Yu Watanabe, Taiju Miyagami*, Taro Shimizu, Yuji Nishizaki, Sho Ukishima, Koichiro Santo, Seiko Furusaka Kushiro, Nozomi Aoki, Mayu Suzuki, Akio Kanazawa and Toshio Naito

Diagnostic errors in patients admitted directly from new outpatient visits

https://doi.org/10.1515/dx-2024-0088
Received May 16, 2024; accepted December 4, 2024;
published online January 1, 2025

Abstract

Objectives: Diagnostic errors frequently represent significant adverse events that can occur in any medical setting, particularly in rushed handovers and constrained timing. Cases that result in emergency hospitalization at the time of the initial outpatient visit are more likely to have complex or serious patient conditions and more detrimental diagnostic errors. Our study investigated diagnostic errors in these under reported situations.

Methods: We conducted a retrospective study using electronic medical record data on patients who were directly admitted to a newly established outpatient clinic at a single university hospital in Japan. Diagnostic errors were assessed independently by two physicians using the Revised Safer Dx instrument. We analyzed patient demographics, symptoms, referrals, and resident doctor (postgraduate-year-1) involvement using logistic regression to compare groups with and without diagnostic error. Additionally, we employed the Diagnostic Error Evaluation and Research (DEER) taxonomy and Generic Diagnostic Pitfalls (GDP) to examine the factors associated with diagnostic errors.

Results: The study included 321 patients, with diagnostic errors identified in 39 cases (12.1 %). Factors contributing to diagnostic errors included the involvement of young residents, male patients, the number of symptoms, and atypical presentation. The most common causes of diagnostic errors were “too much weight given to competing/coexisting diagnosis” as indicated by DEER and “atypical presentation” by GDP.

Conclusions: The frequency of diagnostic errors in this study was higher than those in previous studies of new outpatient visits, underscoring the imperative for heightened scrutiny in cases involving medical residents especially when patients present with multiple or atypical symptoms. This vigilance is crucial to mitigating the risk of diagnostic inaccuracies in these settings. Cases that result in emergency hospitalization at the time of the initial outpatient visit are more likely to have complex or serious patient conditions and more detrimental diagnostic errors.

Keywords: diagnostic error; education; recognition

Introduction

Diagnostic errors are common phenomena that occur across a wide range of patient encounters [1] and represent the most frequent cause of medical litigation [2]. The frequency of diagnostic errors varies across practice settings. For instance, in outpatient settings, approximately 5 % of new patients experience diagnostic errors, and nearly half of these errors result in harm to the patient [3, 4]. In primary care settings, such as clinics, error rates are notably higher for serious conditions like cancer and heart diseases (such as myocardial infarction) [5]. In the field of general internal medicine, cancers, pulmonary embolisms, and infections are commonly misdiagnosed [6]. Additionally, it has been noted that a significant proportion of preventable adverse events in initial outpatient care are related to diagnostic issues [7]. Furthermore, unexpected hospitalizations or return visits within 14 days following an initial outpatient consultation are also linked to diagnostic errors [8].

It is important to understand, however, that although an outpatient provider may recognize the need for hospitalization, specific diagnostic errors themselves do not necessarily “lead to” this decision. More commonly, it is the recognition of the severity of a patient’s condition that prompts hospitalization. Differentiating between these two

***Corresponding author: Taiju Miyagami**, MD, PhD, Department of General Medicine, Juntendo University Faculty of Medicine, 2-1-1 Hongo, Bunkyo-ku, Tokyo, Japan, E-mail: tmiyaga@juntendo.ac.jp. https://orcid.org/0000-0002-4893-2224
Yu Watanabe, Sho Ukishima, Koichiro Santo, Seiko Furusaka Kushiro, Nozomi Aoki, Mayu Suzuki, Akio Kanazawa and Toshio Naito, Department of General Medicine, Juntendo University Faculty of Medicine, Bunkyo-Ku, Tokyo, Japan
Taro Shimizu, Department of Diagnostic and Generalist Medicine, Dokkyo Medical University Hospital, Tochigi, Japan
Yuji Nishizaki, Department of General Medicine, Juntendo University Faculty of Medicine, Bunkyo-Ku, Tokyo, Japan; and Division of Medical Education, Juntendo University School of Medicine, Bunkyo-Ku, Tokyo, Japan

factors provides a clearer understanding of the relationship between diagnostic errors and hospital admissions. Consequently, diagnostic errors in outpatient care settings have been studied from various perspectives [6–8].

That said, studies focusing on patients admitted directly following their first outpatient visit are lacking. In general, diagnostic errors occurring during initial outpatient visits that do not result in immediate hospitalization are considered less severe than those that do require urgent hospitalization [9]. This is particularly concerning because these situations involve patients whose conditions were not fully anticipated during the initial consultation, requiring emergency hospitalization. The emergency nature of these hospitalizations suggests that the patient's condition is complex or severe, making outpatient management challenging [10]. In cases where the patient's vital signs deteriorate, the likelihood of diagnostic errors increases, especially during handoffs between caregivers [11]. Therefore, it can be stated that patients admitted directly from their first outpatient visit are in a situation that heightens the risk of diagnostic errors.

As such, analyzing diagnostic errors in patients requiring immediate hospitalization after their first consultation is essential for patient safety and aligns with the broader objective of improving diagnostic accuracy [1]. Understanding the errors associated with such hospitalizations requires an in-depth examination of both the mechanisms of these errors and the underlying factors that contribute to them. Therefore, this study will explore diagnostic errors in detail, including error mechanisms and the conditions that precipitate them. The insights gained are expected to help inform strategies for reducing diagnostic errors in patients transitioning from outpatient visits to hospitalization, ultimately enhancing diagnostic quality, supporting patient safety, and improving overall healthcare outcomes.

Patients and methods

Setting and patients

Juntendo University Hospital, a 1,051-bed university-affiliated hospital in Tokyo, is a tertiary referral hospital in the region. The Department of General Medicine and the Department of Emergency Medicine are at the forefront of emergency primary care, providing new outpatient consultations and emergency responses 24 h a day. The Department of General Medicine has overseen new outpatient consultations since 2010, with an average of approximately 3,000 new outpatients per year (an average of 12.1 patients per day). Data of patients who visited the tertiary referral hospital between May 1, 2014, and March 31, 2022, and subsequently admitted to the hospital were retrospectively extracted from electronic medical records. The Primary Care Center is controlled by The Department of General Medicine and the Department of Medicine and accepts initial outpatient visits to the hospital. The patient background is mainly walk-in patients with referrals, rather than cases transported by ambulance. All patients in this study were walk-in and off-appointment patients. Patients who visited outside regular hours or emergency outpatient services were excluded from the study. In addition, cases where the final diagnosis was not recorded or was unknown were excluded.

Ethics

Participants' clinical data were retrospectively retrieved from the institution's database. All tests included in this study were performed as part of a routine program, and no data were collected specifically for the purpose of this research. Participants' records and information were anonymized and de-identified prior to analysis. As this was a retrospective study, patient consent was obtained via opt-out on the Juntendo website. The protocol for this study was approved by the Institutional Review Board of Juntendo University Hospital (Approval number E22-0175). This study was conducted in accordance with the Declaration of Helsinki. All the study methods were followed the Ethical Guidelines for Medical and Health Research Involving Human Subjects.

Variables

The measured variables included patients' age, sex, chief complaint, number of symptoms, presence of referral letter from another hospital or clinic, resident doctors' involvement, AM or PM visit, initial diagnosis, final diagnosis, disease category, diagnostic error, and length of stay. The number of chief complaints and symptoms were obtained from the medical questionnaire answered by the patient, and the number of symptoms were counted according to the International Classification of Primary Care Second Edition (ICPC-2) Japanese version [12]. The reason for surveying the difference between AM and PM is that more errors have been reported during shift changes in the emergency room, and more diagnostic errors occur during weekends [13, 14]. We defined 8:00-12:59 as a.m. and 13:00-17:00 as p.m., as regards the timing of the start of the examination. To minimize bias during the case review, we used common definitions of diagnostic error: "delay in diagnosis" and "missed or wrong diagnoses" [15].

After graduating from medical school, resident doctors in Japan are required to spend two years in residency, during which they are required to rotate through internal and emergency medicine, surgery, pediatrics, obstetrics and gynecology, psychiatry, and community medicine, and spend at least one month in outpatient care [16]. The resident doctor's basic duties are ward rounds. The residents at this hospital have a month in which they only do outpatient work in their postgraduate year (PGY) 1 during which they train at the new outpatient clinic of the hospital's general medicine department. Resident doctors' involvement was confirmed if there was a medical record entry by a resident doctor rotating through the outpatient clinic. The initial and final diagnoses were limited to the primary disease associated with the chief complaint leading to admission.

The International Classification of Diseases, 11th edition (ICD-11) was used for the disease category [17]. ICD-11 categorizes diseases into 26 broad categories, including infectious diseases, neoplasms, and cardiovascular diseases, and this study used these categories accordingly. The initial diagnosis for the disease according to ICD-11 was assigned by the physician at the time of patient admission, and the final diagnosis was assigned at the time of patient discharge. Based on the final diagnosis, the medical record was reviewed to assess whether a case was a common or uncommon condition and typical or atypical disease. The disease was considered uncommon if the final diagnosis was registered in Orphanet [18]. Otherwise, if prevalence or incidence was listed in UpToDate™ or DynaMed™, the disease was classified as common or uncommon/rare, with a cutoff of 1 per 2,000 patients. For those not listed, diseases routinely encountered in outpatient practice were classified as common and all others as uncommon. Orphanet is an online service where rare diseases are registered. UpToDate™/DynaMed™ are both reliable medical information databases that hold epidemiological information and are used by healthcare professionals to quickly obtain evidence-based information in clinical practice.

Two physicians (YW and TM, general medical doctors in their 8th and 14th year of practice) independently assessed diagnostic errors using the Revised Safer Dx instrument, and the concordance rate was 98.1 % [19, 20]. Cases where the two physicians disagreed were assessed by a third physician (TS), to determine whether an error had occurred. Revised safer Dx, devised by Hardeep Singh in 2019, examines whether a loss of diagnostic opportunity for 13 items occurred and assesses whether a diagnostic error occurred as a result. The two physicians carefully evaluated the cases based on the medical records, paying attention to the differences between common and uncommon, typical and atypical diagnoses. Diagnostic Error Evaluation and Research (DEER) taxonomy and Generic Diagnostic Pitfalls (GDP) were used to analyze error cases [21, 22]. Both DEER and GDP are methods created by Schiff GD to evaluate the causes of diagnostic errors. DEER is evaluated in the chronological sequence from the patient's visit to the follow-up after the consultation; GDP is evaluated in 12 situations that are likely to result in pitfalls. The two together provide a better understanding of the characteristics of diagnostic errors. Two physicians also evaluated these independently at a rate of 93.6 %. The two physicians discussed the results and made a final decision.

Data analysis

All patient background data and diagnostic error analysis contents were manually entered by the authors after checking the electronic medical records. The results are presented as means ± standard deviation for continuous variables, and as prevalence (%) for categorical variables. A logistic regression analysis was performed with the non-DE group as the reference for intergroup comparisons between the groups with diagnostic error (DE) and the group without diagnostic error (non-DE). The covariates were age, sex, number of symptoms, referral, resident's involvement, afternoon visit, length of stay, mortality within the period of hospitalization, and uncommon or atypical; all variables were entered simultaneously into the logistic regression model. The Kenward–Roger method was used to estimate the degrees of freedom; 95 % confidence intervals (CIs) were calculated using Cox proportional hazards model [23]. All statistical analyses were performed using SAS version 9.4 (SAS Institute, Cary, NC, USA). Statistical significance was set at $p<0.05$.

Results

A total of 898 patients were examined during our new outpatient/emergency department visits. Subsequently, 279 patients who visited the emergency department, 99 who visited the specialized outpatient clinic for COVID-19, and 193 who visited the outpatient clinic after hours were excluded. A total of 327 patients were included in the study; six cases in which the final diagnosis remained unknown were excluded, and 321 patients were included in the study (Figure 1). The number of errors reported during this period was 39 (12.1 %).

The mean age of the patients was 59 ± 22 years. There were 149 (46.4 %) female patients. The mean number of symptoms reported by the patients was 3 ± 1. 67 patients (20.9 %) were referred to the hospital, and 124 (38.8 %) were treated by residents; 153 (47.7 %) patients were examined in the afternoon at an outpatient clinic. The length of hospital

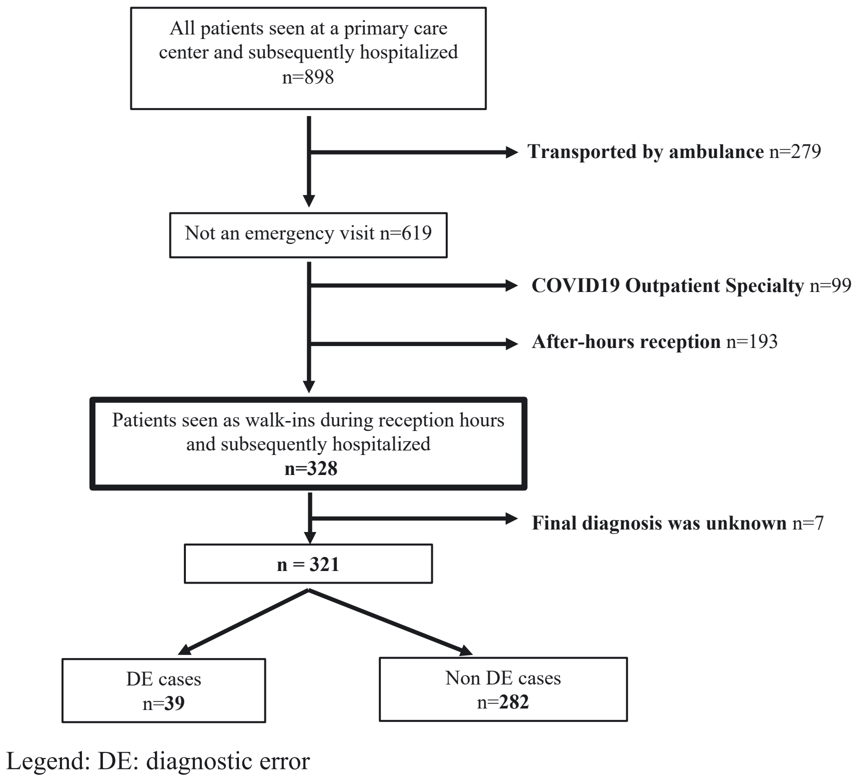

Figure 1: Participant flow. DE, diagnostic error.

stay was 20 ± 33 days; 73 (22.7 %) patients had an uncommon course, and 108 (33.6 %) had an atypical course. The number of errors reported during this period was 39 (12.1 %). In the group comparison, the error group predominantly had fewer reported errors among female patients, more errors in cases involving residents, afternoon outpatients, uncommon/atypical cases, and longer hospital stays (Table 1).

The total number of chief complaints in the study population was 455. Altogether, 130 (28.6 %) of the study participants had a chief complaint of fever. Fever was the most common complaint among the 39 patients in the DE group (23 patients, 59.0 %). All other chief complaints were named by fewer than four patients. The main complaints of the overall group and the diagnostic error cases are shown in Supplementary Material, Appendix 1 (Table).

Logistic regression analysis in the DE and non-DE groups showed that female patients were predominantly less prone to diagnostic errors (adjusted odd ratio (aOR 0.22), 95 % confidence interval (CI) (0.09–0.52), p<0.001). Meanwhile diagnostic errors tended to be higher with greater number of symptoms (aOR 1. 33, 95 %CI(1.01–1.75), p=0.043), more resident doctors (PGY-1) involved (aOR 2.34, 95 % CI(1.08–5.06), p=0.032), more atypical cases (aOR 4.12, 95 % CI(1.81–9.40), p<0.001) (Table 2).

The category of disease with the most diagnostic errors was infectious diseases (20 cases, 51.3 %). The most common of these were pneumonia (4 cases), pharyngitis (2 cases), and deep abscess (2 cases). The next most common category was blood disorders (5 cases, 12.8 %), with coagulation disorders being the most common (3 cases). The next most common category was autoimmune disorders (4 cases, 10.3 %), with autoinflammatory syndrome being the most common (3 cases).

Table 1: Total cases of DE and non-DE.

	Total	DE (39)	Non-DE (282)	p-Value
Age	59 ± 22	58 ± 18	59 ± 23	0.616
Sex: women	149 (46.4)	8 (5.4)	141 (50.0)	<0.001
Number of symptoms	3 ± 1	3 ± 1	3 ± 1	0.060
Referral letter available	67 (20.9)	11 (28.2)	56 (19.9)	0.244
Involvement of residents in the practice	124 (38.8)	22 (56.4)	102 (36.3)	0.017
Visited in the afternoon	153 (47.7)	22 (56.4)	131 (46.5)	0.243
Duration of hospitalization	20 ± 33	32 ± 55	18 ± 29	0.021
Number of deaths	13 (4.1)	4 (10.3)	9 (3.2)	0.068
Uncommon	73 (22.7)	18 (46.2)	55 (19.5)	<0.001
Atypical	108 (33.6)	25 (64.1)	83 (29.4)	<0.001

Continuous variables are presented as mean ± SD and nominal variables as number (%).

Table 2: Comparison between DE group and non-DE group.

	OR	Lower CI	Upper CI	p-Value	aOR	Lower CI	Upper CI	p-Value
Age	1.00	0.99	1.01	0.658	0.99	0.98	1.01	0.480
Sex women	0.26	0.11	0.58	<0.001	0.22	0.09	0.52	<0.001
Number of symptoms	1.24	0.99	1.56	0.065	1.33	1.01	1.75	0.043
Referral letter available	1.59	0.74	3.38	0.244	1.12	0.47	2.63	0.803
Involvement of residents in the practice	2.27	1.15	4.47	0.017	2.34	1.08	5.06	0.032
Visited in the afternoon	1.49	0.76	2.93	0.243	1.88	0.84	4.19	0.125
Duration of hospitalization	14.58	1.07	198.87	0.042	1.01	1.00	1.02	0.993
Number of deaths	1.08	0.96	1.20	0.068	1.21	0.28	5.28	0.802
Uncommon	3.54	1.77	7.09	<0.001	1.93	0.88	4.26	0.102
Atypical	4.28	2.12	8.64	<0.001	4.12	1.81	9.40	<0.001

Non-DE group as a reference. CI, confidence interval; OR, odd's ratio; aOR, adjusted odd's ratio.

The categories of initial and final diagnoses were evaluated. When only diagnostic error cases were evaluated, 21.7 % cases showed a different category between the initial and final diagnoses. The disease category with the most diagnostic errors was infectious diseases.

In DEER, "too much weight on competing/coexisting diagnosis" (category 5C) was the most common cause of diagnostic error, whereas in GDP, atypical presentation (category 5) was the most common cause of diagnostic error. When fever was the chief complaint, "Erroneous lab/radiology reading of test" (category 4H) was most common in DEER, and "counter-diagnosis cues overlooked (e.g., red flags)" (category 6) was most common in GDP. As for the other chief complaints, "Too much weight on competing/coexisting diagnosis" (category 5C) was most common in DEER, and "Atypical presentation" (category 5) was most common in GDP (Table 3).

Discussion

Diagnostic errors accounted for 12.1 % of all cases in this study. Overall, female patients experienced a lower frequency of diagnostic errors. Higher number of symptoms, residents' involvement, and the more atypical diagnoses were all associated with more diagnostic errors. In addition, 21.7 % of cases with diagnostic errors showed a change in disease group. The most common cause for error was the stage of examination and assessment in DEER, whereas atypical presentation or missed findings that could be a clue to the diagnosis were more common in GDP.

Frequency of diagnostic errors

The incidence of diagnostic errors in new outpatient visits is estimated to be approximately 5 %. However, in this study, it was 12.1 %, indicating a higher than usual risk of error [3]. Compared with previous studies that examined the incidence of errors on a case-by-case basis, the incidence of diagnostic errors in patients admitted to the hospital and transferred to the ICU was 23 % [24]. This study found a difference in the length of stay between DE and non-DE and mortality rates tended to be higher, although not significantly different due to small patient numbers.

Table 3: Diagnostic error cause analysis.

Method	Category	Description	Diagnostic error total, n (%)	Fever as chief complaint, n (%)	Other chief complaint, n (%)
DEER	2A	Elicit history	2 (5.1)	0 (0)	2 (12.5)
DEER	3A	Elicit exam	2 (5.1)	2 (8.7)	0 (0)
DEER	4A	Order tests	2 (5.1)	0 (0)	2 (12.5)
DEER	4H	Erroneous read	6 (15.4)	4 (17.4)	2 (12.5)
DEER	4K	Error in interpretation	3 (7.7)	3 (13.0)	0 (0)
DEER	5C	Competing dx	8 (20.4)	4 (17.4)	4 (25.0)
GDP	2	Limitations of tests	6 (15.4)	4 (17.4)	2 (12.5)
GDP	3	A mistaken for B	5 (12.8)	4 (17.4)	1 (6.3)
GDP	5	Atypical presentation	11 (28.2)	5 (21.7)	6 (37.5)
GDP	6	Counter dx cue overlooked	10 (25.6)	7 (30.4)	3 (18.8)

The causes of diagnostic errors were analyzed by chief complaint. DEER, diagnostic error evaluation and research taxonomy. 2A: failure/delay in eliciting critical piece of history data. 3A: failure/delay in eliciting critical physical exam findings. 4A: failure/delay in ordering needed test(s) 4H: erroneous lab/radiology reading of test. 4K: error in clinician interpretation of test. 5C: too much weight on competing/coexisting diagnosis. GDP, generic diagnostic pitfalls. 2: limitations of a test or exam finding not appreciated. 3: disease A repeatedly mistaken for disease B. 5: atypical presentation. 6: counter-diagnosis cues overlooked.

Gender differences in susceptibility to diagnostic errors

Women were significantly less prone to diagnostic errors. A report on the frequency of error susceptibility according to patient sex reported that males had more undocumented malignancies in an analysis of autopsy cases, with no other significant difference [25]. Although we were not able to conduct a rigorous survey, it is possible that the sex of doctors may also affect their practice. It has been reported that when female physicians cared for female patients, they had lower 30-day mortality and readmission rates, with no significant differences in other gender differences [26]. A previous study reported that the mortality and readmission rates were lower for patients treated by female physicians than among those treated by male physicians [27]. It is possible that the physician's gender may have some influence in the diagnostic field as well. In addition, we compared the results of the measurements taken on male and female patients (age, length of hospitalization, common/uncommon disease, typical/atypical disease, etc.), but there were no significant differences in any of the measurements (Supplementary Material, Appendix 2). We also evaluated the initial and final diagnoses for male and female patients. Bacterial pneumonia was the most common diagnosis for both men and women (25 [14.5 %] for men and 19 [12.8 %] for women), followed by pyelonephritis (7 [4.1 %] for men and 17 [11.4 %] for women) and aseptic meningitis (7 [4.1 %] for men and 10 [6.9 %] for women). When comparing men and women in this study, the final diagnoses were all ranked the same, and there were no differences in diseases. An important limitation of this study is that we were unable to link the data with the gender of the doctors; multiple doctors often write the medical records, and it is not always clear who the main doctor is. For this reason, the gender of the doctors involved in this study remains unknown. In the future, we would like to further evaluate patient-physician factors in diagnosis.

Diagnostic errors associated with the number of symptoms

Based on our finding, we assumed that a greater number of symptoms increases diagnostic complexity. We cannot rule out the possibility that many symptoms lead to atypical presentations. In addition, the complexity of the cases may increase when there are many symptoms because researchers have to consider differential diagnosis for each symptom, and their combination. As the patients were seen in a university hospital, the high complexity levels of the patient base may have led to errors due to decision fatigue among the physicians [28]. In situations prone to decision fatigue, it is advisable to take intermittent breaks to rejuvenate the mind and collaboratively assess the issue using the collective intelligence of multiple physicians [29].

Resident doctor involvement

Residents are more prone to diagnostic errors, especially during history taking, physical examinations, and assessments [30]. In this hospital, when residents are involved in the new outpatient care, they are often seen as the first point of contact. It is possible that the diagnostic accuracy may be influenced by the accuracy of the resident's initial diagnosis. Previous studies have reported that the likelihood of misdiagnosis increases if another diagnosis is given upfront [31]. In addition, time course is more important in the outpatient setting than in the inpatient setting, this is because, even if only provisionally, a caregiver has to make a diagnosis and decision on whether to admit the patient to the hospital as well as to provide some initial treatment within the limited time available during the outpatient consultation. In the United States, the participation of medical students in outpatient consultations has been reported to delay the completion of the consultation [32]. On the other hand, it has been reported that the use of UpToDate® can reduce diagnostic errors to improve the care of early clinical residents [33]. Factors such as the accuracy of diagnosis at the time of resident consultation, and tools used by residents for diagnosis are also important for future analyses. On the other hand, senior physicians also assist novices to mitigate the effects of cognitive biases, by providing more extensive educational supports based on the resident doctor's level of confidence and their diagnostic skills.

Atypical presentation

Previous studies have shown that atypical cases are more prone to diagnostic errors [34]. The fact that 21.7 % of cases with diagnostic errors were categorized differently at the time of initial and final diagnoses, and that the number of symptoms was large, suggests that atypical cases are likely to have influenced them. Atypical presentation had the greatest impact on diagnostic errors in the GDP analysis in this study, suggesting that they were strongly associated. General practitioners tend to recall typical diagnoses when making diagnostic errors [35]. Even in incident report evaluations, atypical cases are said to be more likely to be related to errors [36]. It is important to know that atypical cases are more prone to errors than usual, especially when patients are admitted directly from a new outpatient visit [37]. One way to reduce such errors is to have a wide range of differential diagnoses and treat cases as if they fit or do not fit the diagnostic name, leaving the possible cases as differential diagnoses. It may also be necessary to use diagnostic aids and artificial intelligence as a diagnostic aid because these assistances possibly support physicians by deploying comprehensive differential diagnoses.

Analysis by DEER/GDP

In previous reports using DEER, the examination phase had the greatest impact, followed by differential diagnosis [21]. In the present study, however, differential diagnosis (Too much weight given to competing/coexisting diagnosis) had a greater impact, albeit by a small margin, followed by testing (Erroneous lab/radiology reading of test).

In addition, this study included more atypical cases; in a similar previous study analyzing the causes of case reports of diagnostic errors using DEER, Failure/delay in considering the diagnosis was the most common cause in atypical cases [38].

The reasons for the difference from previous reports include the possibility that the diagnosis was given in advance due to the involvement of residents, and that the relatively short time course from outpatient to inpatient resulted in the physician was unable to consider the differential diagnosis, and the treatment flow moved with a tentative diagnosis for the time being.

The trend for chief complaints other than fever was the same as the overall trend, but when only the chief complaint of fever was considered, the survey and assessment reversed, albeit by a small margin. This is because patients presenting with fever often need to be managed with infection control in mind, as in a previous report on delayed antimicrobial therapy during COVID-19 outbreaks, making testing more likely to influence the diagnosis [39]. Atypical presentations and missed diagnostic clues were most commonly associated with GDP. Missed laboratory findings inconsistent with the diagnosis have been reported in the past [21, 40]. We found that atypical presentation was still difficult to diagnose and influenced diagnostic errors. The results were similar in each group except for the overall group and the group with fever as the chief complaint, but only in the group with fever as the chief complaint, there were many 6: counter-diagnostic clues overlooked. This may be because during an epidemic, patients may believe they have the disease even if the tests are negative.

Limitations

The first limitation was the institutional background. This was a single-center study, with selection bias in the patient population due to it being a university hospital that has high thresholds for receiving care. In addition, the Department of General Medicine often sub-specializes in infectious diseases, which results in a bias toward infectious diseases in the outpatient population and other disease categories. Since the number of cases is limited to internal medicine cases, fewer orthopedic and minor diseases were seen. Additionally, the complexity of the consultation process and the division between departments of university hospitals may also have had an impact on the results. It is possible that physicians affiliated with the same department evaluated the diagnostic errors, which may have resulted in some bias. In future work, physicians from multiple departments should evaluate the data. Furthermore, this was a cross-sectional study, and there was no clarity regarding causal relationships. Additionally, diagnostic errors are greatly influenced by the physician's circumstances, environment, and conditions; however, many details of the circumstances are not as clear as in the past. This study is a retrospective study, and patients were not examined at the time of their initial visit with an eye toward diagnostic error. The initial diagnosis was assigned to the disease at the time of admission, and the final diagnosis was assigned to the disease at the time of discharge. In addition, we did not have access to information other than the medical records such as the patient's condition before their outpatient visit or the diagnosis made during that visit. The number of cases of diagnostic errors in this study is small and does not lead to generalization.

Despite these limitations, this is the first study to analyze diagnostic errors in situations with a high risk of hospitalization from new outpatient visits, and is therefore of high value. The analysis of diagnostic errors and the proposed measures to mitigate them prompt physicians to reduce future errors. In the future, we would like to conduct studies that prospectively evaluate situational and environmental factors, including facilities of different sizes, regional characteristics, and urban hospitals, and clinics.

Conclusions

The frequency of diagnostic errors in patients admitted directly from new outpatient visits was 12.1 %, which was higher than in previous studies of new outpatient visits. It is necessary to pay attention to situations in which residents are involved in medical care, especially when the number of symptoms is large or atypical, or when the patient is male. Recognizing and being vigilant about this situation and context is extremely important, especially in a high-risk setting where the patient is admitted directly through new outpatient visits.

Research ethics: The protocol for this study was approved by the Institutional Review Board of Juntendo University Hospital (Approval number E22-0175). This study was conducted in accordance with the Declaration of Helsinki. All the study methods were followed the Ethical Guidelines for Medical and Health Research Involving Human Subjects.
Informed consent: As this was a retrospective study, patient consent was obtained via opt-out on the Juntendo website.
Author contributions: All authors were involved in study design, data interpretation, and manuscript preparation. YW and TM are the principal investigators and were responsible for regulatory compliance, participant recruitment, data collection, data analysis, and manuscript preparation. WW, TM, TS, YN, SU, KS, SK, NA, MS, AK and TN contributed to the study coordination and data collection, entry, and analysis. All authors read and approved the final manuscript.
Use of Large Language Models, AI and Machine Learning Tools: Not applicable.
Conflict of interest: The authors state no conflict of interest.
Research funding: TM received a grant from the National Academic Research Grant Funds (JSPS KAKENHI: 22K17310).
Data availability: The raw data can be obtained on request from the corresponding author.

References

1. Committee on diagnostic error in health C, board on health care S, Institute of M, the national academies of sciences E, medicine. In: Balogh EP, Miller BT, Ball JR, editors. Improving diagnosis in health care. Washington (DC): National Academies Press (US); 2015.
2. Saber Tehrani AS, Lee H, Mathews SC, Shore A, Makary MA, Pronovost PJ, 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;22:672–80.
3. Singh H, Meyer, Thomas EJ. The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations. BMJ Qual Saf 2014;23:727–31.
4. Singh H, Schiff GD, Graber ML, Onakpoya I, Thompson MJ. The global burden of diagnostic errors in primary care. BMJ Qual Saf 2017;26: 484–94.
5. Fernholm R, Pukk Härenstam K, Wachtler C, Nilsson GH, Holzmann MJ, Carlsson AC. Diagnostic errors reported in primary healthcare and emergency departments: a retrospective and descriptive cohort study of 4830 reported cases of preventable harm in Sweden. Eur J Gen Pract 2019;25:128–35.
6. Matulis JC, Kok SN, Dankbar EC, Majka AJ. A survey of outpatient Internal Medicine clinician perceptions of diagnostic error. Diagnosis 2020;7:107–14.
7. Avery AJ, Sheehan C, Bell B, Armstrong S, Ashcroft DM, Boyd MJ, et al. Incidence, nature and causes of avoidable significant harm in primary care in England: retrospective case note review. BMJ Qual Saf 2021;30: 961–76.
8. Singh H, Giardina TD, Forjuoh SN, Reis MD, Kosmach S, Khan MM, et al. Electronic health record-based surveillance of diagnostic errors in primary care. BMJ Qual Saf 2012;21:93–100.
9. Alahmary K, Kadasah S, Alsulami A, Alshehri AM, Alsalamah M, Da'ar OB. To admit or not to admit to the emergency department: the disposition question at a tertiary teaching and referral hospital. Healthcare 2023;11:667.
10. Low LL, Wah W, Ng MJ, Tan SY, Liu N, Lee KH. Housing as a social determinant of health in Singapore and its association with readmission risk and increased utilization of hospital services. Front Public Health 2016;4:109.
11. Gleason KT, Jones R, Rhodes C, Greenberg P, Harkless G, Goeschel C, et al. Evidence that nurses need to participate in diagnosis: lessons from malpractice claims. J Patient Saf 2021;17:e959–63.
12. World Health Organization. International classification of primary care, 2nd ed Medical Database Committee; 2020. http://icpc.nosmokeworld.com/.
13. Tyler PD, Fossa A, Joseph JW, Sanchez LD. Later emergency provider shift hour is associated with increased risk of admission: a retrospective cohort study. BMJ Qual Saf 2020;29:465–71.
14. Watari T, Gupta A, Amano Y, Tokuda Y. Japanese internists' most memorable diagnostic error cases: a self-reflection survey. Intern Med 2024;63:221–9.
15. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. Arch Intern Med 2005;165:1493–9.
16. Ministry of Health LaW. Guideline for clinical training of medical practitioners -fiscal year 2020 edition;2020. https://www.mhlw.go.jp/content/10800000/ishirinsyokensyu_guideline_2020.pdf.
17. Japan Primary Care Association. International classification of diseases 11th revision; 2024. https://icd.who.int/en.
18. Auerbach AD, Lee TM, Hubbard CC, Ranji SR, Raffel K, Valdes G, et al. Diagnostic errors in hospitalized adults who died or were transferred to intensive care. JAMA Intern Med 2024;184:164–73.
19. Al-Mutairi A, Meyer AN, Thomas EJ, Etchegaray JM, Roy KM, Davalos MC, et al. Accuracy of the safer Dx instrument to identify diagnostic errors in primary care. J Gen Intern Med 2016;31:602–8.
20. Singh H, Khanna A, Spitzmueller C, Meyer A. Recommendations for using the Revised Safer Dx Instrument to help measure and improve diagnostic safety. Diagnosis 2019;6:315–23.
21. Schiff GD, Hasan O, Kim S, Abrams R, Cosby K, Elstein AS, et al. Diagnostic error in medicine: analysis of 583 physician-reported errors. Arch Intern Med 2009;169:1881–7.
22. Schiff GD, Volodarskaya M, Ruan E, Lim A, Wright A, Singh H, et al. Characteristics of disease-specific and generic diagnostic pitfalls: a qualitative study. JAMA Netw Open 2022;5:e2144531.
23. Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. Biometrics 1997;53:983–97.
24. Orhanet. Orphanet: an online database of rare diseases and orphan drugs; 2024. https://www.orpha.net/consor/cgi-bin/index.php.
25. Goldman BI, Bharadwaj R, Fuller M, Love T, Metlay L, Dignan C. Error codes at autopsy to study potential biases in diagnostic error. Diagnosis 2023;10:375–82.
26. Miyawaki A, Jena AB, Rotenstein LS, Tsugawa Y. Comparison of hospital mortality and readmission rates by physician and patient sex. Ann Intern Med 2024;23. https://doi.org/10.7326/M23-3163.
27. Tsugawa Y, Jena AB, Figueroa JF, Orav EJ, Blumenthal DM, Jha AK. Comparison of hospital mortality and readmission rates for medicare

patients treated by male vs female physicians. JAMA Intern Med 2017;177:206–13.
28. Vohs KD, Baumeister RF, Schmeichel BJ, Twenge JM, Nelson NM, Tice DM. Making choices impairs subsequent self-control: a limited-resource account of decision making, self-regulation, and active initiative. J Pers Soc Psychol 2008;94:883–98.
29. Barnett ML, Boddupalli D, Nundy S, Bates DW. Comparative accuracy of diagnosis by collective intelligence of multiple physicians vs individual physicians. JAMA Netw Open 2019;2:e190096.
30. Harada T, Miyagami T, Watari T, Hiyoshi T, Kunitomo K, Tsuji T, et al. Analysis of diagnostic error cases among Japanese residents using diagnosis error evaluation and research taxonomy. J Gen Fam Med 2021;22:96–9.
31. 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. J Clin Med 2021;10:826.
32. Adams M, Eisenberg JM. What is the cost of ambulatory education? J Gen Intern Med 1997;12:S104–10.
33. Shimizu T, Nemoto T, Tokuda Y. Effectiveness of a clinical knowledge support system for reducing diagnostic errors in outpatient care in Japan: a retrospective study. Int J Med Inf 2018;109:1–4.
34. Harada Y, Otaka Y, Katsukura S, Shimizu T. Prevalence of atypical presentations among outpatients and associations with diagnostic error. Diagnosis 2024;11:40–8.
35. Ely JW, Kaldjian LC, D'Alessandro DM. Diagnostic errors in primary care: lessons learned. J Am Board Fam Med 2012;25:87–97.
36. Schiff GD, Volodarskaya M, Ruan E, Lim A, Wright A, Singh H, et al. Characteristics of disease-specific and generic diagnostic pitfalls: a qualitative study. JAMA Netw Open 2022;5:e2144531.
37. Ely JW, Kaldjian LC, D'Alessandro DM. Diagnostic errors in primary care: lessons learned. J Am Board Fam Med 2012;25:87–97.
38. Harada Y, Watari T, Nagano H, Suzuki T, Kunitomo K, Miyagami T, et al. Diagnostic errors in uncommon conditions: a systematic review of case reports of diagnostic errors. Diagnosis 2023;10:329–36.
39. Miyagami T, Uehara Y, Harada T, Watari T, Shimizu T, Nakamura A, et al. Delayed treatment of bacteremia during the COVID-19 pandemic. Diagnosis 2021;8:327–32.
40. Schiff GD. Finding and fixing diagnosis errors: can triggers help? BMJ Qual Saf 2012;21:89–92.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/dx-2024-0088).