*Burden of Serious Harms from Diagnostic Error in the USA*        *Newman-Toker et al., BMJQS 2023*

## Supplemental File #1 for *Burden of Serious Harms from Diagnostic Error in the USA* by Newman-Toker, et al. BMJ Quality & Safety, 2023 (doi:10.1136/bmjqs-2021-014130)

Enclosed below are the main Supplementary Materials (Supplement 1) to the peer-reviewed, scientific journal publication entitled *Burden of Serious Harms from Diagnostic Error in the USA (2023).* Some of the methods descriptions (particularly in Supplement 1, section A5 about Monte Carlo analysis) are very similar to methods descriptions from the related, previously published manuscripts from earlier project phases. This is unavoidable since the statistical methods for the current manuscript were the same. An ancillary appendix (Supplement 2) provides the full statistical code for the Monte Carlo analysis.

Table of Contents for Supplement 1

*Abbreviations Not Necessarily Defined in the Text*

N/n – number
M – million
B – billion
CI – confidence interval
PR – plausible range
PPR – probabilistic plausible range

*Burden of Serious Harms from Diagnostic Error in the USA*      *Newman-Toker et al., BMJQS 2023*

## A) Methodological Details

### A1. Study Methods – Rationale for Overall Architecture

Seeking valid estimates of disability and death from diagnostic error is important for public policy, yet methodologically challenging.[1] Diagnostic errors will necessarily be more frequent than diagnostic adverse events (of any severity), which, in turn, will be more frequent than serious misdiagnosis-related harms (i.e., permanent disability or death). Key concerns in estimation of serious harms[1] include (1) proper quantitative synthesis of the literature on error and harm rates, rather than extrapolations based on single studies conducted in non-representative settings; (2) judgments about attributable harms (i.e., the extent to which diagnostic errors result in serious harms or are potentially confounded by comorbidity); (3) judgments about preventability, including whether prevention will result in meaningful gains in healthy life years; (4) methodological risks of undercounting or overcounting serious harms, including double counting of deaths in patients who suffer more than one error; and (5) applying error and harm rate estimates to the appropriate population at risk and conducting methodologically robust statistical analyses to account for uncertainty associated with relatively low frequency events (i.e., serious harms).

Taking a disease-specific (rather than disease-agnostic) approach to measurement helps address many of these methodological challenges. It is easier to more rigorously and precisely measure diagnostic errors, harms, or preventability in disease-specific than disease-agnostic fashion since research studies that cut across diseases cannot incorporate rigorous reference standards for diagnostic accuracy or error for every possible condition. Synthesis of multiple studies via systematic review with meta-analysis is also more straightforward since disease-specific studies are more homogeneous in their disease and error definitions. Finally, a consistent finding from the literature on diagnostic error, whether derived from malpractice claims or clinical practice, is that vascular events, infections, and cancers (together known as the "Big Three" dangerous-disease categories) are responsible for three-quarters of serious misdiagnosis-related harms.[2,3] This permits extrapolation from these specific disease categories to all diseases.

As noted in the main manuscript Methods section, the overall study was designed to estimate the total annual burden (incidence) of serious harms from diagnostic error in the US. It was conducted in three study phases (Figure S1): (1) identify top diseases misdiagnosed that cause serious misdiagnosis-related harms (from a large, nationally representative malpractice data set previously coded for claim type, outcome/harm severity, and disease then comparing the proportion of "Big Three" diseases to clinical practice-based [non-claims] studies)[2]; (2) find disease-specific diagnostic error and harm rates for top harm-causing diseases (from clinical literature and vetted by experts)[4,5]; (3a) measure annual population incidence for each disease (from public use, nationally representative data sets) and (3b) combine error and harm rates with incidence to estimate total annual US incidence of serious diagnostic errors and harms. The first two study phases were published previously[2,4]; the final phase (3a, 3b) is included here.
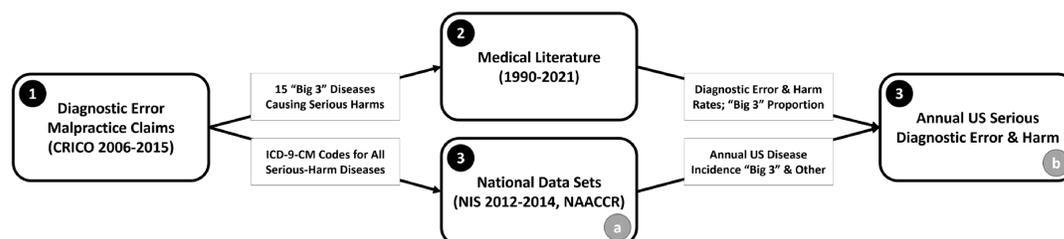


**Figure S1. Overview of scientific methods for estimating total diagnostic errors and serious misdiagnosis-related harms.** *Abbreviations: CRICO – Controlled Risk Insurance Company, Ltd.; ICD-9-CM – International Classification of Diseases 9th Revision, Clinical Modification; NAACCR – North American Association of Central Cancer Registries; NIS – National Inpatient Sample.*

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

*BMJ Qual Saf*

*Burden of Serious Harms from Diagnostic Error in the USA*          *Newman-Toker et al., BMJQS 2023*

### A2. Role of Malpractice Claims in Estimation and Independence of Grand Total Harm Estimates

Some readers may wonder how it can be that our study began with a first phase that used malpractice claims data, yet the final estimates are somehow independent? As noted in the main manuscript Methods section, our scientific approach was constructed such that the final grand total estimates for errors and harms in the US are based on clinical literature and US population incidence, not malpractice claims. This is because (a) no error or harm rates were taken from claims-based studies, (b) the extrapolation from specific "Big Three" disease estimates to the grand total were based on the proportion of "Big Three" diseases causing errors and harms from clinical studies, and (c) any impact of having used malpractice claims to construct the original disease list or weights are mathematically unrelated to the grand totals.

**Why were malpractice data used in the first place?** Malpractice data were used to construct the initial list of diseases likely to be responsible for the greatest numbers of total misdiagnosis-related harms. We needed a starter list of such diseases in order to be able to take the disease-based approach (A1 above).

**If malpractice data were good enough to make the disease list, why not just stop there?** First, it is known that many medical errors never lead to malpractice claims, so it is hard to extrapolate even from a representative, national claims database to a true national estimate. In the Harvard Medical Practice Study, which compared negligent medical errors to malpractice claims, the chance that an injury caused by medical negligence would result in litigation was just 1.5% (95% CI 0-3.2).[6] Second, malpractice claims data are known to represent a biased sample. Some forms of bias in malpractice claims are well known, while others may be hidden. The most well-known bias in malpractice claims data is towards higher-severity harms. This is not necessarily a problem for estimating serious harms[7] (as in this study), but there may also be maldistributions of claims (i.e., non-representativeness) based on other factors as well. For example, myocardial infarction is probably overrepresented in claims relative to stroke as a cause of misdiagnosis-related harms—disease incidence is similar, diagnostic error rates are ~10-fold higher for stroke, and disabling neurologic injuries result in the highest claims payouts (so are more likely to spark a claim), yet numbers of claims are only ~1.5-fold higher for stroke; this could be because legal "standard of care" expectations for accuracy of heart attack diagnoses are higher than those for stroke.[5]

**Were malpractice data used for any other purpose in constructing the estimates?** Yes, as noted in the main manuscript Methods section, misdiagnosis-related harm rates were derived by combining high-quality data on disease-agnostic (non-disease specific) harms per diagnostic error from well-respected clinical studies then applying disease-specific harm-severity weights from malpractice claims.[4] A disease-agnostic approach was required because there were not a sufficient number of disease-specific studies examining attributable harm rates. We weighted the disease-agnostic, per-diagnostic-error serious harm rate for each disease to get a more realistic estimate of harms (e.g., aortic dissection is more likely lethal than pneumonia when initially missed, so assigning the same risk of serious harms *per diagnostic error* for each of the two diseases would have been inappropriate). For each disease, we multiplied the disease-specific, clinical literature-based diagnostic error rate ***by*** the clinical literature-based disease-agnostic per-error harm rate ***by*** a disease-specific, claims-based harm-severity weight. This weight was based on the disease-specific proportion of malpractice cases resulting in serious vs. non-serious harms (e.g., higher weight for aortic dissection than pneumonia). The weighting procedure was also used to prevent overcounting of harms from "other" (non-top 5) diseases. Full statistical details of this approach can be found in our prior publication's Supplementary materials (*Supplement A2, Requirements R1 and R4[4]*).

**If that is true, then how can the final total estimates be claims-independent?** The final results are independent of malpractice claims because we mathematically "forced" the proportion of errors and serious harms attributable to all combined Big Three diseases to be equal to the known attributable fractions found in the clinical literature (*see our prior publication's Supplement A2, Requirements R2 and*

*Burden of Serious Harms from Diagnostic Error in the USA*     *Newman-Toker et al., BMJQS 2023*

R3[4]). This is described, in brief, in the Methods section of the main manuscript, "…"Big Three" results were used to calculate a grand total (including non-"Big Three" dangerous diseases) using the clinical proportion of diagnostic errors (58.5%) and serious harms (75.8%) attributable to "Big Three" diseases.[2,3] These proportions derive exclusively from research studies based in clinical practice (i.e., not malpractice claims studies) (*see Table 3 from our prior citation[2]*). Mathematically, the grand total of diagnostic errors was calculated by dividing the "Big Three" total number of diagnostic errors by 0.585. Similarly, the grand total of serious misdiagnosis-related harms was calculated by dividing the "Big Three" total number of serious misdiagnosis-related harms by 0.758." This forces independence from claims.

**Then what are the implications of malpractice-claims based intermediate steps for the results?**
There are two main potential impacts of these claims-based steps. First, it is possible that the lower-ranked "top 5" diseases might be over-ranked (e.g., it is possible that the unnamed 6th-ranked disease categorized in the "other" subcategory in one of the "Big Three" categories might actually be the *real* 5th-ranked disease in that category). For example, the 6th-ranked disease in the infection category in malpractice claims was appendicitis. There were more than twice as many claims for endocarditis as appendicitis, which is why we searched out data on error rates for endocarditis rather than appendicitis. However, if malpractice claims were somehow biased towards endocarditis or away from appendicitis, it is still potentially conceivable that appendicitis might outrank endocarditis as a cause of misdiagnosis-related harms in clinical practice, since appendicitis has more than twice the real-world incidence of endocarditis. However, this is unlikely, because endocarditis is initially missed an estimated ~26% of the time[4] and appendicitis is initially missed no more than ~5% of the time, more than compensating for the higher incidence of appendicitis.[5] Note that, in our final analysis, appendicitis is still accounted for in the "other" infections subcategory (so it has not gone uncounted). Second (and related), it is possible that the relative proportion of "other" (non-top 5) diseases are underrepresented relative to the top 5.

*In summary, serious harms estimated for individual diseases named (or unnamed) in Table 1 of the main manuscript are potentially impacted by unknown biases that could be present in malpractice claims. This could impact disease-specific rankings or the proportion of "other" (non-top 5) harms. However, the grand total harm estimates are mathematically fully independent of malpractice claims.*

### A3. Double Check of HCUP CCS Code Level Groupings Prior to NIS Incidence Analysis

Prior to NIS analysis of disease incidence in this third phase of the project, we performed a final cross-check at the code level using the HCUP CCS Level 3 groupings for vascular diseases and infections derived from the claims analysis.[2] The code lists were reviewed and any ICD codes unrelated to new, acute events (e.g., 438 "late effects of cerebrovascular disease") were removed prior to NIS analysis. We also reviewed all codes in the CRICO CBS data set to address issues of coding migration over time and reduce the risk that any specific codes might be missed because of sampling error in CRICO data during the years of analysis. From the wider code list, we found 18 related codes that belonged in the top 5 groupings (e.g., 433.0 "occlusion and stenosis of the basilar artery" and 435.3 "vertebrobasilar artery syndrome" for stroke) and added these before conducting the final NIS analysis. We did not consider Level 3 codes present in the parent HCUP CCS classification but <u>not</u> found in the malpractice claims data, to avoid any risk of overcounting non-life-or-limb-threatening diseases unlikely to cause harms. NIS analysis was run at both the ICD-9-CM code level as well as rolled up by disease and category to both (a) ensure sensibility and coherence and (b) identify any coding errors or gaps before being finalized.

### A4. NIS Sampling & Weighting Procedures to Derive Nationally Representative Estimates

We followed standard procedures for NIS data to derive nationally representative estimates, which use pre-specified discharge weights to convert an unweighted sample of hospital discharges into a weighted,

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

*BMJ Qual Saf*

*Burden of Serious Harms from Diagnostic Error in the USA*      *Newman-Toker et al., BMJQS 2023*

nationally representative sample.[8] The result is a weighted estimate for both disease incidence and patient demographics. Each year in the US there are roughly 36M inpatient hospitalizations[9] at more than 6,000 hospitals.[10] For each year since 2012, NIS has sampled more than 7M hospital discharge records from more than 4,000 acute-care hospitals (excluding long-term acute care hospitals). The discharge weights are calculated by NIS data curators by first stratifying the NIS hospitals on the same variables that were used for creating the sample. These variables are hospital Census division, urban/rural location, teaching status, bed size, and ownership. A weight is then calculated for each stratum, by dividing the number of universe discharges (i.e., all discharges) in that stratum, obtained from HCUP and American Hospital Association data, by the number of NIS discharges (i.e., sampled discharges) in the stratum. Discharge weights are assigned to each sampled discharge by NIS data curators and are stored in the NIS data set for use in constructing nationally representative estimates. When discharge weights are applied to the unweighted NIS data, the result is an estimate of the number of discharges for the entire universe (i.e., an estimate of all acute care hospitalizations in the US).

### A5. Monte Carlo Analysis to Determine Probabilistic Plausible Ranges (PPRs) *(reported previously[4])*

The main outcome measures were estimates of total US annual diagnostic errors and serious misdiagnosis-related harms. Annual incidence from NIS and NAACCR were multiplied by literature-derived estimates of disease-specific and category diagnostic error and harm rates,[4] an approach analogous to "minimal modeling" methods in cost-effectiveness or value-of-information analysis.[11]

To obtain the variability of these combined estimates, we used a probabilistic sampling approach based on Monte Carlo simulations[12] (Supplement 2). These simulations produce statistically valid 95% CIs that account for variability in both number and sample sizes for each disease. In the current manuscript, most of these uncertainty estimates are denoted as "probabilistic plausible ranges" (PPRs), rather than 95% CIs. This is because they rely, in part, on diagnostic error rates that utilize literature-derived (and expert-validated) plausible ranges (n=5 cancers) rather than statistically derived 95% CIs as their uncertainty range, reflecting some uncertainty beyond mere sampling error.[4] Specifically, experts felt that for the top five cancers, PRs should be wider than the statistical 95% CIs. For each cancer, this was because different studies defined diagnostic delays of different lengths—defining shorter delays as errors created an upper PR bound, while defining longer delays created a lower PR bound.[4] As part of the same Monte Carlo simulations, we also calculated PPRs around error and harm point estimates for the "other" (non-top 5) subcategories and combined categories (e.g., top 5 vascular events, total "Big Three," grand totals).

For the Monte Carlo analysis, skew-normal distributions were used to approximate the distributions of disease incidence rate, diagnostic error rate, and serious misdiagnosis-related harm rate, separately for each quantity. The location parameter of the skew-normal was set to be the point estimate of the corresponding rate. The scale and skewness parameters were determined such that the lower and upper bounds of 95% CI of the resulting skew-normal distribution coincided with each of the 95% CI or probabilistic range bounds of the corresponding rate. Due to the extreme skewness of plausible ranges for some diagnostic error rates,[4] all approximations were performed on logit-transformed distributions.[13] Monte Carlo samples were drawn independently from the resulting distributions. The population affected by the diagnostic error and the subsequent misdiagnosis-related harm were calculated for each Monte Carlo replica. The PPRs were given by the 2.5% and 97.5% quantiles based on 10,000,000 simulations. The large number of simulations was used to ensure tail probability and reduce Monte Carlo error due to the very skewed sampling distributions. Rates and other parameters have been published previously,[4] although diagnostic error rate estimates for stroke, myocardial infarction, venous thromboembolism, aortic aneurysm and dissection, and sepsis were updated to reflect the most robust estimates available from a systematic review with meta-analysis that was conducted by members of the authorship team.[5]

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

*BMJ Qual Saf*

## B) Estimated Impact of Methodological Choices & Assumptions on Results

### B1. Sensitivity Analysis of Errors & Harms (Impact of Uncertainty in Model Parameters)

We conducted a one-way sensitivity analysis to assess the impact of parameter uncertainty on the final point estimates (Figure S2). Harm results were most sensitive to parameters for common infections. The three highest leverage parameters for potentially overestimating serious harms were the (1) other infection harm rate, (2) sepsis harm rate, and (3) pneumonia harm rate. However, even if each of these harm rates (and the one for stroke [4th for overestimates]) were placed at the lower plausible bound of harms assessed in this one-way sensitivity analysis, the point estimate of serious harms would still be over 500,000.
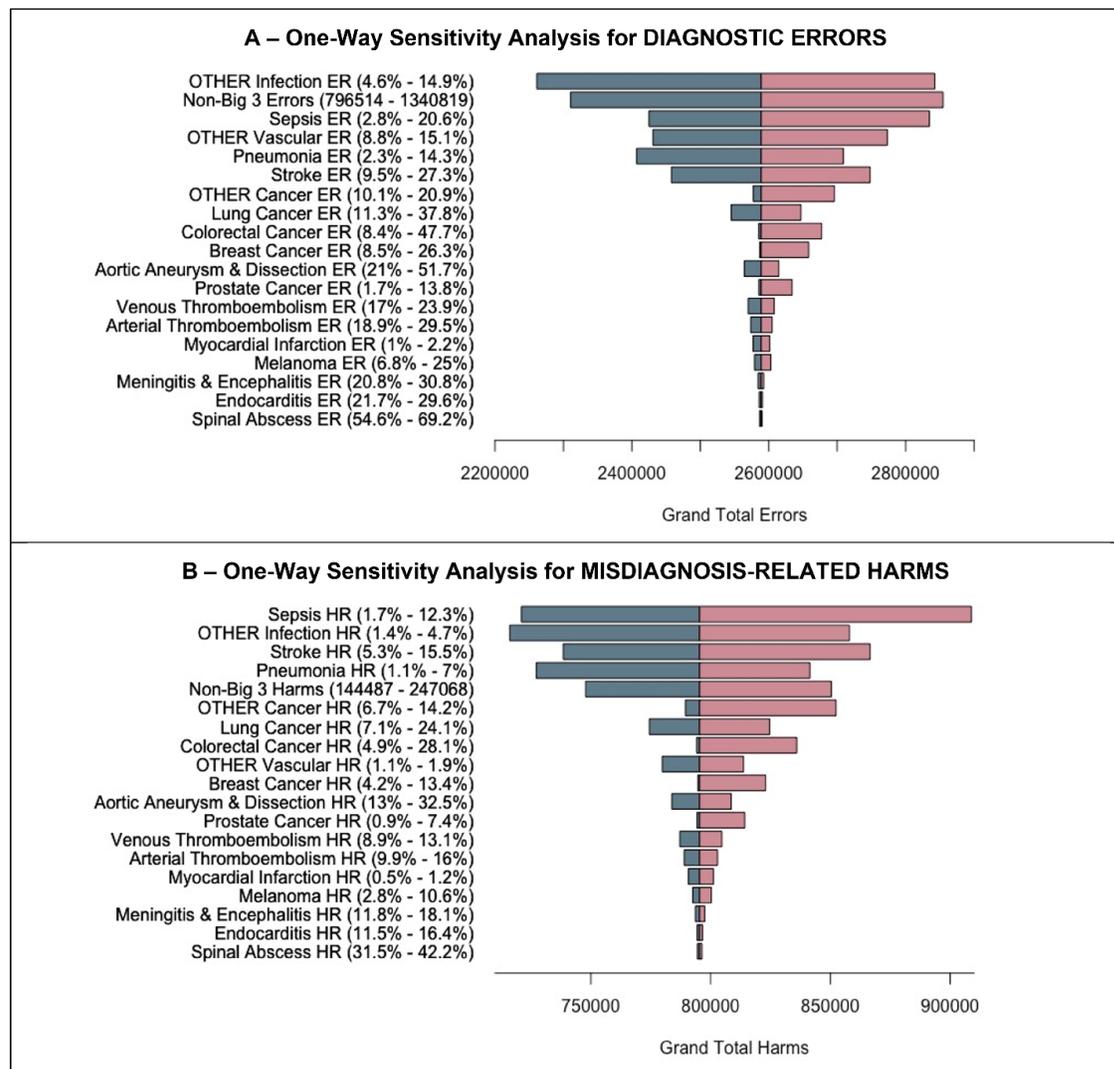


**Figure S2. Sensitivity analysis of model parameters on diagnostic errors and serious harms.** This one-way (as opposed to multi-way, probabilistic sensitivity analysis) varies one parameter value at a time while holding all other model parameter values constant. For each parameter, its value is ranged between the lower and upper uncertainty bounds (i.e., confidence interval or plausible range [*shown*]) and the impact of uncertainty on the total is illustrated, with highest impact parameters shown at the top and lowest at the bottom, yielding a "tornado" appearance. The

point estimate value used for each parameter is fixed as the "midline" of the tornado. The impact on the final total of using the lower bound parameter value is shown in *blue to the left* (reflecting possible *overestimation* in the point estimate). Conversely, the impact on the final total of using the upper bound parameter value is shown in *red to the right* (reflecting possible *underestimation* in the point estimate). Panel A shows a tornado diagram for diagnostic errors. The three parameters with the greatest potential for overestimation of errors were the (1) other infection error rate, (2) non-"Big Three" error rate, and (3) pneumonia error rate. Panel B shows a tornado diagram for serious harms. The three parameters with the greatest potential for overestimation of harms were the (1) other infection harm rate, (2) sepsis harm rate, and (3) pneumonia harm rate. *Abbreviations: ER – error rate; HR – harm rate.*

## B2. Sensitivity Analysis of Errors & Harms (Impact of Using Only Principal Diagnosis)

We estimated disease incidence for vascular events and infections from HCUP data from the NIS. In the primary analysis, we counted discharge (or in-hospital death) diagnoses coded in either the principal or first-listed secondary diagnosis positions, as these two diagnoses are often of equal, competing weight.[14]

We conducted a sensitivity analysis using only first-position NIS diagnosis codes to assess the impact of this methodological decision on the final results. Primary vs. sensitivity results are shown in Table S1. Overall estimates of errors and harms were about 30% lower when using only principal diagnosis codes.

**Table S1. Comparison of diagnostic errors and harms in the primary analysis vs. principal-only analysis\***

| Category | Primary Analysis<br>N in thousands (PPR) | Principal Diagnosis-Only<br>N in thousands (PPR) |
|---|---|---|
| Big Three Total Diagnostic Errors | 1,514 (1,122-1,889) | 1,044 (852-1,365) |
| Big Three Total Serious Harms | 603 (454-776) | 416 (344-550) |
| Grand Total Diagnostic Errors | 2,588 (1,919-3,230) | 1,785 (1,457-2,335) |
| Grand Total Serious Harms | 795 (598-1,023) | 549 (454-725) |

\* The primary analysis counted NIS diagnosis codes in either the principal or first-listed secondary positions. The sensitivity analysis counted NIS diagnosis codes in only the principal position (so are necessarily lower).
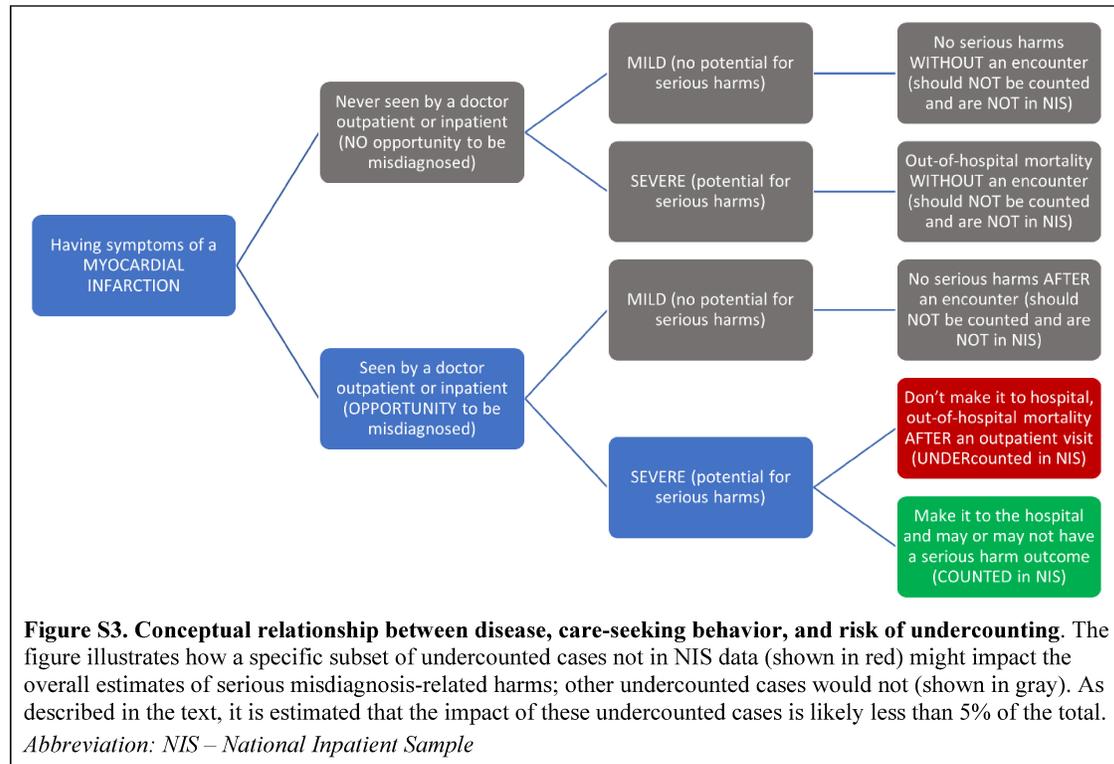
## B3. Estimated Impact of Undercounting (Conservative Assumption re: Out-of-Hospital Deaths)

The conservative assumption was made that incident cases of dangerous (life or limb-threatening) vascular events and infections in the US would eventually involve a hospitalization, even if the patient was initially misdiagnosed in an ambulatory care setting. Outpatient (e.g., primary care, emergency department) visit diagnoses were not included separately in the disease incidence calculations because they would risk inflating disease incidence estimates through double counting. For example, had "myocardial infarction" cases that were correctly diagnosed in outpatient care (and then later confirmed as an inpatient) been included in the analysis, the same incident cases would be counted twice.

This methodological decision, chosen to avoid overcounting, does risk some degree of undercounting; this is principally via patients seen in outpatient settings (e.g., primary care or emergency department) who are misdiagnosed and then die in the community of their underlying illness, without ever reaching the hospital as an inpatient (i.e., out-of-hospital deaths following a missed diagnosis). Reliable data are not available to directly measure out-of-hospital deaths across all conditions (because death certificate data may be unreliable[15]). However, we can estimate the incidence of such deaths using pre-hospital death data, which are available for some of the more common dangerous diseases. For myocardial infarction, the proportion of all true cases that result in pre-hospital death was estimated in a rigorous population-based study from Germany to be 13.6%.[16] For stroke, we can estimate the rate by combining data on pre-hospital stroke deaths with data on stroke hospitalizations. In 2014 (the reference year for our study) there were 980 stroke hospitalizations per 100,000 among Medicare beneficiaries (who are predominantly patients aged 65 and older) in the state of New York.[17] With a 2014 New York state population aged 65

and older of 2,898,094,[18] the total number of stroke hospitalizations for those aged 65 and older in 2014 was 28,401. Since ~38% of stroke hospitalizations in 2014 occurred in patients below age 65,[19] the total number of New York state stroke hospitalizations in 2014 was approximately ~45,800. In 2012-2014, the crude rate of out-of-hospital stroke deaths in New York state was 11.6 per 100,000.[20] With a 2014 New York state total population of ~19,750,000,[18] the total number of pre-hospital stroke deaths was about ~2,300. Therefore, the proportion of strokes that resulted in death without making it to the hospital was about 4.8% of all strokes. The heart attack and stroke estimates cohere well with what is known about the natural history of these two diseases—harms from myocardial infarction are disproportionately deaths, while harms from stroke are disproportionately disabilities. Accordingly, it is sensible that a higher proportion of myocardial infarctions would result in pre-hospital deaths. If we take a disease incidence-weighted average of these two numbers (13.6% for heart attack and 4.8% for stroke), the estimated proportion of undercounted incident cases is ~10%.

If we postulate a similar overall rate (~10% out-of-hospital deaths) for the remaining vascular events and infections, and if 100% of these out-of-hospital death cases were deemed "misdiagnosed," our overall estimate of serious misdiagnosis-related harms would be under-counted by approximately 8% (since this particular problem is not likely to impact incident cancers, which are measured by different means of estimation). However, many such cases involve patients who die without ever having had the opportunity to be misdiagnosed because they never reached medical attention (e.g., previously asymptomatic sudden cardiac death from myocardial infarction[21] or strokes in which premonitory transient ischemic attack symptoms do not prompt the individual to seek attention[22]) (Figure S3). ***Thus, 8% represents a likely upper bound on undercounting and the true value is probably less than 5%.***



**Figure S3. Conceptual relationship between disease, care-seeking behavior, and risk of undercounting**. The figure illustrates how a specific subset of undercounted cases not in NIS data (shown in red) might impact the overall estimates of serious misdiagnosis-related harms; other undercounted cases would not (shown in gray). As described in the text, it is estimated that the impact of these undercounted cases is likely less than 5% of the total.
*Abbreviation: NIS – National Inpatient Sample*

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

*BMJ Qual Saf*

*Burden of Serious Harms from Diagnostic Error in the USA*       *Newman-Toker et al., BMJQS 2023*

### B4. Estimated Impact of Overcounting (Based on Patients with More than One Hospitalization)

There is some possibility that our method of using inpatient hospitalizations to measure dangerous disease incidence might lead to overcounting. NIS data track hospitalizations, not patients, so some patients could have been admitted more than once (e.g., admitted for a myocardial infarction and later a stroke in the same year). Although a single person could suffer permanent disability in more than one way, one patient cannot die twice, so this could theoretically lead to overestimates of deaths using our method.

It is not possible to estimate the impact of such potential overcounting directly using NIS data, but we can estimate the potential extent of the problem by combining NIS with other data sources, such as the National Health Interview Survey (NHIS). Using NHIS, the Centers for Disease Control and Prevention (CDC) reports that, in 2014, ~17.5M (corresponding to ~76% of patients hospitalized at least once that year) were hospitalized only once, ~3.2M (corresponding to ~14% of patients hospitalized at least once that year) were hospitalized twice, and ~2.3M (corresponding to ~10% of patients hospitalized at least once that year) were hospitalized three or more times.[23] With ~35.4M total hospitalizations that year in NIS,[24] that means ~50% of hospitalizations involved "repeat visitors." Using NIS data from our current analysis (average for 2012-2014), 34% of hospitalizations were for vascular events or infections (i.e., 12.1M of 35.8M total hospitalizations). The clinical proportion of serious misdiagnosis-related harms represented by deaths is 46.7%.[25,26] Thus, the potential impact on overcounting deaths from missed vascular events and infections is ~8% (i.e., ~50% x ~34% x ~47% = ~8%). However, patients who did, in fact, die of a misdiagnosis from one of these illnesses could not have been counted again past their death date (i.e., they could not have been a "repeat visitor"), *so the true value is likely to be lower (e.g., <5%).*

## C) Additional Validity Arguments

### C1. Comparison with Independent Estimates from Diagnostic Errors in Hospital Autopsies

To gauge the plausibility of our overall serious misdiagnosis-related harms estimate (~795,000), we can derive the misdiagnosis-associated mortality in our data and compare it to that found in hospital autopsy data. We estimate total deaths from our current study using the previously published proportion of harms representing deaths across inpatient and outpatient settings—46.7%[25,26] (~795,000 x 46.7% = ~371,000). We can do the same for the principal-only analysis (~549,000 x 46.7% = ~256,000). Table S2 shows total expected US deaths and all serious harms (death plus permanent disability) in 2014 (our year of analysis), depending on the hypothesized proportion of deaths associated with diagnostic error.

**Table S2. Anticipated US deaths and serious harms due to diagnostic error depending on hypothesized risk**

| Hypothetical Proportion of Deaths Associated with Diagnostic Error | Potential Misdiagnosis-Related Deaths in the US in 2014* | Corresponding Misdiagnosis-Related Serious Harms in 2014† |
|---|---|---|
| 5% | 131,321 | 281,027 |
| 10% | 262,642 | 562,053 |
| 15% | 393,963 | 843,080 |
| 20% | 525,284 | 1,124,107 |

* According to the Centers for Disease Control and Prevention (CDC), there were 2,626,418 (2.6M) US deaths in 2014,[27] which is the reference year chosen for our manuscript's analysis of serious misdiagnosis-related harms.
† Calculated using the clinical proportion of serious misdiagnosis-related harms represented by deaths (46.7%).[25,26]

The primary analysis estimate (~371,000 deaths [of ~2.6M deaths]) would represent a 14.1% (10.6-18.2) overall misdiagnosis-associated mortality nationally. The principal-only analysis (~256,000 deaths [of ~2.6M deaths]) would represent a 9.8% (8.1% to 12.9%) overall misdiagnosis-associated mortality

*Burden of Serious Harms from Diagnostic Error in the USA*      *Newman-Toker et al., BMJQS 2023*

nationally. Either rate is higher than estimates from studies of hospital autopsies that consider epoch of diagnosis and adjust for bias from submaximal autopsy rates. A large meta-analysis of hospital autopsy studies projected that a modern US hospital which autopsied 100% of in-hospital deaths would find 8.4% (95% CI 5.2-13.1) suffered a major diagnostic error, half considered Class I (deaths directly attributed to the diagnostic error) and half Class II (diagnostic errors that would have changed clinical management and could have altered the patient's clinical course).[28]

However, it is expected that the population-based proportion of misdiagnosis-related deaths would be higher than that found in hospital autopsies. One reason is that hospital autopsies consider diagnostic errors not yet recognized at the time of death, but usually not errors occurring pre-hospitalization when a prompt intervention might have been lifesaving. For example, consider a patient with a new, abrupt-onset headache who is sent home from a primary care clinic as "migraine." If the patient were to return a week later to the emergency department in a coma, they might be promptly diagnosed with aneurysmal subarachnoid hemorrhage and admitted to the intensive care unit. Were they to die, a hospital autopsy would then indicate that <u>no</u> diagnostic error had occurred *in the hospital*. The risk of death from brain aneurysm is increased nearly 5-fold after an initial misdiagnosis, and misdiagnosis disproportionately occurs in outpatient clinics with isolated headache clinical presentations.[29,30] Given highly effective treatments for brain aneurysm and the knowledge that prognosis post-operatively is almost entirely tied to clinical severity at the time of surgery,[29] this case should clearly count as a potentially preventable death due to diagnostic error, but would be considered a <u>correct</u> ante-mortem diagnosis in hospital autopsy data.

Another reason is that the proportion of deaths associated with diagnostic error/delay is probably higher for out-of-hospital than in-hospital deaths. Estimated diagnostic error rates in primary care (2.4% per visit [n=5,126/212,165][31,32]) exceed those in hospitals (0.7%[33]). This makes sense, since (a) hospitalization tends to occur relatively late in the natural course of illness, when a patient has become sick enough to merit inpatient care, and often after the underlying cause for their symptoms is more obvious, and (b) hospitalized patients undergo more intensive diagnostic testing and monitoring than ambulatory patients. As a result, it would not be surprising if outpatient deaths were more often pursuant to diagnostic errors. Since >80% of healthcare visits occur in non-ED ambulatory care[34,35] and >65% of all US deaths occur outside the hospital,[36] an overall misdiagnosis-associated mortality of 9.8% to 14.1% seems plausible.

How many of these misdiagnosis-associated deaths are preventable and how much (or little) longevity might potentially be reclaimed for affected patients is uncertain.[37] Nevertheless, individual cases of otherwise healthy young patients who die from treatable causes that were misdiagnosed make it clear that this could be a half-century or more in years of quality life lost for a given patient.[38-40] The same is true for lifelong disability in young patients after missed opportunities to promptly treat disabling diseases.[41-43] For some of the most harmful diseases in our list, correct initial diagnosis has been associated with clear and substantial reductions in morbidity or mortality (e.g., ischemic stroke [~5-fold],[5] aneurysmal subarachnoid hemorrhage [~5-fold],[29] and ruptured abdominal aortic aneurysm [~2-fold][44]).

## <u>C2. Comparison with Independent Estimates from Diagnostic Adverse Events in Hospitals</u>

We can also gauge the plausibility of our serious harm results in light of diagnostic adverse event data from inpatient hospital stays. Gunderson et al. recently published a systematic review of hospital-based studies of diagnostic adverse events (n=22), two of these US-based.[33] They estimated a pooled hospital misdiagnosis-related harm rate (counting any harm severity) of 0.7% (95% CI 0.5-1.1) with high levels of heterogeneity ($I^2$=95%, p<0.001) (overall range across studies 0.1-2.7).[33] Most of these studies did not report specific diseases missed, but eight did (n=136 cases). Authors listed 70 diseases or categories with at least two instances (Table 2[33]). Among these 70, 78.6% were attributed to "Big Three" diseases (this distribution is very similar to the attributable % used in our current population-based study [75.8%[2]]).

*Burden of Serious Harms from Diagnostic Error in the USA*          *Newman-Toker et al., BMJQS 2023*

If applied to US-based hospitals, they estimated ~250,000 patients harmed annually from diagnostic error. They were unable to assess harm severity based on the available literature. In the well-designed 2010 Dutch study by Zwaan et al.,[25] which measured a similar rate of hospital-based diagnostic adverse events (0.4%), they found that, of diagnostic adverse events, 29.1% resulted in death and ~25.6% (estimated from their Figure 1) resulted in disability at discharge, for an overall rate of ~54.7% serious harms.[25]

Using NIS 2014 US hospitalizations (~35.4M),[24] this translates to ~135,000 (~97K-213K) (uncertainty estimated using 95% CI from Gunderson[33]) serious misdiagnosis-related harms in US hospitals. Relative to our 2014 primary analysis estimate of ~795,000 suffering death or permanent disability, this suggests that ~17% (~12-27) of serious misdiagnosis-related harms occur among inpatients and ~83% (~73-88) among outpatients. Although in 2014 only 2.7% of the roughly 1.3B US healthcare visits were inpatient hospitalizations,[4] severity of illness and diagnostic error adverse events are both higher than outpatient.[45] To help gauge this effect, the proportion of high-severity misdiagnosis-related harms linked to inpatient care in malpractice claims is ~28% (i.e., >10-fold over-representation relative to visit proportion).[2] It is reasonable to expect that inpatient malpractice claims for diagnostic adverse events would be artificially over-represented in claims relative to clinical care proportions, since, relative to outpatient care, outcome severity is higher[45] (a known predictor of legal action[7]) and a "paper trail" of documentation to establish a legal action is more readily available (another likely predictor[2]). Therefore, the ~28% represents an "upper bound," of sorts, on the inpatient-attributable serious harms fraction. Thus, an estimate that ~17% (~12-27) of US serious misdiagnosis-related harms occur in inpatient settings seems quite plausible.

Using these numbers, we can also estimate that the total annual hospital-based deaths from diagnostic error in the US in 2014 would be ~35.4M (NIS 2014[24]) x 0.7% (95% CI 0.5-1.1) (Gunderson et al.[33]) x 29.1% (Zwaan et al.[25]) = ~72,000 (~51K-113K). This value is squarely within the range projected by Leape, Berwick, and Bates (i.e., 40,000-80,000 hospital deaths per year)[46] derived by multiplying total hospital deaths by the rate of hospital autopsy-determined diagnostic errors. The ~72,000 misdiagnosis-related hospital deaths estimate is also squarely in the range of what is expected based on a rigorous systematic review of hospital autopsies by Shojania et al.[28] They calculated the combined Goldman Class I/II diagnostic error rate for an average, modern, US-based hospital that autopsied 100% of its deaths— 8.4% (95% CI 5.2-13.1).[28] According to the CDC, there were 2,626,418 (2.6M) US deaths in 2014,[27] of which 37.3% were hospital-based,[47] for a total of ~980,000 hospital deaths. The ~72,000 would therefore correspond to a 7.4% (n=~72,000/~980,000) misdiagnosis-attributable fraction of hospital deaths.

The resulting estimates comparing inpatient-only harms to those across settings are shown in Table S3.

**Table S3. US deaths and serious harms due to diagnostic error in 2014 comparing inpatient to all settings**

| Misdiagnosis-Related Harms | Inpatient Only (Prior Studies[25,33]) | Across Settings (Current Study) |
|---|---|---|
| Total Serious (Death + Disability) | ~135,000 (~97K-213K)* | ~795,000 (~598K-1,023K)† |
| Deaths Only | ~72,000 (~51K-113K)* | ~371,000 (~279K-478K)‡ |
| Disability (*calculated difference*) | ~63,000 (~46K-100K) | ~424,000 (~319-545) |

* Uncertainty accounted for using 95% CI from Gunderson[33] plus serious harms or death % from Zwaan.[25]
† Uncertainty accounted for using PPR from primary analysis, which used Monte Carlo simulations (see A5).
‡ Uncertainty accounted for using PPR from primary analysis and point death % from Zwaan[25] & Singh[26] combined.

## C3. Triangulation of Available Data across Sources and Methods

We have described three separate methods of estimation that all yield compatible results:

- Method 1 (Manuscript): disease incidence x literature-based misdiagnosis-related harm rate
- Method 2 (Section C1): hospital deaths x % of deaths attributable to diagnostic error
- Method 3 (Section C2): hospital adverse events x % of adverse events resulting in harm or death

*Burden of Serious Harms from Diagnostic Error in the USA*       *Newman-Toker et al., BMJQS 2023*

The three distinct methods can be used to derive inpatient serious harms and misdiagnosis-related deaths. Point estimates for total inpatient serious harms across the three methods range from ~135,000-225,000. All three methods produce point estimates for deaths that fall within the tight range of ~72,000-105,000.

**Method 1** gives us total serious misdiagnosis-related harms (i.e., death + disability) in 2014 for inpatient and outpatient settings (~795,000). From this, we can estimate total deaths using the previously published proportion of harms representing deaths across inpatient and outpatient settings—46.7%[25,26] (~371,000) (see C1). Combining this with the proportion of total serious harms attributed to inpatient settings from a large, nationally representative sample of malpractice claims (28%[2]) gives ~225,000 total serious harms and ~105,000 deaths in US hospitals annually. Because of likely bias towards legal action for inpatient claims (see C2), these are presumed to be slight overestimates. Despite this, they are still close to results estimated by Methods 2 and 3, below. Method 1 serious harms (~225,000) fall within the uncertainty range by Method 2 (~96K-241K) and just beyond that by Method 3 (~97K-213K). Method 1 deaths (~105,000) fall within the uncertainty ranges by Method 2 (~51K-128K) and Method 3 (~51K-113K).

**Method 2** gives us total misdiagnosis-related hospital deaths directly. Hospital deaths in 2014 (~980,000 [see C2]) were published by the CDC and the misdiagnosis-attributable fraction (8.4% [95% CI 5.2-13.1]) is from a rigorous meta-analysis of 53 autopsy studies whose final estimates account for study country, study epoch, and submaximal autopsy rate.[28] The estimate is ~82,000 (~51K-128K) misdiagnosis-related hospital inpatient deaths. Using the previously published proportion of serious harms representing deaths in the inpatient setting (~53.2% [29.1% deaths of ~54.7% serious harms][25]), we can estimate total serious harms of ~155,000 (~96K-241K). These are close to results from Method 3, despite different derivations.

**Method 3** gives us total misdiagnosis-related hospital harms or deaths. US hospitalizations in 2014 (~35.4M[24]) derive from the NIS. The diagnostic adverse event rate (0.7%) is from a meta-analysis of 22 studies of hospital-based diagnostic adverse events,[33] and the proportion of adverse events resulting in serious harms (~54.7%) or death (29.1%) are from a rigorous, population-based sample of inpatient diagnostic adverse events from 21 hospitals. The resulting hospital estimate is ~135,000 (~97K-213K) serious misdiagnosis-related harms, which includes an estimated ~72,000 deaths (~51K-113K) (see C2).

Thus, our current results triangulate well across data sources and methods (convergent construct validity). When the consistency of these estimated misdiagnosis-related harms is combined with the consistent proportion of serious harms (~75-80%[3]) accounted for by "Big Three" diseases across settings (n=44 studies in primary care, emergency department, hospital[3]), this enhances the validity of our study results.

Finally, our results are bolstered by coherence with another recent systematic review and meta-analysis of diagnostic errors in the emergency department,[5] which permits a rough estimate across clinical settings as a final cross-check. As shown in Table S4, data combined from other sources (~855,000 [plausible range ~490K-1,659K]) align well with those found in the current study (~795,000 [PPR 598K-1,023K]). Both values translate to a US per-healthcare-visit serious misdiagnosis-related harm rate of about 0.08%.

**Table S4. Serious misdiagnosis-related harms and serious harms rates by clinical setting (alternate sources)**

| Clinical Setting in which Error Occurs | Annual US Visits per Year in 2014 (n) | Total Serious Misdiagnosis-Related Harms (n) | Estimated Serious Harms Rate (%) |
|---|---|---|---|
| Inpatient | 35,400,000[24] | ~135,000 (~97K-213K)[25,33]* | ~0.38% (~0.27-0.60)* |
| Emergency Department | 137,800,000[48] | ~430,000† (~259K-1,042K)† | ~0.31% (~0.19-0.76)[5] |
| Primary Care Clinics | 461,800,000[49] | ~206,000† (~103K-309K)‡ | ~0.04%§ (~0.02-0.07)‡ |
| Specialty Care Clinics | 423,000,000[49] | ~85,000† (~42K-127K)‡ | ~0.02%§ (~0.01-0.03)‡ |
| TOTAL (£) | 1,057,900,000 | ~855,000 (~490K-1,659K) | ~0.08% (~0.05-0.16) |

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance
placed on this supplemental material which has been supplied by the author(s)

*BMJ Qual Saf*

*Burden of Serious Harms from Diagnostic Error in the USA*      *Newman-Toker et al., BMJQS 2023*

\* Point estimate and uncertainty combine US inpatient harms estimate from Gunderson[33] with serious harms
proportion from Zwaan[25] as described in Section C3 above. Harms rate is then calculated using visits per year.

† Total serious misdiagnosis-related harms calculated as serious harms rate x annual US visits.

‡ When precise estimates of uncertainty were lacking, we arbitrarily assigned it as +/- 50% of the point estimate.

§ Although the rate of serious harms from diagnostic error in ambulatory clinic-based care is not well characterized,
generally the risks of a serious harm event (on a per visit basis) are much lower than either inpatient or emergency
department care, simply because the severity of illness is much lower. From data in patients with missed stroke
(erroneously called "benign" dizziness), we can estimate that the risks of serious harms after discharge from primary
care are approximately 7-fold lower than those seen in the emergency department[50]; we can also approximate that
the rate of serious harms after discharge from specialty care is likely about half that seen in primary care.[51]

£ Totals were calculated by summing n's for visits and harms, then dividing harms by visits to get rates (weighted
average). As a result, the lower and upper uncertainty bounds are wider than if they were sampled probabilistically.

## D) Supplementary References *(N.B. – citation numbers differ from main manuscript)*

1.  Shojania KG, Dixon-Woods M. Estimating deaths due to medical error: the ongoing controversy
    and why it matters. *BMJ Qual Saf.* 2017;26(5):423-428.

2.  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(3):227-240.

3.  Newman-Toker DE, Schaffer AC, Yu-Moe CW, et al. Corrigendum to: Serious misdiagnosis-
    related harms in malpractice claims: The "Big Three" - vascular events, infections, and cancers.
    *Diagnosis (Berl).* 2020;8(1):127-128.

4.  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).* 2020;8(1):67-84.

5.  Newman-Toker DE, Peterson SM, Badihian S, Hassoon A, Nassery N, Parizadeh D, Wilson LM,
    Jia Y, Omron R, Tharmarajah S, Guerin L, Bastani PB, Fracica EA, Kotwal S, Robinson KA.
    *Diagnostic Errors in the Emergency Department: A Systematic Review.* Comparative
    Effectiveness Review No. 258. (Prepared by the Johns Hopkins University Evidence-based
    Practice Center under Contract No. 75Q80120D00003.) AHRQ Publication No. 22(23)-EHC043.
    Rockville, MD: Agency for Healthcare Research and Quality; December 2022 (addendum in
    press, anticipated August 2023). DOI: https://doi.org/10.23970/AHRQEPCCER258.

6.  Localio AR, Lawthers AG, Brennan TA, et al. Relation between malpractice claims and adverse
    events due to negligence. Results of the Harvard Medical Practice Study III. *N Engl J Med.*
    1991;325(4):245-251.

7.  Studdert DM, Mello MM, Gawande AA, et al. Claims, errors, and compensation payments in
    medical malpractice litigation. *N Engl J Med.* 2006;354(19):2024-2033.

8.  Producing National HCUP Estimates - Accessible Version. Agency for Healthcare Research and
    Quality, 2015. (Accessed May 20, 2023 at https://www.hcup-
    us.ahrq.gov/tech_assist/nationalestimates/508_course/508course_2018.jsp.)

9.  National Trends in Inpatient Stays. Healthcare Cost and Utilization Project (HCUP) Fastats.
    Agency for Healthcare Research & Quality, 2023. (Accessed May 20, 2023 at
    https://datatools.ahrq.gov/hcup-fast-stats; https://datatools.ahrq.gov/sites/default/files/2022-10/
    DownloadTable_NationalTrendsInpatientStays_10-3-22.xls.)

10. Fast Facts on U.S. Hospitals, 2022. American Hospital Association. (Accessed May 20, 2023 at
    https://www.aha.org/statistics/fast-facts-us-hospitals.)

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

*BMJ Qual Saf*

*Burden of Serious Harms from Diagnostic Error in the USA*      *Newman-Toker et al., BMJQS 2023*

11.    Meltzer DO, Hoomans T, Chung JW, Basu A. Minimal modeling approaches to value of information analysis for health research. *Med Decis Making.* 2011;31(6):E1-E22.

12.    Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika.* 1970;57(1):97–109.

13.    Bartlett MS. The use of transformations. *Biometrics.* 1947;3(1):39-52.

14.    ICD-10-CM Official Guidelines for Coding and Reporting FY 2019. Centers for Medicare & Medicaid Services, 2019. (Accessed May 20, 2023 at https://www.cms.gov/Medicare/Coding/ICD10/Downloads/2019-ICD10-Coding-Guidelines-.pdf.)

15.    McCormick N, Lacaille D, Bhole V, Avina-Zubieta JA. Validity of myocardial infarction diagnoses in administrative databases: a systematic review. *PLoS One.* 2014;9(3):e92286.

16.    Amann U, Heier M, Thilo C, Linseisen J, Meisinger C. Determinants of prehospital coronary heart disease death. *Sci Rep.* 2021;11(1):17134.

17.    Interactive Atlas of Heart Disease and Stroke. Centers for Disease Control and Prevention. (Accessed May 20, 2023 at https://nccd.cdc.gov/DHDSPAtlas/reports.aspx?geographyType=county&state=NY&themeId=20&filterIds=1,5,3,6,7&filterOptions=6,1,1,1,1#report.)

18.    Annual Report of Vital Statistics: New York State 2014. New York State Department of Health. (Accessed May 20, 2023 at https://www.health.ny.gov/statistics/vital_statistics/docs/vital_statistics_annual_report_2014.pdf.)

19.    Stroke Facts. Centers for Disease Control and Prevention. (Accessed May 20, 2023 at https://www.cdc.gov/stroke/facts.htm.)

20.    Wales K. Pre-Transport Stroke Death Rates by EMS Region and Location of Stroke Designated Hospitals, New York State. Centers for Disease Control and Prevention. (Accessed May 20, 2023 at https://www.cdc.gov/dhdsp/maps/gisx/mapgallery/NY-stroke.html.)

21.    O'Neill L. Estimating out-of-hospital mortality due to myocardial infarction. *Health Care Manag Sci.* 2003;6(3):147-154.

22.    Giles MF, Flossman E, Rothwell PM. Patient behavior immediately after transient ischemic attack according to clinical characteristics, perception of the event, and predicted risk of stroke. *Stroke.* 2006;37(5):1254-1260.

23.    Summary Health Statistics: National Health Interview Survey, 2014. Centers for Disease Control and Prevention, National Center for Health Statistics. (Accessed May 20, 2023 at https://ftp.cdc.gov/pub/Health_Statistics/NCHS/NHIS/SHS/2014_SHS_Table_P-10.pdf.)

24.    Trends in Hospital Inpatient Stays in the United States, 2005-2014 (HCUP Statistical Brief #225). Agency for Healthcare Research & Quality. (Accessed May 20, 2023 at https://www.hcup-us.ahrq.gov/reports/statbriefs/sb225-Inpatient-US-Stays-Trends.jsp.)

25.    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(12):1015-1021.

26.    Singh H, Giardina TD, Meyer AN, Forjuoh SN, Reis MD, Thomas EJ. Types and origins of diagnostic errors in primary care settings. *JAMA Intern Med.* 2013;173(6):418-425.

27.    Deaths: Final Data for 2014. Centers for Disease Control and Prevention, National Center for Health Statistics. (Accessed May 20, 2023 at https://www.cdc.gov/nchs/data/dvs/LCWK1_2014.pdf.)

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

*BMJ Qual Saf*

*Burden of Serious Harms from Diagnostic Error in the USA*          *Newman-Toker et al., BMJQS 2023*

28.    Shojania KG, Burton EC, McDonald KM, Goldman L. Changes in rates of autopsy-detected diagnostic errors over time: a systematic review. *JAMA.* 2003;289(21):2849-2856.

29.    Kowalski RG, Claassen J, Kreiter KT, et al. Initial misdiagnosis and outcome after subarachnoid hemorrhage. *JAMA.* 2004;291(7):866-869.

30.    Tarnutzer AA, Lee SH, Robinson KA, Wang Z, Edlow JA, Newman-Toker DE. ED misdiagnosis of cerebrovascular events in the era of modern neuroimaging: A meta-analysis. *Neurology.* 2017;88(15):1468-1477.

31.    Singh H, Giardina TD, Forjuoh SN, et al. Electronic health record-based surveillance of diagnostic errors in primary care. *BMJ Qual Saf.* 2012;21(2):93-100.

32.    Singh H, Meyer AN, 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(9):727-731.

33.    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(12):1008-1018.

34.    Ambulatory Care Use and Physician Office Visits. Centers for Disease Control and Prevention, National Center for Health Statistics. (Accessed May 20, 2023 at http://www.cdc.gov/nchs/fastats/physician-visits.htm.)

35.    Emergency Department Visits. Centers for Disease Control and Prevention, National Center for Health Statistics. (Accessed May 20, 2023 at http://www.cdc.gov/nchs/fastats/emergency-department.htm.)

36.    Trends in Inpatient Hospital Deaths: National Hospital Discharge Survey, 2000–2010. Centers for Disease Control and Prevention, National Center for Health Statistics. (Accessed May 20, 2023 at https://www.cdc.gov/nchs/products/databriefs/db118.htm.)

37.    Hayward RA, Hofer TP. Estimating hospital deaths due to medical errors: preventability is in the eye of the reviewer. *JAMA.* 2001;286(4):415-420.

38.    From a Small Scrape to Sepsis - Rory Staunton's Story. Society to Improve Diagnosis in Medicine. (Accessed May 20, 2023 at https://www.improvediagnosis.org/stories_posts/from-a-small-scrape-to-sepsis/.)

39.    Unbelieved - Jessica Barnett's Story. Society to Improve Diagnosis in Medicine. (Accessed May 20, 2023 at https://www.improvediagnosis.org/stories_posts/unbelieved-jessica-barnetts-story/.)

40.    Tested Too Late - Chad Becken's Story. Society to Improve Diagnosis in Medicine. (Accessed May 20, 2023 at https://www.improvediagnosis.org/stories_posts/tested-too-late-chad-beckens-story/.)

41.    Missed Stroke Diagnosis - John Michael Night's Story. Society to Improve Diagnosis in Medicine. (Accessed May 20, 2023 at https://www.improvediagnosis.org/stories_posts/missed-stroke-diagnosis/.)

42.    The Missed Bilirubin Test - Cal Sheridan's Story. Society to Improve Diagnosis in Medicine. (Accessed May 20, 2023 at https://www.improvediagnosis.org/stories_posts/the-missed-bilirubin-test-cal-sheridans-story/.)

43.    The Lingering Lump - Diana Cejas' Story. Society to Improve Diagnosis in Medicine. (Accessed May 20, 2023 at https://www.improvediagnosis.org/stories_posts/the-lingering-lump-diana-cejas-story/.)

44.     Smidfelt K, Nordanstig J, Davidsson A, Torngren K, Langenskiold M. Misdiagnosis of ruptured abdominal aortic aneurysms is common and is associated with increased mortality. *J Vasc Surg.* 2021;73(2):476-483 e473.

45.     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(8):672-680.

46.     Leape LL, Berwick DM, Bates DW. Counting deaths due to medical errors (in reply). *JAMA.* 2002;288(19):2404-2405.

47.     QuickStats: Percentage Distribution of Deaths, by Place of Death — United States, 2000–2014. *Morbidity and Mortality Weekly Report.* 2016;65:357. (Accessed May 20, 2023 at https://www.cdc.gov/mmwr/volumes/65/wr/mm6513a6.htm.)

48.     Hooker EA, Mallow PJ, Oglesby MM. Characteristics and Trends of Emergency Department Visits in the United States (2010-2014). *J Emerg Med.* 2019;56(3):344-351.

49.     National Ambulatory Medical Care Survey: 2014 State and National Summary Tables. Centers for Disease Control and Prevention, National Center for Health Statistics. (Accessed May 20, 2023 at https://www.cdc.gov/nchs/data/ahcd/namcs_summary/2014_namcs_web_tables.pdf.)

50.     Nassery N, Mane K, Liu F, et al. Measuring missed strokes using administrative and claims data: towards a diagnostic performance dashboard to monitor diagnostic errors. Diagnostic Error in Medicine 2016; November 6-8, 2016, 2016; Hollywood, CA.

51.     Chang TP, Bery AK, Wang Z, et al. Stroke hospitalization after misdiagnosis of "benign dizziness" is lower in specialty care than general practice: a population-based cohort analysis of missed stroke using SPADE methods. *Diagnosis (Berl).* 2021;9(1):96-106.