Jacky Hooftman*, Andrew P.J. Olson, Casey N. McQuade, Sílvia Mamede, Cordula Wagner and Laura Zwaan

Time pressure in diagnosing written clinical cases: an experimental study on time constraints and perceived time pressure

https://doi.org/10.1515/dx-2024-0125
Received July 19, 2024; accepted November 1, 2024;
published online November 28, 2024

Abstract

Objectives: Time pressure and time constraints have been shown to affect diagnostic accuracy, but how they interact is not clear. The current study aims to investigate the effects of both perceived time pressure (sufficient vs. insufficient time) and actual time constraints (lenient vs. restricted time limit) with regard to diagnostic accuracy.

Methods: Residents from two university-affiliated training programs in the USA participated in this online within-subjects experiment. They diagnosed cases under two perceived time pressure conditions: one where they were told they had sufficient time to diagnose the cases and one where they were told they had insufficient time. The actual time limit was either restricted or lenient (± one standard deviation from the mean time to diagnose). Participants provided their most likely diagnosis and a differential diagnosis for each case, and rated their confidence in their most likely diagnosis.

Results: A restricted time limit was associated with lower accuracy scores (p=0.044) but no effects of perceived time pressure on diagnostic accuracy were found. However, participants self-reported feeling more time pressure when they thought they had insufficient time (p<0.001). In addition, there was an effect of the actual time limit (p=0.012) and perceived time pressure (p=0.048) on confidence.

Conclusions: This study showed that a restricted time limit can negatively affect diagnostic accuracy. Although participants felt more time pressure and were less confident when they thought they had insufficient time, perceived time pressure did not affect diagnostic accuracy. More research is needed to further investigate the effects of time pressure and time limits on diagnostic accuracy.

Keywords: time pressure; diagnostic accuracy; diagnostic reasoning; experiment; internal medicine

Introduction

In most hospitals and general practices, physicians are allowed a fixed amount of time to see each patient. Several studies into working conditions among health care professionals have reported that physicians often complain about time pressure [1] or express the need for more time to provide patients with good quality of care [2]. Time pressure relates to the amount of time allotted for a task compared to the amount of time one considers necessary for the task. Furthermore, working under time pressure is associated with job dissatisfaction, stress, fatigue and burnout [1–5].

Unsurprisingly, time pressure can have an impact on a physician's clinical work as well. A clinical study by Welker and colleagues [6] showed that reducing the 'time to first antibiotic dose' for patients with community-acquired pneumonia from 8 h to 4 h, had an effect on diagnostic accuracy. This measure is used as a core quality indicator and is publicly reported as a percentage of patients who are diagnosed and treated within that timeframe. A lower score could impact the hospital's quality ratings, and thus could

***Corresponding author: Ms Jacky Hooftman**, MSc, Department of Public and Occupational Health, Amsterdam UMC, Vrije Universiteit Amsterdam, Van der Boechorststraat 7, 1081 BT, Amsterdam, The Netherlands; Quality of Care, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands; and Erasmus Medical Center, Institute of Medical Education Research Rotterdam, Rotterdam, The Netherlands,
E-mail: j.hooftman@amsterdamumc.nl. https://orcid.org/0000-0002-8375-6609
Andrew P.J. Olson, Departments of Medicine and Pediatrics, University of Minnesota, Minneapolis, MN, USA. https://orcid.org/0000-0002-7226-5783
Casey N. McQuade, Division of General Internal Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. https://orcid.org/0000-0002-6139-5531
Sílvia Mamede and Laura Zwaan, Erasmus Medical Center, Institute of Medical Education Research Rotterdam, Rotterdam, The Netherlands. https://orcid.org/0000-0003-1187-2392 (S. Mamede). https://orcid.org/0000-0003-3940-1699 (L. Zwaan)
Cordula Wagner, Department of Public and Occupational Health, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; Quality of Care, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands; and Netherlands Institute for Health Services Research (Nivel), Utrecht, The Netherlands. https://orcid.org/0000-0002-2301-2640

potentially create a sense of time pressure to diagnose a potential patient and give them the antibiotic treatment, even in the face of diagnostic uncertainty. The study showed that diagnostic accuracy decreased when the timeframe was shortened, despite the fact that the mean time to first antibiotic dose did not change significantly and was far below the 4 h time limit.

Experimental studies using clinical case vignettes have shown mixed results of the effect of time pressure on diagnostic accuracy. In two studies by AlQahtani [7, 8], resident participants were told, independent of their actual time, that they were either on or behind schedule with diagnosing their cases. Results showed diagnostic accuracy was negatively influenced in these group that thought they were behind schedule. In two studies, time pressure was implemented by allotting resident participants a fixed amount of time to diagnose 20–25 cases [9, 10]. Not all cases needed to be completed. Results showed that residents who were asked to diagnose 25 cases in 30 min had similar accuracy rates compared to residents that were asked to diagnose 20 cases in 40 min [10]. Furthermore, residents were more accurate when they had a lower time-to-diagnose [9].

In psychology, the effects of a time limit on performance have been studied using the Iowa Gambling Task, a psychological decision-making task [11]. In this task, participants choose cards from card decks that are either 'good' or 'bad', determined by their in-game money reward and/or punishment. The goal is to learn the rules of the decks and obtain as much in-game money as possible. Results showed that participants with a time limit had an impaired performance compared to participants without a time limit. In a follow-up study [12], the effect of the perception of time pressure was measured by manipulating the instructions for picking a card. Participants were instructed to select a card within 2 s and half of the participants were told that this provided time was insufficient. Results showed that participants who were informed that the time allotted was insufficient performed worse than those who were informed that they had sufficient time, despite the actual time allotted being the same.

While both time limits and instructions to induce subjective time pressure have been shown to affect diagnostic accuracy, the effects are not consistent and the combination of these factors has not been studied together. The current study aims to fill this gap by investigating the effects of perceived time pressure (sufficient vs. insufficient) in combination with a restricted or lenient time limit with regard to diagnostic accuracy, number of differential diagnoses, and confidence. We hypothesized that diagnostic accuracy, number of differential diagnoses, and confidence are lower when participants think they have insufficient time, regardless of actual time constraint. Furthermore, we hypothesized that the manipulations enhance each other when they 'match' (i.e. diagnostic accuracy, number of differential diagnoses, and confidence are highest in cases where participants think they have enough time and the time limit is lenient, and lowest in cases where participants think they do not have enough time and the time limit is restricted).

Materials and methods

Participants

A sample size calculation based on a two-by-two within-subjects design, an alpha-level of 0.05 and a small effect size (0.2) was performed in G*power, which resulted in an estimation of 24 participants for a desired power of 0.8. We aimed for 32–48 participants to account for possible missing data. Participants were internal medicine residents and medicine-pediatrics residents in year one to four of their residency. Medicine-pediatrics residents start with internal medicine and are therefore comparable to internal medicine residents. Participants were recruited at University of Minnesota medical school and University of Pittsburgh medical school. Participation was voluntary but participants were rewarded with a $25 gift card after completion of the task.

Materials

The experiment was programmed in Qualtrics (Qualtrics, Provo, UT), an online survey tool. Twelve cases, were written by Dutch internal medicine physicians and translated to English for use in a different study with similar participants (i.e. internal medicine residents) [13]. Eight of these cases were developed to be challenging and were used as the experimental cases. Four cases were easier and acted as filler to prevent participants from identifying the manipulations.

Study design

This study was a two-by-two within-subjects design which aimed to test the effects on diagnostic accuracy of two factors: 'time constraint' and 'perceived time pressure'. Each factor had two levels, creating four experimental conditions (see Table 1).

Table 1: Four conditions based on the 2 × 2 within-subjects design.

		Time constraint	
		Lenient time limit	**Restricted time limit**
Perceived time pressure (instruction)	**Sufficient**	Condition A: match 4 cases (2 experimental, 2 filler)	Condition B: mismatch 2 cases
	Insufficient	Condition C: mismatch 2 cases	Condition D: match 4 cases (2 experimental, 2 filler)

The time constraint was implemented by setting a time limit that was either restricted or lenient, which was respectively one standard deviation below or above the mean time to diagnose (see Supplementary Material, Appendix 1). The mean time to diagnose was based on the earlier study by Staal and colleagues that used the same cases [13]. Two exceptions were made for cases 6 and 7, where subtracting one standard deviation from the mean time to diagnose would have resulted in too little time to read the case at a normal reading pace. The timings for the first five participants were checked to adjust if necessary, but this was not needed.

Perceived time pressure was manipulated through instruction. The instruction for the perceived sufficient time read: "For the following cases, you will have enough time to diagnose the patients". The instruction for the perceived insufficient time read: "For the following cases, you will find that the allotted time is restricted".

Each case was presented on screen together with a time bar to help participants keep track of how much time they had left to read and process the case. This bar was green and became red to indicate time was running out. The bar stayed green for 85 % of the allotted time for the cases that were under the perceived-sufficient instruction, and for 60 % of the allotted time for cases under the perceived-insufficient instruction, in order to add to the perceived time pressure. If the bar (and thus the time limit) ran out, the page was forwarded to the next page where participants were able to type their most likely diagnosis and a differential diagnosis.

The two-by-two design results in four unique conditions: sufficient-lenient (A), sufficient-restricted (B), insufficient-lenient (C), and insufficient-restricted (D). For condition A and D, the length of the time limit matched what participants could expect based on the instruction. For example, in condition A, participants were told they would have enough time, and the actual time limit was also lenient. In contrast, conditions B and C had instructions that were mismatching the actual time limit. For example, in condition B, participants were told they would have enough time, but the actual time limit was restricted. To ensure that participants believed the instruction, we added two filler cases to conditions A and D.

Procedure

This study was conducted in accordance to the Declaration of Helsinki (as revised in 2013) and approved by the Medical Ethics Board of the Erasmus Medical Centre.

Before participating, participants were asked to read the information letter and sign an informed consent form. They were not fully informed of the aim of the study beforehand, but were told that the goal of the study was to measure diagnostic performance under time pressure.

The experiment was divided in two blocks (see Figure 1) which were shown in random order. After reading the instruction, participants would diagnose six cases: two matching filler cases, two matching experimental cases, and two mismatching experimental cases, in that order. We chose to have the first four cases always be 'matching' (i.e. condition A or D) so participants would not grow suspicious of the instruction. The participant either clicked forward to the next page themselves, or were forwarded automatically to the next page by the timer. They were then asked for their most likely diagnosis and for other diagnoses they had considered. Participants were not restricted in their time to enter the diagnosis, although time spend on the page was recorded.

The cases within the blocks were counterbalanced in eight versions to prevent order effects. Participants were randomly assigned to one of the versions. After participants finished both blocks, they were asked how confident they were of their diagnosis for each case on a scale of 0–100. At the end of each block and at the end of the survey, participants were asked several questions about the time pressure they felt (see Supplementary Material, Appendix 2). After completing the survey, they were given the link to the correct answers of the cases.

Data processing

Data points were excluded if participants took more than two standard deviations above the mean time to enter their diagnosis, as this could indicate that participants walked away during the case.

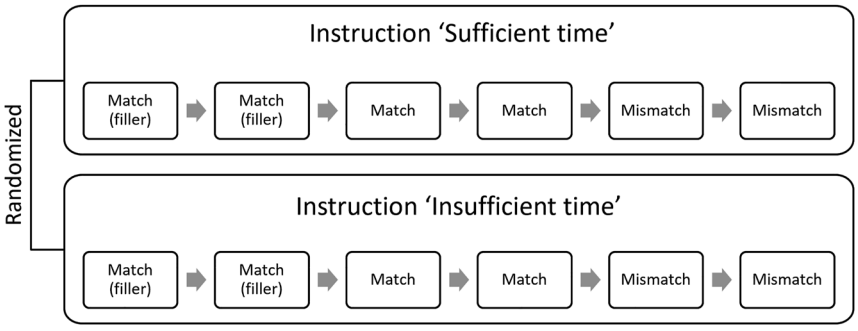

Figure 1: Order of the cases within the two blocks.

Diagnostic accuracy (1=(partially) correct, 0=incorrect) and differential diagnosis accuracy (i.e. did the differential diagnosis contain the correct diagnosis; 1=(partially) correct, 0=incorrect) were scored independently by two physicians (AO and CM) and disagreements were resolved in a meeting. A score of one could mean that the participant's diagnosis was either identical to the correct diagnosis, or close to the correct diagnosis with a small element missing. Number of diagnoses in the differential diagnosis were counted by a researcher (JH) and uncertainties were discussed with the two physicians until consensus was reached.

Analyses

Data cleaning and data analyses were performed using R (R Core Team, 2023) and R studio. Data were loaded using the qualtRics package [14]. Analyses comprised mixed models for dependent variables 'diagnostic accuracy', 'number of differential diagnoses', and 'confidence', with time constraint and perceived time pressure as independent variables, unique participant number as random intercept, and various characteristics of participants (i.e. university, specialization type, residency year) as fixed effects. The mixed models were performed using the lme4 package [15] and summarized using the gtsummary package [16].

Results

Forty-two residents participated in this study (see Table 2). They were mostly female (54.8 %), internal medicine residents (90.5 %) at University of Pittsburgh Medical School (85.7 %). Most residents were in Year 1 (38.1 %) or Year 3 (28.6 %) of their residency.

Table 2: General participant characteristics.

Gender, n (%)	Male	18 (42.9)
	Female	23 (54.8)
	Other	1 (2.4)
Specialization, n (%)	Internal medicine	38 (90.5)
	Medicine-pediatrics	4 (9.5)
Residency year, n (%)	Year 1	16 (38.1)
	Year 2	8 (19.0)
	Year 3	12 (28.6)
	Year 4	6 (14.3)
University, n (%)	Minnesota	6 (14.3)
	Pittsburgh	36 (85.7)

Manipulation checks, case performance, and outliers

Outliers

Outliers were calculated for the time spend on entering the diagnosis and differential diagnosis. The mean time on this page was 58.5 s with a standard deviation of 98.5 s, meaning that all data points above 255.5 s were considered outliers. A total of 10 data points from 6 participants were deleted as they took between 261 and 1,203 s (i.e. 4.4–20 min) to write down their diagnosis.

Manipulation checks

A paired-samples t-test revealed that participants reported experiencing more time pressure in the block of cases with the 'insufficient time' instruction (mean=49.7, SD=21.2) than in the block of cases with the 'sufficient time' instruction (mean=30.5, SD=20.0; t (41)=6.69, p<0.001). Most participants felt that the time pressure varied between cases within a block; during some cases they experienced more time pressure than during other cases (n=31, 73.8 % for sufficient block, and n=37, 88.1 % for insufficient block).

Table 3: Results of the mixed model logistic regressions for diagnostic accuracy, showing Odds Ratio, 95 % confidence intervals, and p-values for all fixed effects.

Dependent variable	Random effects (intercept)	Fixed effects	OR	95 % CI	p-Value
Diagnostic accuracy	Participant	Perceived time pressure	1.57	0.83 to 2.97	0.2
		Time constraint	1.93	1.02 to 3.65	0.044[a]
		Interaction	0.52	0.21 to 1.27	0.2
		University	0.98	0.48 to 2.00	> 0.9
		Residency year	1.16	0.93 to 1.44	0.2
		Specialization type	0.32	0.13 to 0.81	0.016[a]

OR, odds ratio; CI, confidence interval.

Most participants (n=37, 88.1 %) indicated that the time bar contributed to their feeling of time pressure. Participants reported in an open text field that the time bar made them feel 'more rushed' and as if they had to 'read faster', especially when it 'neared the end' (i.e. if the timer almost ran out). Some also reported on the perceived-insufficient condition, that the knowledge of having less time made them 'skim much faster than with the other cases'. Some participants reported on the time bar in general, saying that it was 'distracting', or that the presence of the time bar made it 'hard to focus'.

Participants were more often automatically forwarded to the next page by the timer when they were working on cases in the restricted time constraint conditions (n=47, 28.0 %) than when they were working on cases in the lenient time constraint conditions (n=13, 7.7 %; p<0.001).

Furthermore, it is important to note that participants in the sufficient condition only saw the bar turn red in 30 % of cases, and participants in the insufficient condition only saw the bar turn red in 53 % of cases. There was a trend toward a significant interaction between the perceived time pressure and the color of the time bar (OR=0.44, 95 % CI: 0.17 to 1.12, p=0.085). Participants in the perceived-sufficient conditions were somewhat more correct when the bar stayed green (n=117, accuracy=51.3 %) and somewhat less correct when the bar turned red (n=51, accuracy=29.4 %) compared to the insufficient conditions (respectively, n=79, accuracy 43.0 %; n=89, accuracy=40.4 %).

Case performance and distribution

We calculated the percentage correct for all cases to estimate how well participants performed in general. Data show participants especially had a low diagnostic accuracy on Cases 1, 4 and 5 (7 , 5, and 29 % correct, respectively. Other cases had accuracy scores between 44 and 64 %. Cases were fairly equally distributed among conditions due to counterbalancing. Because of this, we chose not to further control for case in the analyses.

Main analyses

Results of the mixed model show only an effect of time constraint (OR=1.93, 95 % CI: 1.02 to 3.65, p=0.044; see Table 3), for which a restricted time limit is associated with a lower accuracy score (39.4 % correct) compared to a lenient time limit (47.2 % correct; see Table 4).

Furthermore, of relevant participant characteristics, specialization type had a significant effect on the data (OR=0.32, 95 % CI: 0.13 to 0.81, p=0.016; see Table 3), indicating that medicine-pediatrics residents performed worse (overall percentage correct=23.3 %) than internal medicine residents (overall percentage correct=45.3 %).

Exploratory analyses

In addition to diagnostic accuracy, we were interested in whether the perceived time pressure and the actual time constraint had an effect on other variables, namely: number of diagnoses in the differential diagnosis and confidence in the most likely diagnoses.

Table 4: Diagnostic accuracy as percentage correct for all conditions and manipulations.

			Time constraint		Total
			Lenient time limit	Restricted time limit	
Perceived time pressure (instruction)	Sufficient	n	80	83	163
		% correct	45.0 %	44.6 %	44.8 %
	Insufficient	n	81	82	163
		% correct	49.4 %	34.1 %	41.7 %
Total		n	161	165	
		% correct	47.2 %	39.4 %	

Number of differential diagnoses

Results show no effect of time constraint or perceived time pressure, nor an interaction effect, on the number of diagnoses participants wrote down in their differential diagnosis, see Table 5.

Confidence

Results show that there was an effect of perceived time pressure (β=6.84, SE=3.45, t=1.98, p=0.048) and time constraint (β=8.75, SE=3.47, t=2.52, p=0.012) on confidence in the most likely diagnosis. No interaction effect was found. This indicates that participants were more confident when they thought they had sufficient time regardless of the actual time limit (mean=57.2; vs. mean=53.9 for the perceived-insufficient conditions), and furthermore that they were more confident when they had more time regardless of how much time they thought they had (mean=58.4; vs. mean=52.8 for the perceived-insufficient conditions, see Table 5).

Discussion

This study investigated the effects of perceived time pressure and actual time constraints on diagnostic accuracy, number of differential diagnoses, and confidence in an online experiment with written clinical cases. In summary, the current study shows an effect of a restricted time limit on diagnostic accuracy. Although participants reported feeling more time pressure in the perceived-insufficient conditions, no effect of perceived time pressure was found and no interaction effect was present. This means that, contrary to the expectations and previous research [7, 8], the perception of time pressure did not influence diagnostic accuracy. In addition to the most likely diagnosis, participants were asked to name other diagnoses that they considered. It was hypothesized that the number of differential diagnoses that were considered would be higher when participants perceived to have enough time. However, the number of submitted differential diagnoses did not differ between conditions. Lastly, participants were asked about their confidence in the most likely diagnosis on a scale of 0–100. Results show that participants were more confident both when they had a lenient time limit and also when they received the instruction that they would have sufficient time.

It seems that most of the effect on diagnostic accuracy in this study stems from the actual time constraint. This means that participants performed worse in cases were the time limit was restricted time compared to the lenient time limit cases. Participants were not aware of the exact time limit for each case, but they were more likely to reach the end of the timer while reading the case, and so were more likely to be automatically forwarded to the next page to fill in their most likely diagnosis and differential diagnosis, possibly disrupting their thought process. In addition, they were possibly aware of the passing of time due to the time bar that was present at the top of the page, which logically moved faster when there was a shorter time limit.

These results for the effect of the time limit on diagnostic accuracy seem to be different from the study by Norman and colleagues [10], in which residents had a time limit for a total of 20–25 cases and participants who were allotted less time did not have a lower diagnostic accuracy than those who were allotted more time. However, in their experiment they chose a larger range of case difficulty that included easier cases where the participants were likely able to use pattern-recognition to diagnose them correctly. The current study

Table 5: Results for all conditions and manipulations on number of diagnoses in the differential diagnosis and confidence in the most likely diagnosis.

			Time constraint		Total
			Lenient time limit	Restricted time limit	
Perceived time pressure (instruction)	**Sufficient**	n	80	83	163
		Number of diagnosis, mean (SD)	2.56 (1.71)	2.43 (1.77)	2.50 (1.74)
		Confidence, mean (SD)	58.3 (24.9)	56.2 (28.7)	57.2 (26.8)
	Insufficient	n	81	82	163
		Number of diagnosis, mean (SD)	2.41 (1.66)	2.57 (1.63)	2.49 (1.64)
		Confidence in most likely diagnosis, mean (SD)	58.5 (25.8)	49.4 (22.4)	53.9 (24.5)
Total		n	161	165	–
		Number of diagnosis, mean (SD)	2.48 (1.68)	2.50 (1.70)	–
		Confidence in most likely diagnosis, mean (SD)	58.4 (25.3)	52.8 (25.9)	–

used cases that were developed to be challenging. This means that participants might have needed more time than they were allotted, because they could not rely on pattern-recognition. The restricted time limit, visible through the live time bar, therefore could have resulted in lower accuracies.

It is somewhat surprising that no main effect of perceived time pressure on diagnostic accuracy was found, as this effect has previously been established [7, 8]. This discrepancy could in part be explained by a difference in method between the current study and the studies by AlQahtani and colleagues [7, 8]. In their studies, participants were shown a message after diagnosing each case telling the participant that they were behind schedule and they needed to hurry. The current study relied on the message introducing each block of six cases. It is possible that this message did not have enough effect, or was forgotten after a couple of cases, diminishing the effect of the perceived-insufficient condition.

Interestingly, results showed an effect of both perceived time pressure and time constraint on confidence, indicating that participants felt more confident about their most likely diagnosis when they thought they would have enough time or when the time limit was more lenient. The reasons for this could be that the lenient time limit and low perceived time pressure allowed for double-checking of the answers.

Additionally, there are some indications that the time bar has inadvertently added a layer of time pressure to all the conditions. When participants were asked if they thought the time bar contributed to their feeling of time pressure, most of them answered 'yes'. They also described that they were feeling rushed by the time bar moving and changing color from green to red, and that it was distracting them from reading the case. Because the time bar was present during all conditions, it is possible that there was some form of perceived time pressure in all conditions induced by the time bar.

Strengths and limitations

In the current study we made an effort to distinguish the effects of perceived time pressure and a restricted or lenient time limit, which has not been tried before with written clinical cases. A relatively large group of participants was recruited, almost twice as many as needed for sufficient power. Furthermore, the experiment was pre-registered, used counterbalancing and randomization to account for order effects, and used cases that were successfully used in a previous study.

However, there were some limitations to this study as well. First, the cases might have been too difficult for the study population, indicated by the generally low diagnostic accuracy. This may have caused floor effect for at least two of the cases where accuracy was below 10 %. Second, some noise may have been caused by the fixed order of conditions within the blocks, possibly adding a fatigue effect to the cases where the actual time limit did not match the instructions of the allotted time, as these were always last. Third, the presence of the time bar might have introduced some time pressure or stress in all conditions.

Lastly, a common limitation of experimental settings is their artificiality, as they do not reflect clinically realistic settings. However, studying the effects of perceived time pressure and time constraints in a more controlled environment allows us to separate them, which helps to gain a better understanding of their effects on accuracy.

Conclusions

In conclusion, this study we investigated the difference between perceived time pressure via instruction and time constraint via a lenient or restricted time limit. Results showed that a restricted time limit has an effect on diagnostic accuracy. Although participants reported feeling more time pressure when they received the perceived-insufficient time instruction and felt less confident on cases in these conditions, no effects of the instruction of perceived time pressure were found for diagnostic accuracy. More research is needed to investigate the robustness of the effects found in this study. In addition, different forms of (perceived) time pressure should be investigated in various settings to gain a better understanding of the effects on diagnostic accuracy.

Acknowledgments: The authors wish to thank Robert Trowbridge for recruiting participants for the pilot of this study.

Research ethics: This study was conducted in accordance to the Declaration of Helsinki (as revised in 2013) and approved by the Medical Ethics board of the Erasmus Medical Centre.

Informed consent: Informed consent was obtained from all individuals included in this study.

Author contributions: JH, APJO, SM, and LZ were involved in the conception of the study. JH, APJO, CNM, and LZ performed the data collection. JH performed the analyses. JH, APJO, CNM, CW, and LZ were involved in the interpretation of the results. JH and LZ drafted the manuscript. APJO, CNM, SM, CW, and LZ gave critical feedback on the manuscript. All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Use of Large Language Models, AI and Machine Learning Tools: None declared.
Conflict of interest: The authors state no conflict of interests.
Research funding: Funding for the gift cards in this study was provided by the Diagnostic Error Curriculum and Research Fund at the University of Minnesota.
Data availability: The raw data can be obtained on request from the corresponding author.
Pre-registration: This study was pre-registered at OSF (https://osf.io/ptdmg).

References

1. Linzer M, Manwell LB, Williams ES, Bobula JA, Brown RL, Varkey AB, et al. Working conditions in primary care: physician reactions and care quality. Ann Intern Med 2009;151:28–36.
2. Linzer M, Konrad TR, Douglas J, McMurray JE, Pathman DE, Williams ES, et al. Managed care, time pressure, and physician job satisfaction: results from the physician worklife study. J Gen Intern Med 2000;15:441–50.
3. Mawardi BH. Satisfactions, dissatisfactions, and causes of stress in medical practice. JAMA 1979;241:1483–6.
4. Groenewegen PP, Hutten JBF. Workload and job satisfaction among general practitioners: a review of the literature. Soc Sci Med 1991;32:1111–9.
5. Sehlen S, Vordermark D, Schäfer C, Herschbach P, Bayerl A, Pigorsch S, et al. Job stress and job satisfaction of physicians, radiographers, nurses and physicists working in radiotherapy: a multicenter analysis by the DEGRO Quality of Life Work Group. Radiat Oncol 2009;4:6.
6. Welker JA, Huston M, McCue JD. Antibiotic timing and errors in diagnosing pneumonia. Arch Intern Med 2008;168:351–6.
7. Alqahtani DA, Rotgans JI, Mamede S, Alalwan I, Magzoub MEM, Altayeb FM, et al. Does time pressure have a negative effect on diagnostic accuracy? Acad Med 2016;91:710–6.
8. Alqahtani DA, Rotgans JI, Mamede S, Mahzari MM, Al-Ghamdi GA, Schmidt HG. Factors underlying suboptimal diagnostic performance in physicians under time pressure. Med Educ 2018;52:1288–98.
9. Sherbino J, Dore KL, Wood TJ, Young ME, Gaissmaier W, Kreuger S, et al. The relationship between response time and diagnostic accuracy. Acad Med 2012;87:785–91.
10. Norman G, Sherbino J, Dore K, Wood T, Young M, Gaissmaier W, et al. The etiology of diagnostic errors: a controlled trial of system 1 versus system 2 reasoning. Acad Med 2014;89:277–84.
11. Cella M, Dymond S, Cooper A, Turnbull O. Effects of decision-phase time constraints on emotion-based learning in the Iowa Gambling Task. Brain Cognit 2007;64:164–9.
12. DeDonno MA, Demaree HA. Perceived time pressure and the Iowa gambling task. Judgment Decis making 2008;3:636.
13. Staal J, Alsma J, Mamede S, Olson A, Prins-van Gilst G, Geerlings S, et al. The relationship between time to diagnose and diagnostic accuracy among internal medicine residents: a randomized experiment. BMC Med Educ 2021;21:1–9.
14. Ginn J, O'Brien J, Silge J. qualtRics R package version 3.2.0; 2024. Available from: https://github.com/ropensci/qualtrics.
15. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J Stat Software 2015;67:1–48.
16. Sjoberg D, Whiting K, Curry M, Lavery J, Larmarange J. Reproducible summary tables with the gtsummary package. The R Journal 2021;13:570–80.

Supplementary Material: This article contains supplementary material (https://doi.org/10.1515/dx-2024-0125).