained diagnostic error(s), we used the Safer Dx Instrument.[13] We used the Diagnostic Error Evaluation and Research Taxonomy (DEER taxonomy) to gain insight into where human errors occur in the diagnostic process.[14,15] Lastly, we used the Model of Unsafe Acts to determine the types of human errors. This last model is based on a well-known theory on human errors described by Reason.[16,17] It differentiates between decisions that were made intentionally and unintentionally. The addition of this model could provide a more in-depth understanding of human errors. Although these methods are well known, their combination has not yet been used in diagnostic error research nor for analysis of SAEs.

Applying these instruments, we aim to determine whether an in-depth analysis of the human errors involved in diagnostic errors in the ED, using information-rich SAE reports, will increase our understanding of diagnostic errors and will provide leads to improve analysis of diagnostic SAEs and specify recommendations for improving diagnostic safety.

METHODS

Serious Adverse Event Reports

The current study was part of a bigger project including an aggregate analysis of SAEs. For this project, a group of cooperating Dutch general hospitals (n = 28) were asked to share recent SAE reports (2017–2019) related to specific safety critical themes.[18] One of these themes was the diagnostic process in the ED. In the current study, we used the reports that were submitted under this theme (n = 23). Because the reports were anonymized, it is unknown how many of the participating hospitals are included in this sample.

These reports are based on a thorough investigation performed by a multidisciplinary independent hospital committee including physicians with various specialties (e.g., emergency physicians, radiologists, cardiologists), other clinicians (e.g., nurses, pharmacists, surgery assistants), and quality and safety officers.[19] These teams are trained in performing root cause analyses. The committee interviewed clinicians involved in the SAE, studied patient records, and analyzed relevant documentation and test results. If necessary, they consulted an external expert. The investigation ended with an analysis of root causes and formulation of measures for improvement. Tools for the analyses of root causes that hospitals used varied per hospital but was either one or a combination of the following 3 conventional tools: Prevention and Recovery Information System for Monitoring and Analysis,[20] Systemic Incident Reconstruction and Evaluation[21] or Tripod Beta.[22] Although hospitals used various methods for root cause analysis, all reports included: general patient information (sex, age, relevant medical history), information about the SAE (date, time, healthcare professionals involved, description of provided care, contributing sociotechnical factors, and relevant events), a description of root causes (including a classification of identified root causes or failing barriers), and formulated measures for improvement. Reports were written according to the standards of the Dutch Health and Youth Care Inspectorate, which evaluates the content and quality of the reports.[19,23]

Review Process

We further analyzed the reports as provided by the hospitals using established methods for analyzing diagnostic (specifically, Safer Dx Instrument and DEER taxonomy) and human errors (Model of Unsafe Acts). The review process consisted of 4 steps as illustrated in Figure 1.

Safer Dx Instrument

First, all reports were checked for the presence of a diagnostic error as operationalized by the Safer Dx Instrument.[13] This screening tool was developed to improve the accuracy of assessing diagnostic errors based on objective criteria. The instrument has been validated in the primary care setting and the pediatric intensive care unit[24] but has recently been revised to be used in other settings as well.[13] The revised instrument can be used by either researchers, safety professionals, or clinicians.[13] It consists of 12 items (see Appendix 1, http://links.lww.com/JPS/A473), each targeting different aspects of the diagnostic process. The items are judged on a 7-point scale (1 = strongly disagree, 4 = neutral, 7 = strongly agree). A final 13th item addresses whether there was a "missed opportunity to make a correct and timely diagnosis." Only SAE reports that were judged 4 and greater on the overall score were considered diagnostic errors and were further analyzed. The purpose of using the Safer Dx Instrument was to filter out any reports that were not diagnosis related. Researchers (M.C.B. and J.H.) independently reviewed the reports and scored the 13 items. The researchers agreed on all cases on whether the report related to diagnostic error and should be included.

Identifying Human Errors

The SAE reports that involved a diagnostic error were further analyzed for human errors. For identifying these human errors, we focused in particular on the descriptions of root causes. We searched these descriptions for "errors" mainly induced by the actions (or reluctance to act) of a person. When a human error including the underlying actions was already specified in the description of root causes, we adopted this description for our analysis. In some cases, we needed to go back to the detailed description of the event to specify the human error and find the corresponding acts. An example of an identified human error is: "a resident fails in reading the electrocardiogram (ECG) correctly and misses a heart failure diagnosis." Missing the diagnosis is caused by the failure of the resident to accurately assess the ECG. The report specified that the resident evaluated the ECG but interpreted it incorrectly.
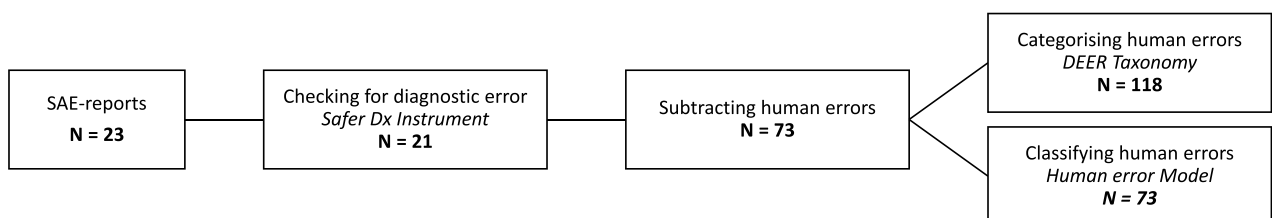

FIGURE 1. Flowchart of the review process. Resulting numbers are represented at each step.

Diagnostic Error Evaluation and Research Taxonomy

The identified human errors (n = 73) were then categorized using the DEER taxonomy (see Appendix 2, http://links.lww.com/JPS/A473).[14,15] This taxonomy helps systematically aggregate diagnostic error cases and reveal patterns of diagnostic failures and areas for improvement. We used the DEER taxonomy to study in which part of the diagnostic process (access to care, history taking, physical exam, diagnostic testing, assessment, consultation/referral, or follow-up) the human errors occurred and specify what exactly went wrong. The DEER taxonomy–based categorization of the human errors was independently executed by 2 of the researchers (M.C.B. and J.H.). They indicated for all 73 identified human errors at which step in the diagnostic process the error occurred and what exactly went wrong (e.g., failure/delay in eliciting a critical piece of history data, error in clinical interpretation of test). Because some categories of the DEER taxonomy overlap or are closely related (e.g., too little weight given to the diagnosis and too much weight on competing diagnosis), a single human error could apply to multiple (sub)categories of the DEER. A third reviewer (L.Z.) was available when consensus could not be met. This was needed for categorizing 9 of the human errors.

Model of Unsafe Acts

Lastly, the identified human errors were classified using the Model of Unsafe Acts by Reason.[16] This model has been occasionally used in health care settings,[25] as well as in studies on the diagnostic reasoning process.[12] The model classifies acts leading to human errors as either intentional or unintentional actions (Fig. 2). It is important to emphasize that this model classifies acts, not the outcomes of acts. The SAE-related outcomes are generally unintended, but the actions causing these outcomes can still be intended. For example, when a nurse decides to not adhere to an infection prevention protocol because compliance takes too much time and the work pressure is high, then, the action of not disinfecting hands is intended, while the outcome (e.g., infecting a fragile patient) is unintended.

Intended actions can be subdivided into mistakes and violations. A mistake occurs if a plan is performed as intended, but the plan was not adequate to reach the outcome that was intended (rule-based mistakes: e.g., misapplication of good rule, application of bad rule, and knowledge-based mistakes). A violation in this context does not imply malicious intentions but is an action that is not in line with the protocols, guidelines, or rules (e.g., routine violations, exceptional violations, efficiency-thoroughness trade-off). An example of a violation is a trade-off between doing your job fast (using erroneous shortcuts) and doing it thoroughly, also known as "efficiency-thoroughness trade-offs."[26]

Unintended actions can be subdivided into lapses and slips, which are related to errors in execution. Slips occur when the correct action is executed poorly (attentional failures: e.g., intrusion, omission, reversal, misordering, mistiming). Lapses occur when the execution involves a failure in memory (e.g., omitting planned items, place-losing, forgetting intentions).[16,27]

All identified human errors (N = 73) were classified by 2 researchers (M.C.B. and J.H.) using this model. Whenever the researchers disagreed on the most appropriate class, a third researcher (L.Z.) was asked to assess the error (n = 18). Furthermore, a 20% sample of the other errors was assessed by the third researcher (n = 12). In case of a discrepancy between the assessments (n = 3), we discussed motivations to reach consensus.

RESULTS

After screening all 23 reports with the Safer Dx Instrument, 2 reports were excluded because the researchers agreed that there was no missed opportunity for a correct and timely diagnosis (judgment on the overall question less than a score of 4). The remaining 21 reports (a Safer Dx Instrument's 13th-item mean score of 6.19) were included in the further analysis.

Serious Adverse Events Characteristics

Most prevalent missed or delayed final diagnoses were cardiovascular (n = 11) or neurological (n = 5) conditions, such as aorta dissection, ruptured abdominal aortic aneurysm, subarachnoid hemorrhage, and spinal cord injury. Initial diagnosis varied widely. An overview of all initial and final diagnoses is presented in Table 1. This table further presents the sociotechnical factors that were identified by the hospitals in the SAE reports, which shows that organizational and patient-related factors contributed alongside the human errors. In our further analysis, we have focused on human errors.

Human Errors in the Diagnostic Process

From the SAE reports, we subtracted 73 human errors. The DEER taxonomy tool was used to specify the errors further. One human error was not suitable for categorization with the DEER taxonomy as it involved forgetting to administer medication after diagnosis was already established and thus was not part of the diagnostic process. A single human error could relate to multiple DEER subcategories. A total of 120 DEER categorizations were made, which is an average of 1.7 categories per error (min = 1, max = 4). These occurred most often in the assessment (n = 49) and testing domain (n = 36). Other human errors occurred during referral or consultation (n = 13), history taking (n = 7), the physical exam (n = 8), follow-up (n = 5), and at access to care or presentation (n = 2).

Most human errors involved a failure or delay in recognizing the urgency of the situation (n = 14; e.g., the supervising physician is not contacted after a patient collapses in the middle of the night) or putting too much weight on a competing or coexisting diagnosis (n = 14; e.g., the working diagnosis of a migraine disrupted the search for other causes of the symptoms of a patient with a subarachnoid hemorrhage). Other recurring themes are the failure or delay to consider a diagnosis (n = 12), failed or delayed
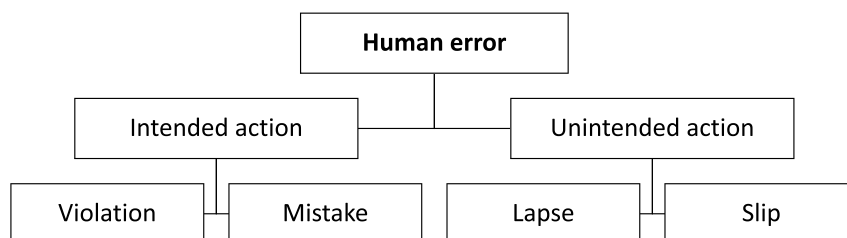

FIGURE 2. Model of Unsafe Acts.[16]

TABLE 1. Overview of the Initial and Final Diagnoses of the Included Cases With the Number of Identified DEER Classifications, the Number of Mistakes and Violations, and the Presence of Sociotechnical Factors

Initial Diagnosis	Final Diagnosis	Socio Technical Factors*				DEER Taxonomy						Human Error			
		Technical	Organizational	Patient-Related	Access	History	Physical Exam	Tests	Assessment	Referral	Follow-up	Mistakes	Violations	Lapses	Slips
Urinary tract infection	Ruptured abdominal ortic aneurysm		X				1	5	1			3	2		1
Diverticulitis	Ruptured abdominal aortic aneurysm		X	X					5			4			
Pneumonia, delirium	Brain injury		X	X		1	1		1			1	1		
Migraine	Subarachnoid hemorrhage			X	1			1	4		1	4		1	
Pulmonary embolism	Aorta dissection	X	X	X			1	4	2	2		2	3		
Pulmonary embolism	Aorta dissection, pericardial tamponade			X				5	3	2		4			
Tendomyogene pain	Aorta dissection		X	X		2			7			5			
Retention bladder	Cardiogenic shock		X					4	1	1		1	2		
Neurological injury	Cardiac infarction		X			1		1	3	1		2	1		1
Muscle strain	Heart and lung perforation, pacemaker lead		X	X			1	1	2	1		3			1
Ileus	Heart attack		X	X				1	1			1			
Tietze syndrome	Heart attack		X	X				4	2	1		4			
Fall, confusion	Myocardial infarction		X	X				1	2			2			
Pneumonia (suspected)	Unknown suspected cardiac/combination pneumonia and sleep apnea		X	X		1	1	1	1	1	1	2	1		1
Hip contusion	Fractured lumbar vertebra, incomplete spinal cord dwarslaesie		X				1		2			2			
Kidney stone	Rib fractures, lung contusion		X	X				1	2			2			
Mild brain damage	Spinal cord injury		X	X		1		2	1	1		3			
Myelopathy	Spinal fractures	X	X		1		1	2	1	2		4	1		
Rent bladder, urinary tract infection	Fournier gangrene			X				1	3	2	1	3	1		
Pneumonia	Airway obstruction, pneumonia in combination with myasthenia gravis		X	X		1	1		5	1		4			
Pericarditis	Lung carcinoma			X				2			2	1	1		
Total					2	7	8	36	49	13	5	56	13	1	3

*An "X" means that hospitals identified at least one sociotechnical factor of that type during that case.

follow-up of (abnormal) test results (n = 8), failure or delay in ordering needed tests (n = 7), wrong test orders (n = 6), and failure or delayed communication or follow-up of a consultation (n = 6). Specified results on all subcategories are summarized in Appendix 2, http://links.lww.com/JPS/A473.

Human Error Classification

The 73 identified human errors were also classified according to the Model of Unsafe Acts. Most human errors were based on intended actions and therefore classified as "intended" (n = 69). Fifty six of the intended actions were classified as mistakes and 13 as violations. Six violations seemed to be efficiency-thoroughness trade-offs (e.g., the physician only read a summary, not the full patient history file, because it is a busy night), one was related to routine-based violations (e.g., the intensivist does not immediately react to a rapid response team call made by a nurse, because they usually get those calls from physicians), the others were not further distinguishable. Unintended actions occurred less often (n = 4) and were most frequently classified as action-based errors (slips). A summary of the results, including examples of each of the human error types, are provided in Table 2.

Classified Human Errors in the Phases of Diagnostic Process

When we combine the results of the DEER taxonomy and the human error classification, we can provide an overview of the DEER categories that were found per type of human error (as presented in Table 2). For example, of the DEER categories found in mistakes (n = 99), approximately half (n = 48) were related to the

assessment phase of the diagnostic process. Human errors classified as mistakes were on average related to a higher number of DEER (sub)categories (1.9 category per mistake) than human errors classified as violations (1.3 category per violation). These insights in the specific types of human error combined with information on where in the diagnostic process they occur may provide important leads to more specifically aim recommendations to prevent recurrence.

DISCUSSION

We studied 21 diagnostic error–related SAE reports that occurred in the ED of several Dutch general hospitals to determine whether an in-depth analysis of human errors could help improve the analysis of SAEs and help aim specific recommendations for improvement.

We found that human errors in our study occurred most frequently within the assessment phase of the diagnostic process. This is congruent with earlier studies.[15,28] More specifically, they occurred because of a failure to consider the correct diagnosis, overweighing of a competing diagnosis, or a failure to recognize urgency. Apart from errors that were categorized in the assessment phase, we also identified a substantial number of errors in the testing phase, particularly in (wrongfully) ordering the needed tests, and the interpretation of test results.

Most of the human errors could be classified as intended actions (i.e., violations and mistakes). The overwhelming number of intended actions shows that actions that contributed to the diagnostic errors were deliberate decisions, such as the decision to skip reading a triage report, or focusing on a specific cause for patients'

TABLE 2. Human Error Classifications, Including an Example of Each Type of Error and How the Errors Are Related to the DEER Taxonomy Categories

| Human Error | N = 73 | Example | DEER Categories, n* | | | | | | | |
			Presentation	History	Physical Exam	Tests	Assessment	Referral	Follow-up	Total
Intended actions	69									
Violations	13	"It is busy in the ED, the physician arrives at the patient and likely makes a efficiency-thoroughness trade-off by not reading the triage report. Because of this, they miss information about fainting and high body temperature."	—	2	2	7	1	4	1	17
Mistakes	56	"The physicians focused on finding a neurological cause of the patients' symptoms. They failed to consider cardiac causes."	2	5	5	27	48	9	3	99
Unintended actions	4									
Lapses	1	"In the hectic situation after the diagnosis of a ruptured aneurysm, clinicians forget to administer vitamin K."	—	—	—	—	—	—	—	—†
Slips	3	"A nurse notes down a body temperature of 36.8°C. Later, it turns out that the temperature was 38.6. The numbers got switched accidentally."	—	—	1	2	—	—	1	4
Total			2	7	8	36	49	13	5	*120*

* A single human error could be categorized in multiple DEER subcategories, which resulted in a higher number of DEER categories than number of human errors.

†The lapse was not suitable for categorization with the DEER taxonomy.

symptoms (see examples in Table 2). Situations in which an action was unintended (i.e., lapses or slips), such as a laboratory result that was overlooked, were rare. These findings can in part be explained by the focus on diagnostic events and the nature of errors in the ED, which is a department where complex tasks are performed and many decisions are made.[9] Unintended actions (i.e., slips and lapses) are more action- and memory-based and are more often found in, for example, medication errors.[29]

Although mistakes and violations are both intended actions, their underlying nature is very different. Mistakes are often based on a lack of specific knowledge (e.g., a physician who mistakes symptoms of a rare neurological disease with pneumonia) or pursuing the wrong diagnosis (e.g., focusing on neurological symptoms when in fact there was something cardiac going on), whereas violations are more often related to organizational and contextual factors (e.g., work and time pressure, which leads to making efficiency-thoroughness trade-offs). This is important, because recommendations for improvement aimed at mistakes will likely not be effective for violations and vice versa. Mistakes could benefit from knowledge-based interventions such as specific training and education to fill in knowledge gaps, feedback and reflection on diagnostic discrepancies, or implementation of diagnostic decision support systems. Violations, on the other hand, need a more system-based approach, such as lowering work pressure and crowding in the ED, improving patient safety culture, and improving teamwork.

A thorough analysis of human errors contributing to diagnostic errors in the ED based on SAE reports for which involved clinicians were interviewed has provided important insights that help understand the role they play in diagnostic error. We would therefore recommend to perform an in-depth analysis of human errors using the DEER taxonomy and Model of Unsafe Acts whenever an SAE occurs. This can assist in further improving recommendations after SAEs occur by specifying what these should focus on to prevent recurrence of similar events. For example, in the cases we studied, most human errors were classified as mistakes and these mainly occurred in the assessment phase and during the interpretation of tests. Recommendations should then thus be aimed at these specific types of errors (i.e., knowledge-based interventions) and moments in the diagnostic process on the ED (e.g., interventions that directly support clinicians in their diagnostic reasoning process, such as a quick-response differential diagnosis generator). These may be more effective recommendations to prevent recurrence of events than those that are generally suggested after root cause analysis.

Hospitals currently typically focus their recommendations on the isolated event (i.e., generic education on the missed disease), and these are often of administrative nature (e.g., reviewing and adjusting ward specific protocols),[30–32] which are considered to be weak. Based on our experience from reading all SAE reports, we noticed that hospitals seldom propose overarching interventions to support the diagnostic process, for example, evidence-based strategies to improve specific knowledge, reforming training methods, structural feedback and reflection on diagnostic discrepancies, implementation of diagnostic decision support systems to improve diagnostic calibration, or implementing team based diagnosis. System-aimed interventions (e.g., lowering work pressure and crowding in the ED, improving patient safety culture, teamwork interventions, and cultural aspects) are rarely proposed, while these types of recommendations may have a better chance of being effective in preventing similar cases and to improve diagnostic safety.

Strengths and Limitations

The available information about SAEs in this study was particularly rich, which allowed for an in-depth analysis of the considerations and decisions of clinicians during the diagnostic process resulting in a diagnostic error. Moreover, we have shown that the combination of multiple established methods (the Safer Dx Instrument, DEER taxonomy, and Model of Unsafe Acts) can help gain insight into human errors. These methods could also be applied in future SAE investigations, which may help hospitals formulate more specifically aimed recommendations to prevent recurrence. This makes our study a valuable addition to the existing literature, as it shows how these methods help analyze human errors in SAEs and provide more specific recommendations.

Limitations of the study include limited generalizability due to a fairly small sample (n = 21) and a focus on general hospitals only. This limits drawing generalizable conclusions about the causes of diagnostic SAEs. Furthermore, although the reports were created with great care and professionalism, they were not created for this study specifically. It was not possible to revisit the involved clinicians with eventual additional questions on the decision making process.

Despite these limitations, the SAE reports have proven themselves suitable for the purposes of this study. For future research, we would recommend using a larger and wider sample to represent and compare different types of hospitals.

Conclusions

An in-depth analysis of human errors contributing to diagnostic errors in the ED provides important insights in *where* in the diagnostic process *what* types of human errors occur. This can help us aim recommendations more specifically to the types of human errors and the phase of the diagnostic process these errors occur to further improve patient safety. The current study suggests that recommendations should be specifically aimed at preventing mistakes in the assessment phase.

Serious adverse event reports have proven to be a useful source to study human errors. Ideally, the human error analysis is adopted into the initial SAE investigation and included in a broader analysis of the event also focusing on other sociotechnical factors.

REFERENCES

1. National Academies of Sciences, Engineering, and Medicine. *Improving Diagnosis in Health Care*. Washington, DC: The National Academies Press. 2015. doi:10.17226/21794.

2. Saber Tehrani AS, Lee H, Mathews SC, et al. 25-Year summary of US malpractice claims for diagnostic errors 1986–2010: an analysis from the National Practitioner Data Bank. *BMJ Qual Saf*. 2013;22:672–680.

3. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med*. 2005;165:1493–1499.

4. Zwaan L, de Bruijne M, Wagner C, et al. Patient record review of the incidence, consequences, and causes of diagnostic adverse events. *Arch Intern Med*. 2010;170:1015–1021.

5. Langelaan M, Broekens MA, de Bruijne MC, et al. *Monitor zorggerelateerde schade 2015/2016: dossieronderzoek bij overleden patiënten in Nederlandse ziekenhuizen*. Utrecht, The Netherlands: NIVEL; 2017.

6. Inspectie Gezondheidszorg en Jeugd. In openheid leren van meldingen 2016-2017. 2018. Available at: https://www.igj.nl/publicaties/rapporten/2018/01/25/in-openheid-leren-van-meldingen-2016-2017. Accessed December 10, 2020.

7. Ilgen JS, Humbert AJ, Kuhn G, et al. Assessing diagnostic reasoning: a consensus statement summarizing theory, practice, and future needs. *Acad Emerg Med*. 2012;19:1454–1461.

8. Zwaan L, Hautz WE. Bridging the gap between uncertainty, confidence and diagnostic accuracy: calibration is key. *BMJ Qual Saf*. 2019;28:352–355.

9. Croskerry P, Sinclair D. Emergency medicine: a practice prone to error? *CJEM*. 2001;3:271–276.

10. Kachalia A, Gandhi TK, Puopolo AL, et al. Missed and delayed diagnoses in the emergency department: a study of closed malpractice claims from 4 liability insurers. *Ann Emerg Med*. 2007;49:196–205.

11. Zwaan L, Monteiro S, Sherbino J, et al. Is bias in the eye of the beholder? A vignette study to assess recognition of cognitive biases in clinical case workups. *BMJ Qual Saf*. 2017;26:104–110.

12. Zwaan L, Thijs A, Wagner C, et al. Relating faults in diagnostic reasoning with diagnostic errors and patient harm. *Acad Med*. 2012;87:149–156.

13. Singh H, Khanna A, Spitzmueller C, et al. Recommendations for using the Revised Safer Dx Instrument to help measure and improve diagnostic safety. *Diagnosi*. 2019;6:315–323.

14. Schiff GD, Kim S, Abrams R, et al. Diagnosing diagnosis errors: lessons from a multi-institutional collaborative project. *Adv Patient Saf Res Implement*. 2005;2:255–278.

15. Schiff GD, Hasan O, Kim S, et al. Diagnostic error in medicine: analysis of 583 physician-reported errors. *Arch Intern Med*. 2009;169:1881–1887.

16. Reason J. *Human Error*. Cambridge, England: Cambridge University Press; 1990.

17. Reason J. Human error: models and management. *BMJ*. 2000;320:768–770.

18. Baartmans M, van Schoten S, Wagner C. *Ziekenhuisoverstijgende analyse van calamiteiten*. Utrecht, The Netherlands: NIVEL; 2020.

19. Leistikow I, Mulder S, Vesseur J, et al. Learning from incidents in healthcare: the journey, not the arrival, matters. *BMJ Qual Saf*. 2017;26:252–256.

20. Van der Schaaf TW, Habraken MMP. *PRISMA-Medical: A Brief Description*. Eindhoven, the Netherlands: Eindhoven University of Technology; 2005.

21. Leistikow IP, Ridder K, Vries B. *Patiëntveiligheid: systematische incident reconstructie en evaluatie*. Houten, The Netherlands: Elsevier gezondheidszorg; 2009.

22. Tan HS. Elk incident heeft een context Het analyseren van een incident heeft het meeste effect als dit de organisatie áchter het voorval blootlegt. Want menselijk falen heeft altijd een context. *Med Contact*. 2010;65:2290.

23. Inspectie Gezondheidszorg en Jeugd. Richtlijn calamiteitenrapportage. 2019. Available at: https://www.igj.nl/publicaties/richtlijnen/2016/01/01/richtlijn-calamiteitenrapportage. Accessed October 18, 2021.

24. Davalos MC, Samuels K, Meyer AND, et al. Finding diagnostic errors in children admitted to the PICU. *Pediatr Crit Care Med*. 2017;18:265–271.

25. Keers RN, Williams SD, Cooke J, et al. Causes of medication administration errors in hospitals: a systematic review of quantitative and qualitative evidence. *Drug Saf*. 2013;36:1045–1067.

26. Hollnagel E. *The ETTO Principle: Efficiency-Thoroughness Trade-off: Why Things That Go Right Sometimes Go Wrong*. Boca Raton, FL: CRC Press; 2009.

27. Pani JR, Chariker JH. The psychology of error in relation to medical practice. *J Surg Oncol*. 2004;88:130–142.

28. Raffel KE, Kantor MA, Barish P, et al. Prevalence and characterisation of diagnostic error among 7-day all-cause hospital medicine readmissions: a retrospective cohort study. *BMJ Qual Saf*. 2020;29:971–979.

29. Nichols P, Copeland T-S, Craib IA, et al. Learning from error: identifying contributory causes of medication errors in an Australian hospital. *Med J Aust*. 2008;188:276–279.

30. Hibbert PD, Thomas MJW, Deakin A, et al. Are root cause analyses recommendations effective and sustainable? An observational study. *Int J Qual Health Care*. 2018;30:124–131.

31. Kellogg KM, Hettinger Z, Shah M, et al. Our current approach to root cause analysis: is it contributing to our failure to improve patient safety? *BMJ Qual Saf*. 2017;26:381–387.

32. Peerally MF, Carr S, Waring J, et al. The problem with root cause analysis. *BMJ Qual Saf*. 2017;26:417–422.