HIGIVARA
Journal of Health Sciences and Policy
Higivara: Journal of Health Sciences and Policy
Vol. 1, No. 1, 2026
E-ISSN: XXXX-XXXX
punggawajournal.com/higivara

Diagnostic Delay in the Emergency Department: Impact on Clinical Outcomes and Patient Prognosis

Waode Fauziah Feby Alfitriani Ary Priady[1]
[1]Hainan Medical University, Haikou, China.
Correspondence: febyaryannas09@gmail.com

Article Info

Article History:
Received July 15th, 2026
Revised July 25th, 2026
Accepted July 29th, 2026

Keyword:
Diagnostic delay; Emergency department; Diagnostic error; Clinical outcomes; Patient safety.

ABSTRACT

Diagnostic delay in emergency care is difficult to judge because a disease that appears obvious after deterioration may have been clinically indistinct during the first encounter. This scoping review with critical narrative synthesis examined where delay becomes visible in the emergency diagnostic process, which mechanisms recur across different conditions, and when later diagnosis is associated with clinically important harm. Across appendicitis, sepsis, vascular and neurological emergencies, tuberculosis, and cross-condition studies, delay was seldom attributable to one failed decision. It more often developed when an early presentation remained nonspecific, a plausible working diagnosis narrowed further assessment, or residual uncertainty was not carried through discharge, pending results, and repeat attendance. Outcomes were most consequential when disease progression was time dependent and effective treatment had been available during the missed interval. The synthesis develops diagnostic revisability as a clinical safety concept. A provisional diagnosis remains revisable when uncertainty is documented, new information can reconnect with the diagnostic process, and return presentations alter rather than reset clinical reasoning. This approach does not demand certainty at first contact. It asks whether emergency care retains a practical route for changing the diagnosis before biological progression makes the correction unavoidable.

© 2026 The Authors. Published by Punggawa Legacy Center. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/)

INTRODUCTION

Serious disease may be clinically indistinct during its early presentation. Diffuse abdominal pain can precede appendicitis, dizziness may be the first expression of posterior circulation stroke, and infection can look minor before organ dysfunction becomes apparent. Emergency clinicians make disposition decisions while that uncertainty is still present. They have to decide whether a patient can leave, needs observation, or requires immediate treatment, often before the illness has acquired its later and more recognisable form. The diagnosis that looks inevitable after deterioration may therefore have been only one possibility among several at the index visit. Hindsight creates a familiar problem for diagnostic-safety research because later knowledge can make earlier ambiguity disappear. Yet uncertainty at the first encounter does not mean that every subsequent delay was unavoidable. The clinically important question is whether there was a reasonable opportunity to recognise risk, gather additional information, or keep the diagnostic question open before harm became harder to prevent. Delay is not simply a wrong diagnostic label held for too long.

The process can fail before a disease enters the differential diagnosis, after it has been considered but inadequately tested, or after discharge when abnormal results and worsening symptoms no longer reconnect with the original assessment. Studies of harmful diagnostic error have therefore moved away from treating discrepancy alone as proof of error. Record review, symptom-disease pair analysis, and electronic triggers ask whether an earlier encounter contained a missed diagnostic opportunity rather than whether the final diagnosis differed from the first one (Gunderson et al., 2020; Liberman et al., 2023). A 2026 multidisciplinary consensus in emergency medicine sharpened that distinction by defining missed diagnostic opportunity around the presence of a reasonable chance to

improve the diagnostic process, not merely the fact that a diagnosis was reached later (Berdahl et al., 2026).

That distinction is essential because natural disease evolution and diagnostic failure can produce the same chronology. Emergency departments make the problem unusually visible. They receive undifferentiated illness, work with incomplete prior information, and often have only one encounter in which to decide whether risk is low enough for discharge. A symptom diagnosis may be entirely appropriate when dangerous causes have been assessed and the plan accounts for what remains uncertain. The same diagnosis becomes unsafe when it is heard as closure, when a pending result has no clear owner, or when a return visit is treated as a new episode rather than additional evidence. Research on diagnostic excellence in emergency care has accordingly shifted attention from the impossible goal of eliminating every miss toward the conditions that make diagnostic performance more reliable (Mahajan et al., 2025).

Excessive testing is not a neutral solution. It can expose patients to radiation, incidental findings, unnecessary admission, treatment cascades, and longer waits. Safety depends on recognising where additional certainty is worth pursuing and where uncertainty itself needs active management. Recent disease-specific studies show a recurring structure despite different biology. Delayed paediatric appendicitis is associated with less typical early findings and, when diagnosis is later established, more perforation and complex care (Mahajan et al., 2020; Michelson et al., 2021, 2023). Sepsis can follow an earlier treat-and-release encounter for nonspecific symptoms, while delayed recognition postpones treatment in a condition where timing can matter (Cifra et al., 2021; Horberg et al., 2021; Husabø et al., 2020). Stroke and cervicocephalic artery dissection are vulnerable to benign explanations when dizziness, headache, or pain dominates the presentation (Jones et al., 2021; Liberman et al., 2020). Aortic dissection presents another version of the same difficulty because many patients do not arrive with a complete textbook pattern (Lovatt et al., 2022).

These conditions cannot be pooled as one outcome, but they expose a common problem. A provisional explanation becomes sufficiently convincing before the clinical course has supplied enough information to justify finality. The review therefore examines diagnostic delay as a clinical trajectory rather than as a catalogue of missed diseases. It asks where delay becomes visible during emergency care, which mechanisms recur across otherwise different conditions, and what can reasonably be said about later clinical outcomes. The analysis is deliberately narrower than malpractice or preventable death.

A late diagnosis counts as clinically important here when evidence supports a missed opportunity within the emergency-care pathway and when earlier recognition could plausibly have changed investigation, monitoring, treatment, or disposition. This framing permits attention to error without assuming fault and makes room for a second problem that is often lost in retrospective review: some uncertainty was legitimate at the first visit, so safer care must preserve a route for revising the diagnosis rather than merely demand that clinicians name the correct disease sooner.

RESEARCH METHODS

A scoping review with critical narrative synthesis was used to examine how diagnostic delay is identified in emergency care and how it relates to clinical outcomes. The approach was selected because the literature applies varied research designs and definitions of diagnostic delay that are not suitable for statistical pooling. The review followed updated JBI guidance for scoping reviews and used PRISMA-ScR principles to support transparent reporting (Peters et al., 2020). Searches were conducted primarily in PubMed/MEDLINE and supplemented by citation tracking and targeted searches. Eligible studies were English-language peer-reviewed publications from 2020 to 2026 that examined diagnostic delay, missed diagnostic opportunities, subsequent diagnosis, or clinically relevant consequences within an emergency-care pathway.

Relevant information was extracted on the initial clinical presentation, diagnostic assessment, missed or delayed diagnostic opportunity, subsequent diagnosis or treatment, and reported outcomes. Studies without a clear emergency-care component, pure diagnostic-accuracy studies without a delay pathway, case reports, and opinion pieces were excluded. Findings were synthesized through repeated comparison to identify recurring points at which diagnostic processes became restricted or failed to remain responsive to new information. Particular attention was given to limitations in recognition, interpretation, and continuity of diagnostic care. Because the review aimed to map patterns and

mechanisms rather than estimate pooled effects, no formal meta-analysis or uniform risk-of-bias scoring was performed; instead, methodological limitations and the strength of diagnostic-error attribution were considered during interpretation.

RESULTS AND DISCUSSION

1. Early Diagnostic Ambiguity and Constraints on Clinical Recognition

Children with delayed appendicitis provide a clear example of limited early discernibility. At the first encounter, children later diagnosed with appendicitis often had fewer classic findings and lower estimated diagnostic probabilities than those diagnosed promptly. On return, the picture was easier to recognise, but perforation and longer admission were more common (Michelson et al., 2021). Claims-based analysis reached a related conclusion. Potentially missed appendicitis was more likely when abdominal pain was absent, constipation was present, or comorbidity created competing explanations, and women and girls had higher odds of a potentially missed diagnosis (Mahajan et al., 2020). The pattern does not support a simple story of clinicians ignoring obvious signs. Some patients reached emergency care before the disease had become easy to distinguish from common alternatives. The missed opportunity, where one existed, often concerned what happened to uncertainty during disposition rather than failure to recognise a fully developed syndrome. Clinical exposure also appears to matter.

Emergency departments seeing more children had lower rates of delayed paediatric appendicitis, and delayed cases experienced more perforation, intensive care, abscess drainage, repeated operations, and sepsis (Michelson et al., 2023). The association cannot establish that greater volume directly improves reasoning. High-volume departments may have different imaging access, paediatric expertise, observation practices, referral patterns, or coding. Even so, it weakens an exclusively individual account of diagnostic delay. Recognition is partly supported by repeated exposure to a patient population and by an organisation's ability to hold an uncertain case long enough for the trajectory to become clearer. Sepsis shows why a nonspecific early presentation cannot be equated with a harmless one.

Children later hospitalised with severe sepsis sometimes had a preceding emergency visit plausibly related to the same illness (Cifra et al., 2021). Among adults, earlier treat-and-release encounters for altered mental status or fluid and electrolyte disorders preceded some sepsis hospitalisations, suggesting that a subset of patients had warning presentations before definitive recognition (Horberg et al., 2021). These findings do not identify a universal point at which sepsis should have been diagnosed. They do show that the diagnostic process needs to respond to accumulation. Observational evidence links earlier recognition with shorter treatment intervals, while the relation between timing and mortality remains vulnerable to confounding by severity and indication (Husabø et al., 2020). Neurological and vascular emergencies are frequently missed when the dominant symptom has a common benign explanation. Posterior circulation strokes and milder deficits are overrepresented among strokes not recognised during initial emergency assessment (Jones et al., 2021).

Cervicocephalic artery dissection may first be treated as headache or neck pain before focal neurological findings appear (Liberman et al., 2020). Systematic review evidence on vulnerable populations suggests that presentation, communication, and clinician interpretation can also interact with age, sex, race, language, disability, and comorbidity, although the direction of these associations is not uniform across studies (Herasevich et al., 2023). Atypical presentation is therefore not simply an unusual patient feature. It becomes clinically important when the available model of the disease is narrower than the way serious illness actually appears in emergency care.

Aortic dissection illustrates the same problem with a low-frequency, high-harm diagnosis. The systematic review by Lovatt et al. (2022) found marked variation in presentation and recurrent initial misdiagnoses. No single symptom or early test safely removes the diagnosis across all patients. The practical implication is limited but important. Vigilance needs to be attached to risk features, discordant findings, and clinical trajectory. Asking clinicians to remember every rare diagnosis in every encounter does not create a workable safety strategy. The more defensible response to weak early discernibility is to identify which unresolved features justify observation, repeat examination, additional testing, or explicit return precautions.

2. Diagnostic Closure, Reassessment, and Longitudinal Continuity

Diagnostic delay becomes more preventable when enough information is available to question the first explanation but the interpretation does not reopen. Retrospective work on diagnostic discrepancies in emergency departments found failures across history, examination, test selection, synthesis, and follow-up rather than one repeated cognitive mistake (Schols et al., 2024). A scoping review of cognitive bias reached a similar conclusion. Anchoring, confirmation bias, availability, and premature closure are useful descriptions, yet their effects depend on workload, experience, team interaction, and the information environment (Jala et al., 2023). Labelling a bias after the event can therefore be less informative than asking what evidence contradicted the working diagnosis, whether that contradiction was visible, and what made revision unlikely at that point. Acute myocardial infarction demonstrates how population-level detection can reveal patterns without proving error in every case. SPADE analysis links a later serious diagnosis to an earlier treat-and-release visit for a related symptom and compares observed with expected associations. Applied to acute myocardial infarction, the method can identify organisations or symptom pathways that warrant review (Sharp et al., 2021).

It cannot establish from claims alone that the first clinician had sufficient evidence for definitive diagnosis. Its value is as a screen. When a pattern is unusually frequent, case review can examine whether electrocardiograms were repeated, biomarker pathways were completed, prior records were available, or a non-cardiac explanation became final before discordant information had been resolved. Negative tests create a different form of closure. A result may be technically correct and still fail to answer the clinical question because testing occurred too early, the chosen test did not match the risk, or qualified findings were simplified during disposition. Sudden severe headache is a useful example. Strategies for subarachnoid haemorrhage depend on clinical assessment, the timing and quality of computed tomography, and selective use of further testing.

A systematic review found that diagnostic performance varies with pathway and context, so no single negative result functions as an unconditional exclusion across all presentations (Walton et al., 2022). The safety problem arises when uncertainty that remains after testing is no longer visible in the plan. Electronic-trigger review supports the view that breakdown often occurs through incomplete integration rather than complete absence of data. Validation studies have used later stroke, unplanned return, or escalation to identify records for focused review (Mahajan et al., 2021; Vaghani et al., 2021). In implementation across a national health system, different eTriggers identified different types of missed opportunities and showed widely varying positive predictive values (Vaghani et al., 2025). The signal is useful precisely because it is imperfect. A return with admission or an unexpected transfer to intensive care enriches the pool for review, but many such cases still represent appropriate initial care followed by disease progression.

Triggering therefore shifts the question from "was there an error" to "is this case worth clinical review. The newest work extends that screening problem to large language models. Marks et al. (2026) tested six commercial models in two emergency-care eTrigger cohorts and found that model performance varied substantially by cohort and operating threshold. None removed the need for physician adjudication. The finding matters less as an argument for artificial intelligence than as a reminder that diagnostic surveillance itself contains judgement. Algorithms decide which records receive attention, and model thresholds alter the balance between review burden and missed cases.

Automated detection may make longitudinal quality review more feasible, but a technically efficient screen can still misclassify natural disease evolution as error or miss harm that falls outside the chosen trigger. Diagnostic uncertainty is also carried through language. Resident physicians have been observed to communicate uncertainty at discharge indirectly or incompletely, leaving patients with a more definitive impression than the clinician intended (Doty et al., 2022). Simulation-based mastery learning improved clinicians' ability to explain uncertainty and safety-net patients without reciting every remote possibility (Rising et al., 2023). These studies locate part of diagnostic safety after the immediate reasoning task. A provisional diagnosis can only remain provisional if the patient understands what has not been established and knows what change should prompt reassessment. Generic advice to return if "worse" may be inadequate when the clinically important change is new weakness, persistent fever, recurrent syncope, or a different pattern of pain. Continuity problems become more visible when patients return repeatedly. Pulmonary tuberculosis is an unusually clear example because delayed recognition affects both the patient and people exposed in crowded care settings.

Repeated emergency-department use before definitive tuberculosis diagnosis has been associated with longer health-system delay, tuberculosis-related mortality, exposure time, and secondary cases (Heffernan et al., 2021). The first respiratory presentation may have been nonspecific. By the third or fourth encounter, however, previous imaging, unresolved symptoms, epidemiological risk, and failure of the initial explanation acquire new weight. A diagnostic process that treats each attendance as independent loses precisely the longitudinal information that makes a serious alternative more likely. Unexpected return with intensive-care admission has therefore been used as a practical trigger for possible diagnostic error (Aaronson et al., 2020). It is not a quality metric on its own. Some patients deteriorate despite appropriate assessment, some seek care elsewhere, and others suffer diagnostic harm without returning.

Its usefulness lies in concentrating review on a small group in which the cost of a missed opportunity may be high. The same caution applies to broader trigger portfolios. Case-finding methods are valuable when they bring clinicians back to the original encounter with a reason to look again. They become misleading when trigger positivity is treated as a verdict. Checklists are a less reliable answer to premature closure than their simplicity suggests. A systematic review using a human-factors framework found that diagnostic checklists can support cognitive work, but effects depend on how and when they are used (Al-Khafaji et al., 2022).

A generic list added to every encounter risks becoming another task completed under time pressure. The reviewed evidence more strongly supports selective interruption when the working diagnosis leaves important findings unexplained, when a high-harm condition remains plausible, or when the patient has returned unexpectedly. The purpose is modest. It is to create a point at which the first explanation can still be revised.

3. Clinical Consequences and Time-Sensitive Harm Associated with Diagnostic Delay

Clinical consequences differ because diseases do not progress at the same speed and earlier treatment is not equally effective in every condition. Appendicitis offers the most concrete sequence in the reviewed literature. Children diagnosed after an earlier missed opportunity had more perforation and longer hospital stays, and multicentre data linked delay with intensive care and additional invasive treatment (Michelson et al., 2021, 2023). These associations do not mean that every perforation could have been prevented at the first visit. They show that the diagnostic interval can coincide with meaningful biological progression. The later diagnosis may become easier because the disease has become more severe. Sepsis has a less discrete diagnostic threshold. Infection, host response, and organ dysfunction develop along a continuum, while early symptoms overlap with common self-limited illnesses. Delayed recognition can nevertheless postpone antimicrobials, monitoring, source control, and escalation.

Antecedent emergency encounters before paediatric and adult sepsis hospitalisation show that some patients pass through a period in which the final diagnosis is not yet secure but risk is already accumulating (Cifra et al., 2021; Horberg et al., 2021). Causal interpretation remains difficult because rapidly progressive illness can both be harder to recognise early and produce worse outcomes despite appropriate care. That limitation should temper claims about preventable mortality. Vascular emergencies convert time into loss of viable tissue or sudden physiological collapse. Missed acute myocardial infarction can delay reperfusion and antithrombotic treatment, while stroke can close a treatment window and leave persistent disability.

Cervicocephalic dissection may precede cerebral ischaemia, making an earlier pain or headache presentation important even when focal neurological findings were absent (Liberman et al., 2020). Aortic dissection can progress to rupture, tamponade, malperfusion, or death before diagnostic certainty is reached (Lovatt et al., 2022). Outcome severity is one reason overall diagnostic accuracy is a weak safety metric. A rare false negative may have far greater clinical consequence than many minor false positives, yet indiscriminate testing can also cause harm. The burden of delay is not distributed evenly. Potentially missed appendicitis has been associated with sex and comorbidity, and systematic review evidence identifies diagnostic-error concerns among patients with cardiovascular or neurological symptoms who are older, disabled, from racial or linguistic minorities, or otherwise vulnerable (Mahajan et al., 2020; Herasevich et al., 2023).

The evidence is heterogeneous and should not be reduced to a single disparity estimate. It does suggest that complexity can obscure a new signal. Multiple chronic conditions generate more plausible

explanations, prior labels can anchor interpretation, and communication barriers may limit how strongly a symptom changes the working diagnosis. Mortality attracts the strongest language and requires the greatest restraint. Death after a delayed diagnosis does not establish that delay caused the death.

Population estimates of serious misdiagnosis-related harm help show scale, but they combine diseases and infer preventability using models that are different from adjudication of an individual case (Newman-Toker et al., 2024). Hospital record-review studies likewise show that estimates depend heavily on case-finding method and reviewer judgement (Gunderson et al., 2020). A stronger causal claim needs more than chronology. There should be a plausible missed opportunity, a clinically meaningful interval, and a reason to believe earlier recognition would have changed management while benefit was still available. Without those elements, "too late" becomes rhetoric rather than a clinical conclusion.

4. Diagnostic Revisability as a Framework for Emergency Diagnostic Safety

Across the reviewed literature, diagnostic delay is better understood as loss of revisability than as one wrong decision. The sequence often begins with legitimate uncertainty. A low-specificity presentation supports several explanations, and one becomes the working diagnosis because it fits most of the available information. Harm becomes more likely when later evidence cannot effectively challenge that first model. A qualified test is heard as negative, a result returns after discharge without a clear owner, or a patient comes back and the earlier assessment is repeated rather than updated. This sequence explains why the same diagnostic label can be reasonable at one point and unsafe later. It also avoids the retrospective fiction that a serious diagnosis was always obvious simply because it eventually became obvious. Diagnostic revisability can be stated more precisely. It is the practical capacity of a provisional diagnosis to change when new clinical information alters risk. That capacity depends on more than an individual clinician remaining open-minded. Residual uncertainty has to be visible in documentation and communication. New results need to reach someone who can act on them.

A repeat presentation has to be linked with the earlier encounter. Observation and reassessment must be available when the immediate probability of disease is low but the consequence of a miss is high. In this sense, revisability is not another name for defensive testing. It is a way of preserving diagnostic options until the available evidence justifies closure. The diagnosis may remain provisional even when discharge itself is appropriate. The concept also clarifies the relationship between cognitive and system explanations. Premature closure occurs through clinical judgement, but the cost of reopening a diagnosis is shaped by the environment.

Crowding shortens the time available for reassessment. Handovers fragment ownership. Separate records make prior uncertainty difficult to see. A pending result that arrives after discharge may be clinically meaningful but operationally homeless. Conversely, describing every failure as systemic can obscure the fact that clinicians still interpret evidence and decide when discordant findings deserve more weight. Diagnostic safety needs both levels because system conditions influence whether clinical judgement can be revised, while judgement determines whether the opportunity to revise is used. Reopening the case then becomes an operational task as well as a cognitive one. Emergency departments already use time diagnostically through observation, serial biomarkers, repeat electrocardiograms, and repeated examination. The relevant challenge is selective use. Observation capacity is finite, and crowding creates pressure toward disposition. A universal instruction to "consider more diagnoses" is therefore impractical.

The reviewed studies point instead to situations where a brief diagnostic pause is more defensible: the patient appears sicker than the working diagnosis suggests, an important finding remains unexplained, a high-harm alternative is still plausible, or the patient has returned unexpectedly. These are not validated universal triggers. They are recurring points where the cost of premature finality appears greater than the cost of reopening the assessment. That distinction matters most when observation resources are already constrained. Safety-netting belongs inside this diagnostic process, not after it.

Communication studies show that patients may leave with a more certain interpretation than clinicians intended (Doty et al., 2022), while structured training can improve explanation of uncertainty and return precautions (Rising et al., 2023). The clinical implication is simple but often neglected. A working diagnosis that remains uncertain should be communicated as uncertain, and the return plan should identify meaningful changes rather than rely on a generic instruction to come back if worse. The

plan also has to be feasible. Advice has limited protective value when follow-up is inaccessible or when another emergency department cannot see the unresolved questions from the first visit. Specificity in the return plan reduces the chance that uncertainty is mistaken for reassurance. Retrospective detection is equally important because diagnostic delays are less likely than medication or procedural errors to be voluntarily reported. The clinician who discharged the patient may never learn the final diagnosis. Electronic triggers address that visibility problem by using later events to select records for review.

Their limitations are not incidental. Trigger definitions determine who becomes visible, positive predictive value can be low, and patients who leave the health system may disappear from the data. The 2026 LLM study by Marks et al. adds another layer. Automated models may reduce screening burden, but differing operating thresholds mean that technology can redistribute which cases receive human attention without resolving the underlying adjudication problem. Human review remains necessary because the reference standard is itself a clinical judgement. A fair review process must therefore resist hindsight. The eventual diagnosis should not erase the uncertainty of the index encounter. Reviewers need to reconstruct what information was available, what alternatives were clinically reasonable, what evidence emerged later, and whether the disposition matched the residual risk at the time.

This is close to the logic of missed diagnostic opportunity adopted in recent emergency-medicine consensus work (Berdahl et al., 2026). It separates quality improvement from an assumption of negligence. Improvement can be justified when a reasonable opportunity was missed even if legal fault is absent, while a diagnosis that could not reasonably have been made earlier should not be converted into error by outcome knowledge alone. The outcome literature supports the same restraint. Associations between delayed diagnosis and perforation, intensive care, prolonged admission, disability, or death are strongest when biological progression is time sensitive and earlier treatment has a plausible benefit.

The evidence is weaker when the interval is short, treatment would not have changed, or the first presentation genuinely lacked discriminating information. This matters for priorities. Diagnostic-safety programs do not need to review every amended diagnosis with equal intensity. Cases involving a serious later diagnosis, unexpected escalation, repeat presentation, or a clinically important unacknowledged result provide a more defensible starting point for learning. Important evidence gaps remain. Most recent studies come from well-resourced systems with longitudinal electronic records. Patients who seek subsequent care elsewhere are undercounted in many trigger methods. Administrative studies can map symptom-disease patterns but cannot fully reconstruct reasoning, while record review is vulnerable to reviewer knowledge of the outcome. Patient experience is also thinly represented. Distress, loss of trust, disability, and the practical burden of repeat care are less consistently measured than admission or mortality.

CONCLUSION

Diagnostic delay in emergency care is not synonymous with diagnostic error. Early disease can be genuinely difficult to distinguish from benign illness, and a later diagnosis may reflect biological evolution rather than a missed opportunity. The reviewed evidence becomes more concerning when clinically meaningful uncertainty was present but could no longer influence what happened next. A provisional diagnosis became functionally final, a return visit did not alter the prior frame, or new information failed to reconnect with the patient. Diagnostic revisability captures that problem without demanding certainty at first contact.

It requires a realistic route for changing the working diagnosis while earlier treatment or monitoring can still matter. Observation, reassessment, explicit uncertainty, result ownership, and longitudinal review are useful only insofar as they preserve that route. For emergency medicine, the more credible safety aim is therefore not zero diagnostic delay. It is earlier recognition of missed opportunities in the cases where time changes what can still be done for the patient during emergency care.

REFERENCES

Aaronson, E., Jansson, P., Wittbold, K., Flavin, S., & Borczuk, P. (2020). Unscheduled return visits to the emergency department with ICU admission: A trigger tool for diagnostic error. American Journal of Emergency Medicine, 38(8), 1584-1587. https://doi.org/10.1016/j.ajem.2019.158430

Al-Khafaji, J., Townsend, R. F., Townsend, W., Chopra, V., & Gupta, A. (2022). Checklists to reduce diagnostic error: A systematic review of the literature using a human factors framework. BMJ Open, 12(4), e058219. https://doi.org/10.1136/bmjopen-2021-058219

Berdahl, C. T., Schiff, G. D., Venkatesh, A. K., Seferian, E. G., Torbati, S. S., Ptacek, M. P., Carrascoza-Bolanos, J., Qureshi, N., Nuckols, T. K., & Multidisciplinary ED Diagnostic Excellence Panel. (2026). Defining diagnostic excellence and missed diagnostic opportunity for the emergency department setting. Annals of Emergency Medicine. Advance online publication. https://doi.org/10.1016/j.annemergmed.2026.04.004

Cifra, C. L., Westlund, E., Ten Eyck, P., Ward, M. M., & Michelson, K. A. (2023). Delayed diagnosis of serious paediatric conditions in 13 regional emergency departments. BMJ Quality & Safety, 32, 603-612. https://doi.org/10.1136/bmjqs-2022-015314

Cifra, C. L., Westlund, E., Ten Eyck, P., Ward, M. M., Mohr, N. M., & Katz, D. A. (2021). An estimate of missed pediatric sepsis in the emergency department. Diagnosis, 8(2), 193-198. https://doi.org/10.1515/dx-2020-0023

Doty, A. M., Rising, K. L., Hsiao, C. J., Barg, F. K., Moyer, C. A., & Santen, S. A. (2022). 'Unfortunately, I don't have an answer for you': How resident physicians communicate diagnostic uncertainty to patients during emergency department discharge. Patient Education and Counseling, 105(7), 2053-2057. https://doi.org/10.1016/j.pec.2021.12.002

Gunderson, C. G., Bilan, V. P., Holleck, J. L., Nickerson, P., Cherry, B. M., Chui, P., Bastian, L. A., & Grimshaw, A. A. (2020). Prevalence of harmful diagnostic errors in hospitalised adults: A systematic review and meta-analysis. BMJ Quality & Safety, 29(12), 1008-1018. https://doi.org/10.1136/bmjqs-2019-010822

Heffernan, C., Paulsen, C., Asadi, L., Egedahl, M. L., Rowe, B. H., Barrie, J., & Long, R. (2021). Individual and public health consequences associated with a missed diagnosis of pulmonary tuberculosis in the emergency department: A retrospective cohort study. PLOS ONE, 16(3), e0248493. https://doi.org/10.1371/journal.pone.0248493

Herasevich, S., Soleimani, J., Huang, C., Pinevich, Y., Dong, Y., Pickering, B. W., & Herasevich, V. (2023). Diagnostic error among vulnerable populations presenting to the emergency department with cardiovascular and cerebrovascular or neurological symptoms: A systematic review. BMJ Quality & Safety, 32(11), 676-688. https://doi.org/10.1136/bmjqs-2022-015038

Horberg, M. A., Nassery, N., Rubenstein, K. B., Certa, J. M., Shamim, E. A., Rothman, R. R., Wang, Z., Hassoon, A., Townsend, J., Galiatsatos, P., Pitts, S., & Newman-Toker, D. E. (2021). Rate of sepsis hospitalizations after misdiagnosis in adult emergency department patients: A look-forward analysis using SPADE methodology. Diagnosis, 8(4), 479-488. https://doi.org/10.1515/dx-2020-0145

Husabø, G., Nilsen, R. M., Flaatten, H., Solligård, E., Frich, J. C., Bondevik, G. T., Braut, G. S., Walshe, K., Harthug, S., & Hovlid, E. (2020). Early diagnosis of sepsis in emergency departments, time to treatment, and association with mortality: An observational study. PLOS ONE, 15(1), e0227652. https://doi.org/10.1371/journal.pone.0227652

Jala, S., Fry, M., & Elliott, R. (2023). Cognitive bias during clinical decision-making and its influence on patient outcomes in the emergency department: A scoping review. Journal of Clinical Nursing, 32(19-20), 7076-7085. https://doi.org/10.1111/jocn.16845

Jones, S. P., Bray, J. E., Gibson, J. M. E., McClelland, G., Miller, C., Price, C. I., & Watkins, C. L. (2021). Characteristics of patients who had a stroke not initially identified during emergency prehospital assessment: A systematic review. Emergency Medicine Journal, 38(5), 387-393. https://doi.org/10.1136/emermed-2020-209607

Liberman, A. L., Navi, B. B., Esenwa, C. C., Lipton, R. B., & Kamel, H. (2020). Misdiagnosis of cervicocephalic artery dissection in the emergency department. Stroke, 51(6), 1876-1878. https://doi.org/10.1161/STROKEAHA.120.029390

Liberman, A. L., Wang, Z., Zhu, Y., Hassoon, A., Choi, J., Austin, J. M., Johansen, M. C., & Newman-Toker, D. E. (2023). Optimizing measurement of misdiagnosis-related harms using symptom-disease pair analysis of diagnostic error: Comparison groups to maximize SPADE validity. Diagnosis, 10(3), 225-234. https://doi.org/10.1515/dx-2022-0130

Lovatt, S., Wong, C. W., Schwarz, K., Borovac, J. A., Lo, T., Gunning, M., Di Salvo, G., & Kotecha, D. (2022). Misdiagnosis of aortic dissection: A systematic review of the literature. American Journal of Emergency Medicine, 53, 16-22. https://doi.org/10.1016/j.ajem.2021.11.047

Mahajan, P., Basu, T., Pai, C.-W., Singh, H., Petersen, N., Bellolio, M. F., Gadepalli, S. K., & Kamdar, N. S. (2020). Factors associated with potentially missed diagnosis of appendicitis in the emergency department. JAMA Network Open, 3(3), e200612. https://doi.org/10.1001/jamanetworkopen.2020.0612

Mahajan, P., Pai, C.-W., Cosby, K. S., Mollen, C. J., Shaw, K. N., Chamberlain, J. M., & Singh, H. (2021). Identifying trigger concepts to screen emergency department visits for diagnostic errors. Diagnosis, 8(3), 340-346. https://doi.org/10.1515/dx-2020-0122

Mahajan, P., Singh, H., & Newman-Toker, D. E. (2025). From diagnostic errors to diagnostic excellence in emergency care: Time to flip the script. Academic Emergency Medicine. Advance online publication. https://doi.org/10.1111/acem.15033

Marks, C. M., Gibney, S., Stenson, B., Sarma, D., Gaudet, C., Mombini, H., Buckley, T. A., Keko, M., Nathanson, L. A., Burke, L. G., Shapiro, N. I., Burstein, J. L., Grossman, S. A., Parab, A., Janke, A. T., Manrai, A. K., Taylor, R. A., Rosen, C. L., Rodman, A., & Haimovich, A. D. (2026). Screening for missed opportunities for diagnosis in the ED using eTriggers and large language models. JAMA Network Open, 9(6), e2620939. https://doi.org/10.1001/jamanetworkopen.2026.20939

Michelson, K. A., Bachur, R. G., Rangel, S. J., Monuteaux, M. C., Mahajan, P., & Finkelstein, J. A. (2023). Emergency department volume and delayed diagnosis of pediatric appendicitis: A retrospective cohort study. Annals of Surgery, 278(6), 905-912. https://doi.org/10.1097/SLA.0000000000005972

Michelson, K. A., Reeves, S. D., Grubenhoff, J. A., Cruz, A. T., Chaudhari, P. P., Dart, A. H., Finkelstein, J. A., & Bachur, R. G. (2021). Clinical features and preventability of delayed diagnosis of pediatric appendicitis. JAMA Network Open, 4(8), e2122248. https://doi.org/10.1001/jamanetworkopen.2021.22248

Newman-Toker, D. E., Nassery, N., Schaffer, A. C., Yu-Moe, C. W., Clemens, G. D., Wang, Z., Zhu, Y., Saber Tehrani, A. S., Fanai, M., Hassoon, A., & Siegal, D. (2024). Burden of serious harms from diagnostic error in the USA. BMJ Quality & Safety, 33(2), 109-120. https://doi.org/10.1136/bmjqs-2021-014130

Peters, M. D. J., Marnie, C., Tricco, A. C., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C. M., & Khalil, H. (2020). Updated methodological guidance for the conduct of scoping reviews. JBI Evidence Synthesis, 18(10), 2119-2126. https://doi.org/10.11124/JBIES-20-00167

Rising, K. L., Doty, A. M., Moyer, C. A., Hsiao, C. J., Barg, F. K., & Santen, S. A. (2023). Communicating diagnostic uncertainty at emergency department discharge: A simulation-based mastery learning randomized trial. Academic Medicine, 98(5), 617-624. https://doi.org/10.1097/ACM.0000000000004993

Schols, L. A., Maranus, M. E., Rood, P. P. M., & Zwaan, L. (2024). Diagnostic discrepancies in the emergency department: A retrospective study. Journal of Patient Safety, 20(6), 420-425. https://doi.org/10.1097/PTS.0000000000001252

Sharp, A. L., Baecker, A., Nassery, N., Park, S., Hassoon, A., Lee, M.-S., Peterson, S., Pitts, S., Wang, Z., Zhu, Y., & Newman-Toker, D. E. (2021). Missed acute myocardial infarction in the emergency department-standardizing measurement of misdiagnosis-related harms using the SPADE method. Diagnosis, 8(2), 177-186. https://doi.org/10.1515/dx-2020-0049

Vaghani, V., Gupta, A., Mir, U., Wei, L., Murphy, D. R., Mushtaq, U., Sittig, D. F., Zimolzak, A. J., & Singh, H. (2025). Implementation of electronic triggers to identify diagnostic errors in emergency departments. JAMA Internal Medicine, 185(2), 143-151. https://doi.org/10.1001/jamainternmed.2024.6214