Clinical diagnosis is complex, but why? Identifying factors contributing to diagnostic complexity of infectious diseases in emergency medicine

Swagatam Dey [1,✉], **Shrirang Joshi** [2], **Prakash Ranjan Mishra** [2] **and Pramod Khadilkar** [1]

[1] Indian Institute of Technology Delhi, India, [2] All India Institute of Medical Sciences New Delhi, India

✉ swagatam.dey@design.iitd.ac.in

ABSTRACT: A major cause of diagnostic errors is the underlying complexity caused by patient presentations and the context in which diagnosis is being undertaken. This is especially true for settings like emergency medicine and disease spectrums like infectious diseases. To design artefacts that counter such errors, it is essential to map the factors contributing to diagnostic complexity. However, existing complexity assessment methods in healthcare are limited in scope. Addressing this gap, our work operationalises a complexity estimation tool to identify factors contributing to the diagnostic complexity of 10 infectious disease cases in an emergency medicine setting. Our objective findings are further validated by a strong correlation with the difficulty perceived by attending doctors. The work provides a basis for the design of targeted interventions aiming to mitigate complexity in diagnosis.

KEYWORDS: complexity, diagnostic complexity, diagnostic errors, decision making, biomedical design

1. Introduction

Over the last decade, the behavioural design (BD) domain has actively promoted health-conscious and preventive behaviours among individuals and communities. The domain synthesises theories from cognitive psychology, behavioural science and design to develop artefacts that ethically influence desired behaviours in areas such as sexual health and contraception, smoking cessation and healthy lifestyle practices (Bay Brix Nielsen et al., 2021; Cowdell & Dyson, 2019; Khadilkar & Cash, 2020). However, beyond preventive health, a critical yet underexplored facet of health and well-being for behavioural designers is clinical diagnosis and its underlying complexity. A crucial component of healthcare, diagnosis helps estimate the systems and pathologies associated with a patient's presentation, further guiding relevant treatment and recovery strategies (Bornstein & Emler, 2001). Diagnostic errors are often consequential, leading to significant patient morbidity and mortality (Croskerry, 2009b, 2009a). They also contribute to increased medico-legal cases and burnout in the medical fraternity. If we consider the US healthcare system alone, Newman-Toker et al. (2019) estimated that 40,000 to 4 million individuals may be exposed to misdiagnosis-related harm. Diagnostic errors are a major threat to the safety and well-being of a population and a critical area for BD to intervene. This is because errors are essentially a result of, or are, "wrong or undesirable behaviours" (Dey et al., 2023) which is either flawed plan formulation (problem-solving behaviour) or incorrect execution of well-formed plans (monitoring behaviour) (Reason, 1990).

Multiple reports have identified diagnostic complexity and its imposed cognitive load as critical antecedents to diagnostic errors (Croskerry & Clancy, 2025; D. Newman-Toker et al., 2023). Clinicians often diagnose under intense emotional and physiological states in stressful and ambiguous contexts.

These situations are more prominent in high-stakes settings like emergency medicine (EM), where effective patient disposition under extreme time pressure is a critical demand (Calder et al., 2015; Wears, 2009). Such contextual factors and a patient's inherent conditions make the diagnostic process complex and error prone. A disease spectrum which is highly vulnerable to such errors is infectious diseases (ID) (D. E. Newman-Toker et al., 2019, 2021). The vulnerability arises from the spectrum's underlying unpredictability, common syndromic presentations across multiple infections and the potential for global repercussions (Roosan et al., 2017). Identifying and mapping the causes of diagnostic complexity within a clinical setting is necessary during the problem identification and framing stages of a design process. Furthermore, the understanding can inform the development of evaluation criteria to test existing and proposed solutions. These insights are crucial for designing behaviour-centred and contextually relevant interventions like information management software, workflows, patient report/form designs, medical device designs, diagnostic aid designs, etc.

Despite the opportunities, current methods utilised for estimating diagnostic complexity in healthcare are limited in scope. As a result, they can't effectively elucidate the factors contributing to the diagnostic complexity of disease cases (Islam et al., 2016; Roosan et al., 2017). Among the few studies addressing diagnostic complexity, Roosan et al's (2017) operationalised a task complexity estimation framework by Liu & Li (2012) to identify the complexity contributory factors (CCFs) associated with the handling of infectious disease cases. However, they only considered interactions within the clinician's team as a data source. To holistically identify the factors contributing to diagnostic complexity, it is essential to observe the diagnostic interactions between clinicians, teams, patients and caregivers across the diagnostic continuum. Another major limitation of the past research was the discrepancy between the perceived difficulty rated by the clinicians and the observed objective complexity determined by the authors.

Addressing these gaps, the paper identifies factors contributing to the diagnostic complexity of 10 ID cases in an EM setting. In doing so, we demonstrate the operationalisation of a complexity estimation tool for diagnostic tasks in the concerned setting. Although this study specifically focuses on ID and EM, it aims to provide a generalised foundation for utilising the tool across various clinical settings and for speciality cases. This can help define the behavioural and human factors issues associated with the diagnostic workflow of a concerned clinical setting and generate actionable insights towards improving diagnostic accuracy. The work is a subset of a collaborative study between the Department of Design, Indian Institute of Technology Delhi and the Department of Emergency Medicine, All Indian Institute of Medical Sciences (AIIMS), New Delhi. Appropriate ethical approval was taken from the Institute Ethics Committee of both institutes to conduct the study.

2. Study site

2.1. Site

The study was conducted in the Emergency Medicine 1 (EM 1) section of the Department of Emergency Medicine, AIIMS New Delhi. Although patients with infectious diseases entering EM 1 can be moved to EM 2 (Gynaecological and Surgical Section) and EM3 (Observation Section), the scope of our study was limited to the diagnostic interactions within the EM 1 facility to maintain consistency in data.

2.2. Patient flow

Following the AIIMS Triage Protocol (Sahu et al., 2020), incoming patients are brought to a *Triage area* (**A** and **B**). Here, the patient's presentation, history, and earlier reports are evaluated by attending clinicians while their current vitals are recorded. This is followed by triaging the patients into Green (**E**mergency **S**everity **I**ndex 5, i.e. *ESI 5*), Yellow (*ESI 3 and 4*) and Red (*ESI 1 and 2*) (Sahu et al., 2020). If marked green, the patient is referred to the outpatient department after primary treatment. A yellow patient marked ESI 3 and 4 is moved to areas marked **D1** (*Yellow Counter*) until there is a bed available in **D2** (*Yellow Area*). Patients marked ESI 2 are moved to **C1** (*Red Counter*) before transferring to **C2** (*Red Area*), while ESI 1 patients are directly moved to **C2**. Area **E** acts as a *coordinating area* between the healthcare providers and other specialized departments, while areas **F** and **G** are *nursing stations*. Figure 1 is a schematic representation of the mentioned patient flow and spatial layout of EM1.
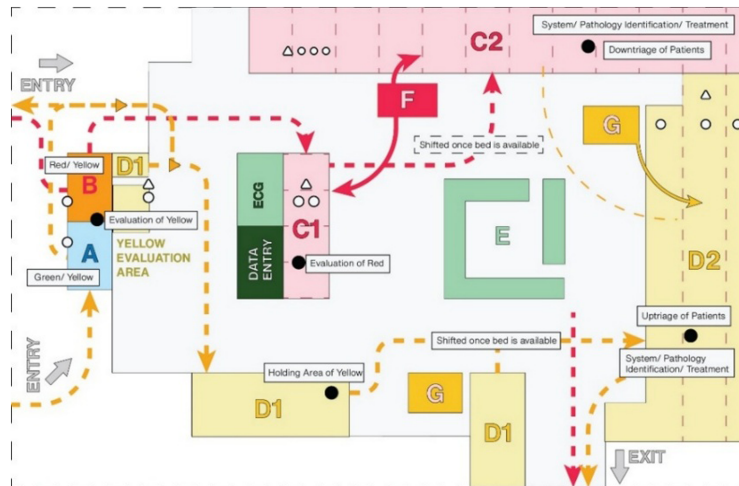
Figure 1. Schematic representation of a patient flow within the EM 1 facility

2.3. Spatial mapping of diagnostic subtasks according to the patient flow

The process of diagnosis consists of four primary subtasks.

a) *Presentation and Perception Formation (P)*   Clinicians are presented with the patient's symptoms and signs (Croskerry, 2009a). Based on the interaction, they begin preliminary examinations. In an emergency context, this is when potentially affected systems are identified.

b) *Hypothesis Generation (H)*   Based on the patient's age, sex, race, and presentations, clinicians further develop several diagnostic hypotheses (Kassirer & Kopelman, 1989). At this phase, the potential pathology of the patient is hypothesised.

c) *Gathering and Interpretation of Evidence (E)*   Clinicians assess the tests and evidence required to modify the relative probability of each hypothesis based on the ones that are developed. This aids in comparing multiple hypotheses (Bornstein & Emler, 2001).

d) *Verification (V)*   Final diagnosis is formulated by assessing the hypotheses to determine the one accommodating most of the patient's manifestation and test results (Bornstein & Emler, 2001).

Although sequential, these subtasks are performed dynamically across multiple locations by different clinicians of varying expertise within a clinical setting (Zwaan & Singh, 2015). Each of the locations within the concerned setting imposes distinct contextual demands that can influence the performance of the subtasks. Hence, to simultaneously observe and identify relevant complexity factors for multiple ID cases, we spatially mapped the diagnostic subtasks according to the patient flow within EM 1. This allowed an optimal positioning of observers to study diagnostic subtasks within their specific environment demands. The mapping of diagnostic subtasks to designated regions was informed by direct observation of clinical workflows within EM 1, undertaken during a week-long pilot study. The insights were further discussed, modified and validated after extensive discussions with the chief medical officer and senior resident clinicians engaged in patient triage and diagnosis. This ensured that our observation framework aligned with the real-world diagnostic workflow within the setting.

For cases marked **ESI 2,3**, and **4**, **A** and **B** were defined as areas where a patient case's presentation and perception formation stage is undertaken. **C1** and **D1** were identified as the areas where hypothesis generation and evidence interpretation take place, and verification can happen in **C1, C2, D1**, and **D2**. For cases marked **ESI 1**, all the diagnostic subtasks were undertaken in area **C2**.

While this mapping reflects the specific workflow of EM 1, it can be modified for other clinical settings based on variables like patient acuity, departmental configuration, and resource allocation. Prospective users of the tool may adjust the mapping according to these factors to align with their respective ward environments and workflow.

3. Objective diagnostic complexity estimation tool

The diagnostic complexity estimation tool developed by the first and last author effectively quantifies the objective complexity of diagnosing a patient presentation and further justifies it through causalities. The tool consists of the 44 identifying criteria and their associated clinical probes derived from the 25 CCFs

mentioned by Liu & Li (2012), arranged in a tabular format. To determine the criteria and probes, we reviewed the key literature through which Liu et al. identified the CCFs. A total of 27 papers were reviewed. Furthermore, these probes are classified into 33 disease-specific (X) and 34 context/ environment-specific (Y) probes, catering to patient presentation and context/ environment-imposed complexity. These probes are then provided with a binary scoring mechanism (1 for presence and 0 for absence of the corresponding CCF) and the scoring criteria to mark their relevance across the four diagnostic subtasks (refer to Table 2). Users are required to score each probe retrospectively based on the observed diagnostic interaction between patients and clinicians across the four subtasks. While the Y probes are generic and can be scored by individuals without clinical expertise based on direct observation, scoring the X probes requires extensive consultation with medical experts and clinicians. Table 1 presents an example of the identifying criteria, probes, and scoring criteria for the complexity contributory factors of Goal/ Output Conflict and Redundancy, Process Clarity and Presentation Heterogeneity, and Process Repetitiveness (Refer Liu & Li (2012) for details).

Table 1. Examples of clinical probes, sources and scoring criteria for CCFs

CCF	Identifying Criteria	Clinical Probes	Method	Probe Type	Scoring Criteria
Goal/ Output Conflict and Redundancy	The subtasks within the primary task are conflicting in nature	A. Is there any instance wherein fulfilling the goal of a diagnostic subtask faces conflict due to the fulfilment of another subtask's goals?	Both	Y	Conflict- 1 Redundant- 0
Process Clarity	One ideal path is absent, and multiple possibilities necessarily need consideration	B. Is differential diagnosis needed, or can the clinician pinpoint one system/ pathology?	Observable	X	Yes- 1No- 0
	Elimination of bad options is not easy	C. Is how the vitals and test results are interpreted, capable of eliminating possible system/ pathologies?	Focus Group	X	Yes- 0No- 1
	The procedures to be undertaken are highly dependent on each other	D. Are the subtask steps to be undertaken highly dependent on each other?	Focus Group	X	Yes-1No-0
	The link between the actions undertaken towards judgment and the outcome of the action is unclear or uncertain	E. There is no clear indication/ feedback between the diagnosis and the subsequent treatment or procedures undertaken	Focus Group	X	Yes- 1No-0
Presentation Heterogeneity, Process Repetitiveness	The representation of information presented is scattered or heterogeneous and doesn't aid in relating and correlating the data with undertaken measures	F. Are the modes of presenting the test results and patient records vitals scattered and not aiding in defining the relation and correlation of patient presentations and other information?	Both	Y	Yes- 1No- 0

An example of a probe's potential relevancy across the four diagnostic subtask stages, along with their highest achievable scores, is provided in Table 2. Probes A, B, C, and D can be relevant across each stage, and hence the highest achievable score (HAS) for these probes is 4, while probe E is only relevant for the verification stage, and hence its HAS is 1. Probe F caters to two CCFs, i.e., Presentation Heterogeneity and Process Repetitiveness, and it is relevant across Presentation and Perception formation and Verification. In this case, even though the HAS for this probe is 4, it is to be divided among the two CCFs,0, i.e., 2 per CCF.

Table 2. Applicability and scoring System of each probe across Four diagnostic subtasks

Clinical Probes	P	H	E	V	HAS
A. Is there any instance wherein fulfilling the goal of a diagnostic subtask faces conflict due to the fulfilment of another subtask's goals?	1	1	1	1	4
B. Is differential diagnosis needed, or can the clinician pinpoint into one system/ pathology?	1	1	1	1	4
C. Is how the vitals and test results are interpreted, capable of eliminating possible system/ pathologies?	1	1	1	1	4
D. Are the subtask steps to be undertaken highly dependent on each other?	1	1	1	1	4

(Continued)

Table 2. Continued.

Clinical Probes	P	H	E	V	HAS
E. There is no clear indication/ feedback between the diagnosis and the subsequent treatment or procedures undertaken	0	0	0	1	1
F. Are the modes of presenting the test results and patient records vitals scattered and don't aid in defining the relation and correlation of patient presentations and other information?	1/1	0	0	1/1	4

4. Protocol

10 infectious patient cases were observed across their diagnostic continuum by three observers, and field notes around each diagnostic subtask were undertaken. Observer 1 was stationed at areas A and B (*Triage*) and would be informed by the attending clinicians at triage once they were presented with cases carrying infectious disease symptoms. If the case was marked for the *yellow area*, observer 2 would be alerted by observer 1, who would further observe the patient's case across areas D1 and D2. If a case was marked for the *red area* at the triage, observer 3 would hand over the case for areas C1 and C2. The observers captured the contextual and environmental conditions in which the diagnostic tasks were being undertaken. The observation of a patient case would close when the attending clinician informed the observers that they had identified a probable diagnosis.

Prior to the main study, the observers underwent a week-long immersion at the site to counter observer bias. Cases observed during this immersion period were utilised to assess the inter-rater reliability of the three observers using *Fleiss Kappa*. This exercise was repeated twice using the immersion period data until moderate agreement (0.449) was reached. Post the observation period, focus group discussions consisting of the Chief Medical Officer, 2 attending Senior Residents and 1 observer were conducted on the observed cases. Based on the discussions, the disease-specific (X) probes inherent to the case were scored. Next, the attending clinicians were asked to rate the difficulty of the case on a scale of 1 to 5 to record their perceived difficulty.

5. Data analysis and results

The HAS for each complexity contributory factor (CCF) was calculated by adding the HAS for all the probes catering to it (see Table 4). Two CCFs, Presentation Compatibility and Presentation Format, were found to be irrelevant in a clinical context. Variety was found to be relevant across all the cases. The total objective complexity score calculated was 205.

Table 3. HAS for CCFs across each diagnostic subtask and total HAS for each probe

CCF and Relevant Probe	HAS (P)	HAS (H)	HAS (E)	HAS (V)	HAS for CCF
1. Quantity (G/O) (Probes- 4,5,6,7,8,9)	4	4	3	4	15
2. Quantity (I) (Probes- 18,19,20,24,25)	4	4	3	4	15
3. Quantity of actions/steps (Pro) (Probes- 5, 45)	1	2	1	2	6
4. Quantity of Paths (Pro) (Probes- 40)	1	1	1	1	4
5. Diversity (I) (Probes- 23)	1	1	1	1	4
6. Goal+ Output Clarity (Probes- 1,2,3,15,16,17)	4	2	3	4	13
7. Input Clarity (Probes- 26,27,28,29)	4	4	3	4	15
8. Process Clarity (Probes- 38,39,41,55)	3	3	3	4	13
9. Unstructured Guidance (Probes- 22)	1	1	0	0	2
10. Goal+ Output Conflict & Redundancy (Probes- 10)	1	1	1	1	4
11. Input Conflict- Redundancy (Probes- 3, 24)	2	1	1	2	6
12. Process Conflict (Probes- 42,43)	2	2	2	2	8
13. Goal+ Output Change (Probes- 11,12,13,14)	4	1	1	1	7

(Continued)

Table 3. Continued.

CCF and Relevant Probe	HAS (P)	HAS (H)	HAS (E)	HAS (V)	HAS for CCF
14. Rate of Input Change (Probes- 21,31)	2	2	2	2	8
15. Inaccuracy of Inputs (Probes- 32,33,34,35)	3	3	4	4	14
16. Non-routine Events as Input (Probes- 21)	1	1	1	1	4
17. Process Repetitiveness (Probes- 21, 61)	2	1	1	2	6
18. Input Mismatch (Probes- 36,37)	1	1	1	1	4
19. Presentation Heterogeneity (Probes- 61)	1	0	0	1	2
20. Pro Cognitive Requirement (Probes- 44,45, 53,54)	3	3	1	4	11
21. Pro Physical Requirement (Probes-46,47,48,49,50,51,52)	6	4	5	5	20
22. Concurrency (Probes- 59,60)	2	2	2	2	8
23. Pressure (Probes- 56,57,58)	4	4	4	4	16

Next, we calculated the contribution percentage of each CCF for the 10 observed cases and the complexity percentage of each diagnostic stage. Due to space constraints, we elaborate on three cases, while the summary of the other seven cases is provided below in Table 7.

Table 4. Diagnostic complexity estimation of case 1

Case Summary- *80-year-old female with decreased responsiveness for 1 day. The patient presented in altered mental sensorium and shock. The patient was intubated in flow Glasgow Coma Scale and resuscitated with fluids and anotropes. liver function test (LFT) and renal function test (RFT) deranged.*

Emergency Diagnosis- *Meningoencephalitis with acute kidney injury and acute liver injury with sepsis with septic shock.*

Perceived Difficulty- 3	Objective Complexity Score- 64 X Score- 43 Y Score- 21
Complexity of Diagnostic Task Stages (in descending order)-	Hypothesis Generation (35.08%), Verification (29.82%), Presentation and Perception Formation (28.07%), Gathering and Interpretation of Evidence (19.29%).
Top 5 CCF (in descending order)-	Input Diversity (100%), Quantity of Paths (100%), Presentation Heterogeneity (100%), Input Quantity (66.67%) and Goal/ Output Quantity (53.33%).

Table 5. Diagnostic complexity estimation of case 4

Case Summary- *21-year-old male with a known case of old pulmonary tuberculosis presented with hemoptysis.*

Emergency Diagnosis- *Reactivation of pulmonary tuberculosis.*

Perceived Difficulty- 1	Objective Complexity Score-41 X Score- 21 Y Score- 20
Complexity of Diagnostic Task Stages (in descending order)-	Hypothesis Generation (22.92%), Gathering and Interpretation of Evidence (20.45%), Presentation and Perception Formation (19.29%) and Verification (17.86%)
Top 5 CCF (in descending order)-	Input Diversity (100%), Goal/ Output Conflict (50%), Input Conflict (50%), Input Quantity (46.66%), Goal/ Output Quantity (40%).

Table 6. Diagnostic complexity estimation of case 7

Case Summary- *70-year-female with hypertension/ diabetes mellitus presented with loose stools, vomiting and fever.*

Emergency Diagnosis- *Acute gastroenteritis with sepsis with septic shock with pneumothorax.*

Perceived Difficulty- 4	Objective Complexity Score- 113 X Score-73 Y Score- 41
Complexity of Diagnostic Task Stages (in descending order)-	Hypothesis generation (60.41%), Presentation and Perception Formation (57.89%), Verification (53.57%) and gathering and Interpretation of Evidence (47.72%)
Top 5 CCF (in descending order)-	Presentation Heterogeneity (100%), Input Conflict (100%), Quantity of Paths (100%), Input Diversity (100%), Non-Routine Events as Inputs (100%)

Table 7. Summarised diagnostic complexity estimation of remaining cases

Case	Complexity Score	Perceived Difficulty	X Score	Y Score	Top 5 CCFs
2	62	3	42	20	Time Concurrency (100%), Input Diversity (100%), Quantity of Paths (100%), Time Pressure (75%), and Input Quantity (73.33%).
3	60	2	42	18	Quantity of Paths (100%), Input Diversity (100%), Input Clarity (53.33%), Process Repetitiveness (50%), Time Pressure (43.75%)
5	29	2	1	28	Input Diversity (100%), Input Quantity (100%), Goal/ Output Quantity (40%), Time Pressure (25%) and Input Clarity (20%)
6	54	2	34	20	Input Diversity (100%), Quantity of Paths (100%), Presentation Heterogeneity (100%), Process Clarity (84.61%), Goal/ Output Quantity (33.33%)
8	27	2	16	11	Input Diversity (100%), Goal/ Output Change (42.82%), Process Clarity (38.46%), Process Cognitive Requirements (2.27%), Input Quantity (26.66%), Goal/ Output Quantity (26.66%)
9	41	3	22	19	Input Diversity (100%), Goal/ Output Quantity (60%), Quantity of Action Steps (50%), Process Clarity (48.15%), Process Cognitive Requirement (36.36%)
10	32	2	26	6	Input Diversity (100%), Input Quantity (73.33%), Input Conflict (50%), Process Clarity (46.15%), Process Cognitive Requirement (27.27%)

6. Discussions

A major cause of diagnostic errors is the underlying complexity of the diagnostic process. To design and deploy behavioural centred artefacts that effectively counter these errors, it is inherent to identify the causes of complexity. Through our work, we successfully demonstrate the operationalisation of a complexity estimation tool to identify the objective complexity of 10 infectious disease cases. Undertaking a *Pearson correlation test* between the perceived difficulty scores mentioned by clinicians attending the cases and our objective scores revealed a strong correlation of 0.739. Our analysis also revealed Hypothesis Generation (H) as the most complex subtask, followed by Verification of Evidence (V), Gathering and Interpretation of Evidence (E), and Presentation and Perception Formation (P) for ID cases. Further, we also elucidate specific complexity contributory factors relevant to each diagnostic subtask (Table 8).

6.1. DICE model as a causality determination for DEER analysis

A major implication of the tool is its ability to define the causality behind diagnostic errors. Schiff et al. (2009) developed a taxonomy of diagnostic errors (DEER) in clinical settings that elucidates the modality of errors and the stage of their occurrence in a diagnostic interaction. Combining our model with the DEER taxonomy framework offers a critical opportunity to correlate the causality and modality of diagnostic errors in a clinical workflow. This can form the foundation for predictive informatics focusing on anticipating potential future diagnostic errors based on current infrastructure and workflow of healthcare settings. Leveraging these insights can help design tailored behavioural interventions that aid diagnostic decision-making. Future research should focus on developing a broader framework that operationalises this combination.

Table 8. Most relevant complexity contributory factors of diagnostic subtasks

Diagnostic Subtask	Most Relevant CCFs Identified
Presentation and Perception Formation (P)	Input Diversity, Input Quantity, Quantity of Paths, Quantity of Action Steps, Input Clarity, Goal/ Output Conflict or Redundancy, Process Cognitive Requirement
Hypothesis Generation (H)	Goal/ Output Quantity, Input Quantity, Quantity of Paths, Input Diversity, Goal/ Output Clarity, Process Clarity, Input Conflict/ Redundancy, Inaccuracy of Inputs, Input Mismatch, Process Cognitive Requirements
Gathering and Interpretation of Evidence (E)	Quantity of Paths, Input Diversity, Goal/ Output Conflict or Redundancy, Goal/ Output Change, Rate of Input Change, Non-routine events as Inputs, Time Pressure
Verification (V)	Quantity of Paths, Input Diversity, Goal/ Output Conflict or Redundancy, Rate of Input Change, Non-routine events as Inputs, Process Repetitiveness, Presentation Heterogeneity, Concurrency and Time Pressure

7. Conclusions and limitations

We operationalised a complexity estimation tool and elicited factors contributing to diagnostic complexity for infectious diseases in an EM setting while aiming to establish a foundation for employing the tool in diverse clinical settings and other disease spectrums. A strong correlation between the measured objective complexity and the difficulty perceived by attending clinicians emphasises the validity of our findings. However, the study is bound by certain limitations. The small data set of 10 cases provides necessary insights but limits the robustness of our findings for ID cases. Future research should focus on expanding the dataset to a larger number to refine and further adjust our findings. Second, considering this was a single-centre study, our findings may not generalise to other EM settings. While the disease-specific complexity might remain constant, the contextually imposed complexity might vary according to the organisational structure and cultural influences within a clinical setting. Therefore, future studies need to be conducted across multiple emergency settings to account for these variations and optimise the generalisability of our findings.

Nevertheless, our work provides a basis for designing behavioural interventions aiming to reduce the complexity of diagnostic processes. Such interventions can help enhance the efficiency of clinical diagnosis, improving patient outcomes and well-being.

Reference

Bay Brix Nielsen, C. K. E., Daalhuizen, J., & Cash, P. J. (2021). Defining the behavioural design space. *International Journal of Design*, 15(1), 1–16.

Bornstein, B. H., & Emler, A. C. (2001). Rationality in medical decision making: A review of the literature on doctors' decision-making biases. *Journal of Evaluation in Clinical Practice*, 7 (2), 97–107. https://doi.org/10.1046/j.1365-2753.2001.00284.x

Calder, L. A., Arnason, T., Vaillancourt, C., Perry, J. J., Stiell, I. G., & Forster, A. J. (2015). How do emergency physicians make discharge decisions? *Emergency Medicine Journal*, 32 (1), 9–14. https://doi.org/10.1136/emermed-2013-202421

Cowdell, F., & Dyson, J. (2019). How is the theoretical domains framework applied to developing health behaviour interventions? A systematic search and narrative synthesis. *BMC Public Health*. https://doi.org/10.1186/s12889-019-7442-5

Croskerry, P. (2009a). A Universal Model of Diagnostic Reasoning. In *Academic Medicine* (Vol. 84, Issue 8, pp. 1022–1028). Lippincott Williams and Wilkins. https://doi.org/10.1097/acm.0b013e3181ace703

Croskerry, P. (2009b). Clinical cognition and diagnostic error: Applications of a dual process model of reasoning. *Advances in Health Sciences Education*, 14 (1 SUPPL), 27–35. https://doi.org/10.1007/s10459-009-9182-2

Croskerry, P., & Clancy, M. (2025). Diagnosis through prisms: Unraveling its complexity. In *Academic Emergency Medicine*. John Wiley and Sons Inc. https://doi.org/10.1111/acem.15120

Dey, S., Dabral, S., & Khadilkar, P. (2023). BEHAVIOURAL DESIGN FOR MEDICAL ERRORS DURING PATIENT DIAGNOSIS PROCESS. *Proceedings of the Design Society*, 3. https://doi.org/10.1017/pds.2023.77

Islam, R., Weir, C., & Fiol, G. Del. (2016). Clinical complexity in medicine: A measurement model of task and patient complexity. *Methods of Information in Medicine*, 55 (1), 14–22. https://doi.org/10.3414/ME15-01-0031

Kassirer, J. P., & Kopelman, R. I. (1989). Cognitive errors in diagnosis: Instantiation, classification, and consequences. *The American Journal of Medicine*, 86 (4), 433–441. https://doi.org/10.1016/0002-9343(89)90342-2

Khadilkar, P. R., & Cash, P. (2020). Understanding behavioural design: barriers and enablers. *Journal of Engineering Design*, 31 (10), 508–529. https://doi.org/10.1080/09544828.2020.1836611

Liu, P., & Li, Z. (2012). Task complexity: A review and conceptualization framework. *International Journal of Industrial Ergonomics*, 42 (6), 553–568. https://doi.org/10.1016/J.ERGON.2012.09.001

Newman-Toker, D. E., Schaffer, A. C., Yu-Moe, C. W., Nassery, N., Saber Tehrani, A. S., Clemens, G. D., Wang, Z., Zhu, Y., Fanai, M., & Siegal, D. (2019). Serious misdiagnosis-related harms in malpractice claims: The "big Three"-vascular events, infections, and cancers. *Diagnosis*, 6 (3), 227–240. https://doi.org/10.1515/dx-2019-0019

Newman-Toker, D. E., Wang, Z., Zhu, Y., Nassery, N., Saber Tehrani, A. S., Schaffer, A. C., Yu-Moe, C. W., Clemens, G. D., Fanai, M., & Siegal, D. (2021). Rate of diagnostic errors and serious misdiagnosis-related harms for major vascular events, infections, and cancers: Toward a national incidence estimate using the "big Three." *Diagnosis*, 8 (1), 67–84. https://doi.org/10.1515/dx-2019-0104

Newman-Toker, D., Peterson, S. M., Badihian, S., Hassoon, A., Nassery, N., Parizadeh, D., WIlson, L. M., Jia, Y., Omron, R., Tarmarajah, S., Guerin, L., Bastani, P. B., Fracica, E. A., Kotwal, S., & Robinson, K. A. (2023). *Diagnostic Errors in the Emergency Department: A Systematic Review Comparative Effectiveness Review*. www.ahrq.gov

Reason, J. (1990). *Human Error*. Cambridge University Press.

Roosan, D., Weir, C., Samore, M., Jones, M., Rahman, M., Stoddard, G. J., & Del Fiol, G. (2017). Identifying complexity in infectious diseases inpatient settings: An observation study. *Journal of Biomedical Informatics*, 71, S13–S21. https://doi.org/10.1016/j.jbi.2016.10.018

Sahu, A., Bhoi, S., Galwankar, S., Aggarwal, P., Murmu, L., Nayer, J., Sinha, T., Mishra, P., Ekka, M., & Kumar, A. (2020). All india institute of medical sciences triage protocol (ATP): ATP of a busy emergency department. In *Journal of Emergencies, Trauma and Shock* (Vol. 13, Issue 2, pp. 107–109). Wolters Kluwer Medknow Publications. https://doi.org/10.4103/JETS.JETS_137_19

Wears, R. L. (2009). What makes diagnosis hard? *Advances in Health Sciences Education*, 14 (1 SUPPL), 19–25. https://doi.org/10.1007/s10459-009-9181-3

Zwaan, L., & Singh, H. (2015). The challenges in defining and measuring diagnostic error. *Diagnosis*, 2 (2), 97–103. https://doi.org/10.1515/dx-2014-0069