

Just how many diagnostic errors and harms are out there, really? It depends on how you count

David E Newman-Toker  ^{1,2}

¹Department of Neurology,
Johns Hopkins University
School of Medicine, Baltimore,
Maryland, USA

²Department of Epidemiology,
Johns Hopkins University
Bloomberg School of Public
Health, Baltimore, Maryland,
USA

Correspondence to

Dr David E Newman-Toker;
toker@jhu.edu

Received 1 November 2024

Accepted 4 February 2025

Published Online First

15 March 2025

The significant adverse consequences of diagnostic errors are well established.^{1 2} Across clinical settings and study methods, diagnostic adverse events often lead to serious permanent disability or death and are frequently deemed preventable.^{3–5} In malpractice claims, diagnostic adverse events consistently account for more total serious harms than any other individual type of medical error,^{5 6} a finding supported by large, population-based estimates of total serious misdiagnosis-related harms.² Despite this, they generally go unrecognised, unmeasured and unmonitored, causing the US National Academy of Medicine to label diagnostic errors as ‘a blind spot’ for healthcare delivery systems.¹

Diagnostic errors have been described as ‘the bottom of the iceberg’ of patient safety. This analogy is intended to connote both their enormous impact and their unmeasured, hidden nature relative to more visible errors such as medication-dosing errors or wrong-site surgeries. The analogy, however, can be extended to an additional aspect: as we dive deeper and remeasure the iceberg’s circumference, it keeps getting bigger. US annual estimates of hospital-based misdiagnosis-related harms have increased from ~40 000–80 000 (deaths only, using hospital autopsies)⁷ to ~250 000 (harms of any severity, based on a systematic review of 22 studies using various methods)⁸ to ~376 000 (high-severity harms, including deaths, based on Dalal *et al*’s new single-centre study of multiple data sources including 90-day follow-up).⁹ Put differently, the apparent magnitude of the problem may be heavily influenced by how we decide to measure it. For error and harm rates, this principally means how we define both the numerator (nature of errors or harms measured) and denominator (population sampled)¹⁰ (figure 1a).

WHAT ARE THE NEW ESTIMATES OF SERIOUS MISDIAGNOSIS-RELATED HARMS IN HOSPITALISED PATIENTS?

In this issue of the journal, Dalal and colleagues add to the growing body of evidence on harms from diagnostic error.⁹ They conducted a retrospective study of hospitalised patients on a general internal medicine service in the US measuring diagnostic errors and harm severity by dual-rater chart review using standardised instruments. Errors were separately judged for process failures and preventability. By design, disagreements underwent consensus discussion, and borderline cases triggered a tertiary panel review by five other clinicians, thereby increasing the overall rigour of assessment. Authors used weighted sampling on known markers of diagnostic failure to enrich their sample with likely diagnostic error cases, while still maintaining the ability to estimate error and harm rates across all inpatients. They sampled 675 of 9147 eligible cases (representing 7390 unique patients). Inter-rater agreement was moderate for diagnostic error determination (kappa 0.52) but excellent after the panel review (kappa 0.87). The main study results included weighted total inpatient population rates for harmful diagnostic errors (7.2%, 95% CI 4.7 to 9.8), preventable harmful diagnostic errors (6.1%, 95% CI 3.8 to 8.5) and severely harmful diagnostic errors (1.1%, 95% CI 0.55 to 1.68). Process failures linked to increased odds of harmful diagnostic error mainly involved bedside diagnostic decision-making (case formulation and diagnostic test ordering), as in prior studies.^{3 6} Roughly 85% of harmful errors were deemed preventable.

If Dalal *et al*’s estimates are representative of all ~34 million US hospital admissions occurring annually, this translates



► <http://dx.doi.org/10.1136/bmjqs-2024-017183>



© Author(s) (or their employer(s)) 2025. No commercial re-use. See rights and permissions. Published by BMJ Group.

To cite: Newman-Toker DE. *BMJ Qual Saf* 2025;**34**:355–360.

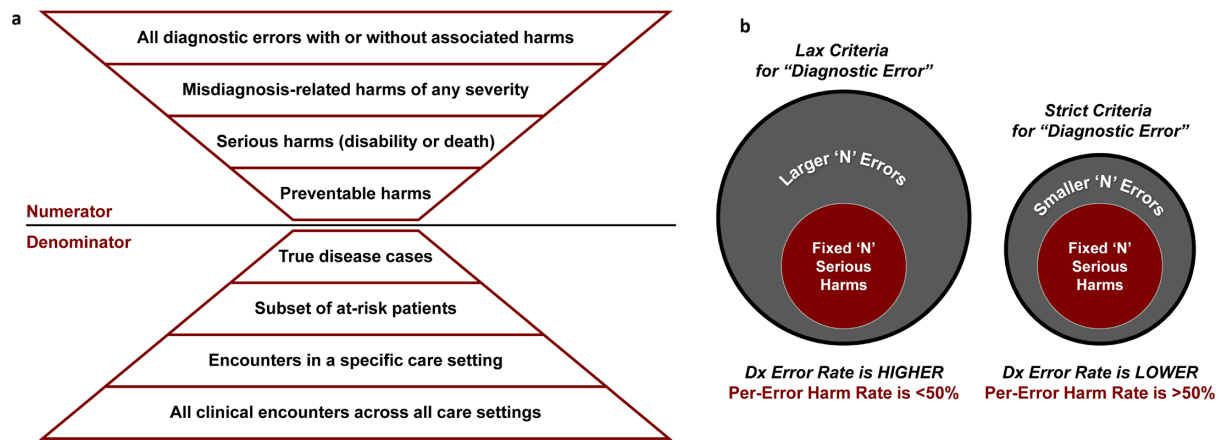


Figure 1 Key factors influencing the estimates of diagnostic error and misdiagnosis-related harm rates (adapted from Newman-Toker *et al*).¹⁰ Panel (A) at left indicates the relative magnitude of defining factors (numerator and denominator) impacting a measured error or harm rate (with wider bars indicating larger numbers).¹⁰ The specific combination of definitions for numerator and denominator can lead to widely divergent estimates of diagnostic error or misdiagnosis-related harm rates. Preventable deaths as a function of all clinical encounters across all clinical settings may be <0.1%, while all diagnostic errors (including very short delays) in true disease cases presenting with a challenging clinical phenotype may be close to 100%. The case mix of diseases or symptoms may powerfully influence rates. For example, stroke is misdiagnosed ~11-fold more often than heart attack in the emergency department,¹¹ while stroke presenting with dizziness or vertigo is misdiagnosed ~14-fold more often than stroke presenting with motor weakness.¹¹ Panel (B) at right shows the relationship between lax vs strict definitions for 'diagnostic error' and the fixed, actual number of serious harms (permanent disability or death) that result from such errors. There is substantial room for differences in diagnostic error definitions across studies, but there is almost no room for differences in serious harm definitions across studies. While ascertainment may vary slightly by study, the clinical definition of death (or loss of a limb) is unambiguous. By contrast, diagnostic error definitions may be operationalised in widely divergent manners across studies. This often includes *objective* criteria (eg, 'How long must a delay be to be defined as a diagnostic delay or error?') and typically also includes *subjective* criteria (eg, 'Was this error due to a preventable process failure?'). This phenomenon can be detected across studies by examining the per-error serious harm rate for one or more dangerous diseases. The most straightforward examples come from studies that define two thresholds for diagnostic error based on different-length delays to diagnosis, as with cancer diagnosis. Measured values for per-error harm rates can vary widely from <1% (in studies defining an 'error' based on a very short delay, for example, <1 month) to values of 25%–50% or more (in studies defining an 'error' based on a longer delay, for example, >6 months). Diagnostic error rates in such studies can vary from <10% (for longer delays) to >50% (for shorter delays).

to ~2.4 million harmful diagnostic errors including ~376 000 inpatients suffering disabling or fatal misdiagnosis each year (~319 000 of these preventable). This is 2.8-fold higher than that reported in our recent US study that combined estimates from Gunderson *et al*⁸ and Zwaan *et al*⁴ (see Newman-Toker *et al*,² supplement C3). Why? Possibly because Dalal *et al* looked harder—using more case triggers, more data sources, more rigorous case assessment and more follow-up. In Gunderson's systematic review, for example, few source studies took such pains to discover diagnostic errors so rigorously.

Of course, it is hard to know whether Dalal *et al*'s estimates should be applied to *all* US hospitalisations. Some hospitalisations are elective surgical admissions, in which, presumably, the risk of diagnostic error or associated harm is much lower. Perhaps those admitted to general medicine have particularly challenging clinical presentations or are less likely to benefit from consultative expertise during their stay. Any of these might render extrapolations from Dalal *et al*'s numbers unjustified or overstated.

Alternatively, such extrapolations might be underestimated. One could argue that, since disabilities and deaths from diagnostic error are less frequent at academic institutions^{11 12} (which account for only a minority of hospital stays), the average US

hospital-based diagnostic error rate is actually much higher than at 'a large academic medical centre in Boston'. Or one might point out that some out-of-hospital/out-of-health-system deaths (or events after 90 days) went uncounted, and cases without trigger events (which have lower harmful error rates but are far more plentiful) were probably undersampled.

Perhaps more importantly, there is evidence that diagnostic errors ascertained by retrospective chart review (as in Dalal *et al*) are likely substantial underestimates.¹³ Schwartz *et al* used standardised patients to assess diagnostic errors and associated costs in general medicine clinics, finding that just 5% of costs attributable to inappropriate tests or treatments were evident from examining charted documentation after the 'secret shopper' visit.¹⁴ In two studies of emergency department diagnostic error from a single academic group, diagnostic adverse events and misdiagnosis-attributable mortality were 18-fold and 27-fold higher, respectively, when measured prospectively (including follow-up calls) versus retrospectively (by triggered chart review) at the same two academic hospitals (Newman-Toker *et al*,¹¹ 'Differences in Estimation Based on Study Design').

One way to gauge the overall plausibility of these hospital-based extrapolations of harm from diagnostic error is to see if they align well with what is known about total serious harms across all clinical settings.

Table 1 Estimated diagnostic errors and serious misdiagnosis-related harms by clinical setting in the USA annually*

Clinical setting	Visits (n)†	Diagnostic errors (n)‡	Serious harms (n)‡
Inpatient	34 million	~2.4 million (95% CI 1.6M to 3.3M) ⁹	~376 000 (95% CI 185K to 566K) [§]
Emergency department	140 million	~7.2 million (95% CI 5.0M to 11.2M) ¹¹	~433 000 (95% CI 266K to 1062K) ¹¹
Primary care clinics	521 million	~32.8 million (95% CI 31.9M to 33.7M) ¹³	~231 000 (speculative) ²
Specialty care clinics	515 million	~10–30 million (speculative)	~114 000 (speculative) ²
All visits combined	1.2 billion	~50–100+ million (speculative)¶	~909 000 (PR 684K to 1170K)**

*The estimates in this table are based on best available evidence. However, there are limitations in making estimates or extrapolations based on studies that use different methods, data sources and measures. In cases where there are extremely high levels of uncertainty or limited data, such estimates are denoted as 'speculative'.

†Annual US visits are from data published by the American Hospital Association (inpatient) and Centers for Disease Control and Prevention National Center for Health Statistics (emergency department, primary care, specialty care) from the most recent years available: inpatient (2022), emergency department (2021), ambulatory clinics (2019).

‡When prior estimates were based on visit data from earlier years, diagnostic errors and serious misdiagnosis-related harms have been scaled from source studies to the total visits from the most recent years available. For Dalal *et al*⁹ and Singh *et al*¹³ values are derived as the mathematical product between published error rates and number of visits. For those listed as 'speculative', a rationale is provided in the text, footnotes or cited sources.

§The proportion of serious harms attributable to the inpatient clinical setting is ~41% (n=376, 000/~909, 000), with a plausible range between ~20% (~185, 000/~909, 000) and ~62% (~566, 000/~909, 000) (using the numerator's upper and lower confidence bound divided by the point estimate for the denominator).

¶Rates of diagnostic error reported across frontline clinical settings in various studies are in the range of 5.1% to 7.2%.^{9 11 13} If we assume specialty clinic error rates (on which there are no aggregate data) are the same as those seen in primary care, then the visit-volume-based weighted, mean error rate across all clinical settings would be 6.2% (PR 5.8 to 6.7). With 1.2 billion healthcare visits annually in the US, the total number of errors would then be ~75 million (PR 70 to 81). If the error rate in specialty care were lower (eg, roughly one-third of what is seen in primary care), the total would be ~50 million. However, most of these studies systematically underestimate diagnostic errors with minor harms either unintentionally or by design (see text). Experts in the field have speculated an overall error rate in the 10% to 15% range, which would translate to 121–181 million errors. Therefore, we offer a speculative range of 50–100+ million US diagnostic errors annually across all clinical settings.

**The combined number of serious harms for all visits (~909 000) should be less than the sum across clinical settings (~1 154 000), mainly because it is expected that some of the serious harm counts are partly overlapping (eg, from a diagnostic delay that begins in the emergency department but carries forward through to the inpatient setting). The proportion of deaths (46.7%)² corresponds to ~424 000 deaths each year due to diagnostic error.

K, thousand; M, million; PR, plausible range.

DOES THIS FIT WITH WHAT WE KNOW ABOUT SERIOUS MISDIAGNOSIS-RELATED HARMS ACROSS ALL SETTINGS?

Robust aggregate estimates at the population level are difficult to construct. No single study can rigorously ascertain diagnostic errors and harms in a representative sample across all clinical contexts.

In 2023, our team published a unique approach to estimating overall serious misdiagnosis-related harms based on diseases, rather than clinical settings.² We combined the US population incidence of dangerous diseases with literature-derived disease-specific diagnostic error and harm rates, validated these through expert input and extrapolated to all diseases using prior clinical studies; we then triangulated our results via other sources and methods and performed multiple sensitivity analyses of key assumptions used in our estimation procedures.² We estimated ~795 000 (598K–1023K) serious harms in the USA in 2014,² which, using more recent data on visit volumes, extrapolates to ~909 000 (684K–1170K) (table 1).

Is our estimate plausible? The short answer is yes. We know some things for sure about the number of misdiagnosis-related harms. The most obvious is that there cannot be more annual US deaths *due to misdiagnosis* than the number of annual US deaths *in total* (~3.3 million in 2022). It is also very likely that the proportion of all deaths due to diagnostic error is between 5% and 20%,² further constraining plausible estimates of misdiagnosis-attributable deaths to

roughly 164 000–656 000 per year. Similar logic might be used to approximate the number of patients experiencing misdiagnosis-attributable serious disability, but the annual US incidence of serious disability is less readily available than national death statistics, and the misdiagnosis-attributable disability fraction is unknown. An alternative is simply to estimate serious, misdiagnosis-attributable morbidity directly from misdiagnosis-attributable deaths, based on their relative proportions in malpractice claims (where deaths slightly outnumber serious disabilities, by 1.1-fold to 1.2-fold)^{5 6} or clinical studies of diagnostic error (where serious disabilities slightly outnumber deaths, by 1.1-fold).² This set of calculations would suggest ~308 000 to ~1 422 000 permanent, serious harms representing deaths plus disabilities (equivalent to permanent loss of an arm, eye, kidney or more severe)⁶ each year in the US from diagnostic error. This defines a plausible range (308K–1422K) slightly wider but overlapping with the one (684K–1170K) (table 1) based on our team's method.

We can gauge the plausibility of extrapolating Dalal *et al*'s serious harm findings to all US hospitalisations by looking at the proportion of serious harms across all clinical settings attributable to inpatient settings. Dividing the extrapolation from Dalal *et al* (~376 000) by the total (~909 000) gives ~41% (range ~20%–62% (see table 1 footnote)) attributable to inpatient settings. This range overlaps the proportion of serious harms attributed to inpatient settings seen in diagnostic

malpractice claims (~28%–34%).^{5 6} Although the absolute number of harms cannot be calculated from malpractice studies, there is little reason to believe that malpractice data are strongly biased with respect to the proportion of harms that are inpatient. Thus, this helps corroborate that extrapolating Dalal *et al*'s findings to all hospitalisations is plausible.

HOW DO SERIOUS HARMS RELATE TO THE NUMBER OF DIAGNOSTIC ERRORS ACROSS CLINICAL SETTINGS?

While the total number of Americans suffering permanent disability or death from diagnostic error is estimated to be between 0.7 and 1.2 million annually (and highly unlikely to be substantially outside that range), there is far less known about the total number of diagnostic errors leading to these harms. Why? At least partly, this is because errors have less consistent definitions than disability or death (figure 1b).

Interestingly, despite some variation in study methods, error rate estimates across frontline care settings have been similar (~4% to 10%, accounting for uncertainty): 7.2% (95% CI 4.7 to 9.8) among inpatients,⁹ 5.2% (95% CI 3.6 to 8.0) in emergency departments¹¹ and 6.3% (95% CI 6.1 to 6.5) in primary care.¹³ There are no estimates of diagnostic errors across specialty care, but if they occur at rates similar to primary care, the visit volume-weighted average rate would be 6.2% and the total number of diagnostic errors across all settings would be roughly ~75 million, including ~2.4 million in the hospital (table 1).

However, most studies underestimate the true frequency of diagnostic errors that lead to minor adverse events. The number of 'near misses' (with minimal or no harm) should generally exceed the number of errors with serious harms, but, by design, Dalal *et al* identified few such milder cases. Their weighted sampling schema oversampled on high-risk scenarios (eg, ICU transfer or death) and, unsurprisingly, identified very few cases (6%) with harms judged minor. In particular, chart review studies tend to underestimate errors in diagnosing common, benign conditions that cannot be readily discerned from chart documentation and instead require prospective data capture. For example, a recent prospective diagnostic trial showed diagnostic errors occur at a rate of more than 90% for patients with the most common self-limited peripheral vestibular disorder, benign paroxysmal positional vertigo (BPPV),¹⁵ which has an annual population incidence of 1.6%.¹⁶ Thus, each year in the US there could easily be 5 million misdiagnosed BPPV cases across clinical settings, most of which probably go uncounted in studies of error because they lack relevant charted evidence. There are presumably many other similar examples.

If the true diagnostic error rate across clinical settings is closer to that which has been previously speculated

by experts (ie, 10% to 15%), then total diagnostic errors would be 120–180 million (table 1).

WHAT DON'T WE KNOW ABOUT DIAGNOSTIC ERRORS AND HARMS?

Most studies do not consider (or massively undercount) false-positive errors or harms associated with overtesting, overdiagnosis or treatment for conditions not actually present. Most also fail to consider 'correct' diagnoses where disease severity is misestimated (staging errors), leading to patient harm. Few account for diagnostic delays in accessing healthcare initially after symptoms begin, a period known as the *patient interval*, which is clearly implicated in delayed cancer diagnosis¹⁷ and likely other conditions. Likewise, they generally fail to assess diagnostic errors in other preclinical settings such as urgent care clinics or nursing homes. None fully account for lesser-severity harms than permanent, severe disability or death.

Very little is known about aggregate burden of misdiagnosis-related harms that are less serious than losing a limb. This includes undiagnosed chronic pain syndromes, mental disorders, vertigo and much more. The WHO rates migraine and depression as two of the top causes of (potentially reversible) global disability, and both affect tens of millions of Americans. Since delays in diagnosis are common, millions each year likely suffer temporary disability from misdiagnosed migraine, depression and BPPV alone. Whether the public health impact of such harms (which are lesser in severity but greater in number) collectively exceeds the magnitude of more serious harms is unknown.

Almost nothing is known about aggregate diagnostic errors and harms (serious or not) worldwide. Error rates in some other high-income countries and regions (eg, Western Europe, Canada and Australia) are similar to those in the USA.¹¹ Those in some low-income and middle-income countries are likely much higher,⁸ if for no other reason than that access to healthcare and even basic diagnostic tests is often lacking.² With a world population 24-fold larger than that of the USA, there could easily be billions of diagnostic errors and tens of millions of serious misdiagnosis-related harms across the globe, although many of these are probably diagnostic delays in the prehealthcare setting (ie, patient interval).

ARE MISDIAGNOSIS-RELATED HARMS REALLY PREVENTABLE?

The preventability of misdiagnosis-related harms will always be questioned by some diagnostic safety sceptics, but three things should give us hope. First, inequities in diagnosis, where women as well as racial and ethnic minorities are at 20%–30% higher risk than white men,¹¹ suggest at least part of the problem should be reducible. Second, concerted efforts have succeeded in the areas of chest pain and acute coronary syndromes—new diagnostic tests, chest pain protocols

and quality metrics attached to regulatory and financial incentives have made myocardial infarction the one well-studied, dangerous disease where error rates (1%–2%) and serious harms (0.5%–1%) are both extremely low.² Third, new diagnostic approaches for other conditions, such as stroke presenting with vertigo, show tremendous promise for increasing diagnostic accuracy while simultaneously decreasing inappropriate testing.¹⁸

HOW SHOULD WE MEASURE GOING FORWARD?

It is hard to gain consensus on how best to define and measure diagnostic errors and harms. In 2015, the US National Academy of Medicine created what is now the most widely used definition of a *diagnostic error* ('the failure to (a) establish an accurate and timely explanation of the patient's health problem(s) or (b) communicate that explanation to the patient'),¹ but some diagnostic error researchers insist on their own terminology and definitions. Measurement methods flow from definitions and are often quite contentious. A more comprehensive set of recommendations on measurement is provided elsewhere.¹¹ Briefly, three types of measurement are needed, all of which would ideally be standardised:

1. **Serious misdiagnosis-related harms.** Death and serious, permanent disability are unambiguous, but the link between diagnostic errors and these serious harms is often disputed, especially if they are determined by retrospective chart review. Newer methods such as the Symptom-disease Pair Analysis of Diagnostic Error (SPADE) have addressed this issue by using clinically valid control groups or assessing observed greater than expected rates of harm that are statistically significant.¹⁹ In most high-income countries, deaths are routinely reported; adding serious disability reporting (using a schema such as the National Association of Insurance Commissioners Severity of Injury Scale)^{5 6} would facilitate standardised SPADE-based measurement of serious misdiagnosis-related harms associated with common, dangerous diseases.
2. **Time-to-diagnosis.** This straightforward, disease-specific measurement framework has been used successfully by the NHS in the UK to reduce time to cancer diagnosis. The most challenging aspect of measurement is systematically ascertaining (usually directly from patients) the time of first *related* symptoms, which may require clinical confirmation or adjudication; this is crucial to establish the patient interval. Importantly, it can be applied to almost any disease, including hyperacute diseases (eg, aortic dissection), as long as the time scale used is minutes to hours rather than days to years.
3. **Patient experience in diagnosis.** Disease-based tracking of serious harms from diagnostic error and time-to-diagnosis cannot assess the patient's experience in diagnosis, including effectiveness of communicating a (correct) diagnosis. Incorporating patient-reported measures allows tracking and aggregation of diagnostic

experience (including communication) across symptoms, diagnoses and clinical settings.²⁰ When assessed in delayed (post-encounter) fashion, direct patient feedback also permits systematic measurement of inaccurate or delayed diagnoses after a follow-up period,²⁰ thereby catching additional diagnostic errors for diseases not already systematically tracked.

If these methods can be standardised, institutional self-monitoring or external benchmarking become straightforward. These measures avoid retrospective chart reviews to adjudicate the presence of clinical diagnostic process failures (ie, whether a correct diagnosis 'could' or 'should' have been made) as well as the preventability of any resulting harms, both of which are notoriously inconsistent parameters plagued by problems of inadequate documentation, the rater's clinical knowledge and hindsight bias. Instead, preventability should be judged by actual prevention—if an intervention is implemented that cuts misdiagnosis-related deaths in a particular clinical setting, then those were preventable deaths.

Local institutions can, and should, measure diagnostic process, but these efforts should be guided first by *diagnostic outcome measures*, which assess the direct impact of diagnostic errors on patient health. From a policy perspective, diagnostic outcome measures (eg, using SPADE) should be the focus of public reporting and incentive programmes, while process measures should be formative rather than summative.

CONCLUSIONS

There is always some uncertainty inherent in extrapolations beyond original study source populations or comparing data from studies that use different methods to measure diagnostic errors or adverse events. So, bearing caveats in mind, what do we *really* know about the frequency of diagnostic errors and associated harms? At the very least, we know that this is an important and impactful patient safety concern—one clearly big enough that we should stop arguing exactly how many errors and harms are out there (or are preventable) and move on to the business of fixing the problem. Someday we may dive deep enough to reach the iceberg's maximum circumference, but we probably have not yet reached that point for either diagnostic errors or total misdiagnosis-related harms. We know the issue is at least partly remediable, and fixing it offers the potential to improve care quality while reducing costs of care. Thus, policy efforts should focus more on diagnostic outcomes and less on diagnostic process. Given that serious harms from diagnostic errors outnumber those from any other medical error category, whatever resources we have devoted to address other safety issues (eg, healthcare-associated infections) over the past 25 years, we should devote at least that much to diagnostic safety and quality over the next quarter century.

Acknowledgements Megan Clark, MWC, provided editing and manuscript preparation assistance.

Contributors DEN-T is responsible for the overall content as guarantor and accepts full responsibility for the finished work.

Funding This work was supported by the US Agency for Healthcare Research and Quality (grant R18HS029350) and the Armstrong Institute Center for Diagnostic Excellence.

Competing interests DEN-T reports grants related to dizziness/stroke diagnosis and diagnostic error from the Agency for Healthcare Research and Quality, Gordon and Betty Moore Foundation, Society to Improve Diagnosis in Medicine, and Natus Inc.; medicolegal consulting for both plaintiff and defence firms related to misdiagnosis of neurological conditions, including dizziness and stroke; and honoraria for lectures on these topics; outside the submitted work.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Commissioned; internally peer-reviewed.

ORCID iD

David E Newman-Toker <http://orcid.org/0000-0003-2789-4115>

REFERENCES

- National Academies of Sciences, Engineering, and Medicine. Improving diagnosis in health care. The National Academies Press; 2015. Available: <https://doi.org/10.17226/21794>
- Newman-Toker DE, Nassery N, Schaffer AC, *et al.* Burden of serious harms from diagnostic error in the USA. *BMJ Qual Saf* 2024;33:109–20.
- Schiff GD, Hasan O, Kim S, *et al.* Diagnostic error in medicine: analysis of 583 physician-reported errors. *Arch Intern Med* 2009;169:1881–7.
- Zwaan L, de Bruijne M, Wagner C, *et al.* Patient record review of the incidence, consequences, and causes of diagnostic adverse events. *Arch Intern Med* 2010;170:1015–21.
- Saber Tehrani AS, Lee H, Mathews SC, *et al.* 25-Year summary of US malpractice claims for diagnostic errors 1986–2010: an analysis from the National Practitioner Data Bank. *BMJ Qual Saf* 2013;22:672–80.
- Newman-Toker DE, Schaffer AC, Yu-Moe CW, *et al.* Serious misdiagnosis-related harms in malpractice claims: The “Big Three” - vascular events, infections, and cancers. *Diagnosis (Berl)* 2019;6:227–40.
- Leape LL, Berwick DM, Bates DW. Counting Deaths Due to Medical Errors—Reply. *JAMA* 2002;288:2405.
- Gunderson CG, Bilan VP, Holleck JL, *et al.* Prevalence of harmful diagnostic errors in hospitalised adults: a systematic review and meta-analysis. *BMJ Qual Saf* 2020;29:1008–18.
- Dalal AK, Plombon S, Konieczny K, *et al.* Adverse diagnostic events in hospitalised patients: a single-centre, retrospective cohort study. *BMJ Qual Saf* 2025;34:377–88.
- Newman-Toker DE, Wang Z, Zhu Y, *et al.* 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 (Berl)* 2021;8:67–84.
- Newman-Toker DE, Peterson SM, Badihian S, *et al.* *Diagnostic Errors in the Emergency Department: A Systematic Review. AHRQ Comparative Effectiveness Reviews.* Rockville, MD: Agency for Healthcare Research and Quality, 2022.
- Obermeyer Z, Cohn B, Wilson M, *et al.* Early death after discharge from emergency departments: analysis of national US insurance claims data. *BMJ* 2017;356:j239.
- Singh H, Meyer AND, Thomas EJ. The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations. *BMJ Qual Saf* 2014;23:727–31.
- Schwartz A, Weiner SJ, Weaver F, *et al.* Uncharted territory: measuring costs of diagnostic errors outside the medical record. *BMJ Qual Saf* 2012;21:918–24.
- Badihian S, Zee DS, Carey JP, *et al.* Remote expert diagnosis by video-oculography is more accurate than in-person ED diagnosis in acute vertigo and dizziness—preliminary results of the AVERT Trial. Abstracts of the 31st Bárány Society Meeting, Madrid, Spain. *J Vestib Res* 2022;S136–7.
- von Brevern M, Radtke A, Lezius F, *et al.* Epidemiology of benign paroxysmal positional vertigo: a population based study. *J Neurol Neurosurg Psychiatry* 2007;78:710–5.
- Lyratzopoulos G, Saunders CL, Abel GA, *et al.* The relative length of the patient and the primary care interval in patients with 28 common and rarer cancers. *Br J Cancer* 2015;112 Suppl 1:S35–40.
- Gold D, Peterson S, Bosley J, *et al.* Diagnostic impact of a device-enabled remote “Tele-Dizzy” consultation service in the emergency department. Abstracts of the 31st Bárány Society Meeting, Madrid, Spain. *J Vestib Res* 2022;32:S71–2.
- Liberman AL, Wang Z, Zhu Y, *et al.* Optimizing measurement of misdiagnosis-related harms using symptom-disease pair analysis of diagnostic error (SPADE): comparison groups to maximize SPADE validity. *Diagnosis (Berl)* 2023;10:225–34.
- Gleason KT, Dukhanian V, Peterson SK, *et al.* Development and Psychometric Analysis of a Patient-Reported Measure of Diagnostic Excellence for Emergency and Urgent Care Settings. *J Patient Saf* 2024;20:498–504.