David E. Newman-Toker*, Adam C. Schaffer, C. Winnie Yu-Moe, Najlla Nassery, Ali S. Saber Tehrani, Gwendolyn D. Clemens, Zheyu Wang, Yuxin Zhu, Mehdi Fanai and Dana Siegal*

Serious misdiagnosis-related harms in malpractice claims: The "Big Three" – vascular events, infections, and cancers

Abstract

Background: Diagnostic errors cause substantial preventable harm, but national estimates vary widely from 40,000 to 4 million annually. This cross-sectional analysis of a large medical malpractice claims database was the first phase of a three-phase project to estimate the US burden of serious misdiagnosis-related harms. **Methods:** We sought to identify diseases accounting for the majority of serious misdiagnosis-related harms (morbidity/mortality). Diagnostic error cases were identified from Controlled Risk Insurance Company (CRICO)'s Comparative Benchmarking System (CBS) database (2006–2015), representing 28.7% of all US malpractice claims. Diseases were grouped according to the Agency for Healthcare Research and Quality (AHRQ) Clinical Classifications Software (CCS) that aggregates the International Classification of Diseases diagnostic codes into clinically sensible groupings. We analyzed vascular events, infections, and cancers (the "Big Three"), including frequency, severity, and settings. High-severity (serious) harms were defined by scores of 6–9 (serious, permanent disability, or death) on the National Association of Insurance Commissioners (NAIC) Severity of Injury Scale. **Results:** From 55,377 closed claims, we analyzed 11,592 diagnostic error cases [median age 49, interquartile range (IQR) 36–60; 51.7% female]. These included 7379 with high-severity harms (53.0% death). The Big Three diseases accounted for 74.1% of high-severity cases (vascular events 22.8%, infections 13.5%, and cancers 37.8%). In aggregate, the top five from each category (n = 15 diseases) accounted for 47.1% of high-severity cases. The most frequent disease in each category, respectively, was stroke, sepsis, and lung cancer. Causes were disproportionately clinical judgment factors (85.7%) across categories (range 82.0–88.8%). **Conclusions:** The Big Three diseases account for about three-fourths of serious misdiagnosis-related harms. Initial efforts to improve diagnosis should focus on vascular events, infections, and cancers.

Keywords: diagnosis; diagnostic errors; health services research; malpractice; medical errors.

Introduction

Diagnostic error is recognized as a major source of preventable harms in US healthcare, but current estimates of

*Corresponding authors: David E. Newman-Toker, MD, PhD, Professor, Department of Neurology, The Johns Hopkins University School of Medicine, The Johns Hopkins Hospital, Meyer Building 2-221, 600 North Wolfe Street, Baltimore, MD 21287, USA; Director of the Armstrong Institute Center for Diagnostic Excellence, Johns Hopkins University School of Medicine, Baltimore, MD, USA; and Professor, Department of Epidemiology, The Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA, Phone: (443) 287-9593, Fax: (410) 614-1746, E-mail: toker@jhu.edu; and Dana Siegal, RN, CPHRM, CPPS, Director of Patient Safety, CRICO Strategies, 1325 Boylston Street, Boston, MA 02215, USA, Phone: (617) 450-5552, Fax: (617) 679-1237, E-mail: DSiegal@rmf.harvard.edu
Adam C. Schaffer: Sr. Clinical Analytics Specialist, Patient Safety-Advanced Analytics and Coding, CRICO, Boston, MA, USA; and Assistant Professor of Medicine, Part-time, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA, E-mail: ASchaffer@rmf.harvard.edu

C. Winnie Yu-Moe: Data Analytics Manager, Patient Safety-Advanced Analytics and Coding, CRICO, Boston, MA, USA, E-mail: WYu@rmf.harvard.edu
Najlla Nassery: Department of Medicine, The Johns Hopkins University School of Medicine, Baltimore, MD, USA, E-mail: nnasser3@jhmi.edu
Ali S. Saber Tehrani and Mehdi Fanai: Department of Neurology, The Johns Hopkins University School of Medicine, Baltimore, MD, USA, E-mail: ali.tehrani@jhmi.edu (A. S. Saber Tehrani); mehdifanai@jhmi.edu (M. Fanai)
Gwendolyn D. Clemens: Department of Biostatistics, The Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA, E-mail: gclemen1@jhu.edu
Zheyu Wang and Yuxin Zhu: Department of Oncology, The Johns Hopkins University School of Medicine, Baltimore, MD, USA; and Department of Biostatistics, The Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA, E-mail: wangzy@jhu.edu (Z. Wang); daisy@jhu.edu (Y. Zhu)

aggregate misdiagnosis-related harms vary widely. Estimates combining autopsy-detected error rates [1] and total hospital deaths suggest perhaps 40,000–80,000 misdiagnosis-related deaths in US hospitals annually [2]. Estimates from national malpractice data suggest that serious morbidity is at least as common as death, translating to roughly 80,000–160,000 serious misdiagnosis-related harms each year [3]. Estimates extrapolating from diagnostic error rates in specific research studies [4] suggest that 12 million Americans suffer a diagnostic error each year in primary care alone [5]. The same studies found 33% of these diagnostic errors resulted in "serious permanent damage" or "immediate or inevitable death [4]." This would translate to at least 4 million seriously harmed, including at least 1.7 million who died from diagnostic error. If correct, then >60% of the 2.7 million deaths annually in the US [6] would be attributable to diagnostic error, which seems implausible, given that previous estimates of attributable deaths from multiple systematic reviews of autopsy studies indicate that the proportion is likely closer to 5–10% [1, 7, 8].

Given the wide range in current estimates for serious misdiagnosis-related harms (40,000 to 4 million individuals per year in the US), a new approach is warranted. Across practice settings, missed vascular events, infections, and cancers (sometimes collectively referred to as the "Big Three" [9]) account for most of the morbidity and mortality attributable to diagnostic errors [7, 8, 10, 11]. As a first step toward a US national epidemiologic estimate of serious misdiagnosis-related harms, we sought to analyze malpractice claims for the most common vascular events, infections, and cancers from a large medical malpractice claims database with closed claims data from institutions across the US. Our main goal was to identify the list of top diseases that, when missed, cause serious harms so that in later research steps we could measure their annual incidence, frequency of diagnostic errors, and risk of harm to approximate incident harms.

Materials and methods

Overall research concept

The overall goal of this three-phase research project was to construct a US national estimate of serious misdiagnosis-related harms (i.e. permanent disability or death). Each phase was designed to answer a key question from a specific data source that would support the final estimate: (1) what diseases account for the majority of serious harms? (using malpractice claims data); (2) how frequent are diagnostic errors causing harm among these diseases? (using medical literature-derived estimates from disease-specific studies); and (3) what is the overall epidemiologic incidence of diagnostic errors and harms among these diseases? (using nationally representative databases). The answer to the first question is presented here, and the answers to the other two questions will be presented elsewhere.

Current study design

This was a cross-sectional analysis of a large medical malpractice claims database. We analyzed closed claims from diagnostic error cases and focused our analysis on identifying and categorizing the principal diseases likely to account for the majority of serious misdiagnosis-related harms. Based on prior studies of malpractice [10, 11] and autopsy data [7, 8, 12, 13], we specifically targeted three major disease categories believed to account for the majority of serious morbidity and mortality attributable to diagnostic errors: (1) vascular events, (2) infections, and (3) cancers, here collectively referred to as the "Big Three" [9]. We then measured the frequency of claims for specific diseases in each Big Three category.

Closed malpractice claims (2006–2015) from patients of all ages were analyzed, grouping diagnoses into vascular events, infections, cancers, or other, using the Agency for Healthcare Research and Quality (AHRQ) Clinical Classifications Software (CCS) [14]. We used CCS to aggregate the International Classification of Diseases 9th Revision, Clinical Modification (ICD-9-CM) diagnostic codes from individual claims into clinically sensible disease groupings that could be used for both category-level and disease-level analyses. A 10-year window was chosen to maximize the sample size for identifying uncommon but important disease conditions in the claims data; the most recent 3 years (2016–2018) of data were not used because a large fraction of recent claims remained open, so data were incomplete.

Definitions

We used published definitions for diagnostic error and harm. The National Academy of Medicine (NAM) defines a *diagnostic error* as 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 [15]. We chose the NAM definition because of its official status vis-à-vis US medical policymaking and because it is effectively equivalent to prior definitions of diagnostic error used by key leaders in the field of improving diagnosis [16–18]. *Misdiagnosis-related harm* is harm resulting from the delay or failure to treat a condition actually present, when the working diagnosis was wrong or unknown [delayed or missed diagnosis (false negative)], or from treatment provided for a condition not actually present [wrong diagnosis (false positive)] [17, 18].

Harm severity was defined according to the National Association of Insurance Commissioners (NAIC) Severity of Injury Scale, a recognized industry standard [19, 20]. NAIC severity codes are organized on a nine-point scale ranging from 0 (legal issue only) to 9 (death) (Supplementary material A1, Box S1). We defined *high-severity harm* according to a standard low-medium-high schema that aggregates NAIC codes 6–9 representing permanent, serious morbidity with mortality. Specifically, this includes as high severity the following NAIC scores: 6 – permanent, significant injury (e.g. deafness, loss of single limb, loss of eye, or loss of one kidney or lung);

7 – permanent, major injury (e.g. paraplegia, blindness, loss of two limbs, or brain damage); 8 – permanent, grave injury (e.g. quadriplegia, severe brain damage, lifelong care, or fatal prognosis); and 9 – death, including fetal/neonatal death when the mother suffers lesser direct harm. When discussing these findings, we use the term "serious misdiagnosis-related harms" synonymously.

Source data

Malpractice claims data derive from the Controlled Risk Insurance Company, Ltd. (CRICO) Comparative Benchmarking System (CBS) [21]. This is a prospectively curated database of well-characterized medical malpractice claims from captive and commercial malpractice insurers across the US. Although it would have been preferable to use a national data repository of claims [i.e. the National Practitioner Data Bank (NPDB)] to support our overall goal of developing a US national estimate of misdiagnosis-related harms, public-use NPDB data do not provide claim-level data about diseases [22], which are required for this analysis. As an alternative to the NPDB, the CRICO CBS database [21] currently holds over 400,000 cases from 165,000 insured physicians and over 400 hospitals (including more than 30 academic/teaching hospitals), and now represents ~35% of all US malpractice claims. The CRICO CBS data analyzed here represent 28.7% of all NPDB claims during the study period. We conducted confirmatory analyses to demonstrate that CRICO data are similar to the broader pool of national malpractice claims from the NPDB (Supplementary material A2, Box S2, Tables S1 and S2, Figures S1–S3).

Though a malpractice case may have multiple defendants (and thus multiple claims), the clinical events in the CBS database are reviewed, categorized, and reported at the case level, and are defined by the individual event (or series of events) and/or patient outcome(s) that triggered the claim(s). For this reason, we present results in this paper using the term "cases" rather than "claims." Closed claims are either "paid" (i.e. non-zero indemnity payment compensating harms) or "unpaid" ($0 indemnity payment). Cases with at least one paid claim include indemnity payments that may have occurred through settlement, arbitration, or legal judgment. Cases with only unpaid claims may still have an expenses payment associated with the claim(s) (i.e. case management cost, including legal fees, which can be substantial, especially for cases going to trial and ending in a defense verdict).

Relevant factors in each case are abstracted based on a complete review of the medical and legal case file including case summaries, medical record data, depositions, and legal proceedings. Cases are reviewed and coded by experienced clinical taxonomy specialists (typically registered nurses with at least 10 years of quality or risk management experience), who abstract data using a multi-tiered coding taxonomy. At the highest level of abstraction, the CRICO clinical taxonomy clusters cases by roughly a dozen case-type (major allegation) codes representing the clinical "essence" of the case (e.g. diagnosis-related, surgical treatment-related, obstetrical treatment-related, and medication-related). At more granular levels of the taxonomy, data elements include specific diagnoses, harm severity, contributing factors, clinical setting, and primary responsible service, each defined in a coding manual with case exemplars to maximize coding consistency. *Primary responsible service* is defined as "the clinical service of the provider determined to be the most responsible for the patient's care at the time of the event."

Clinical coding is peer reviewed and audited on an ongoing basis. Coding is typically performed by a single clinical taxonomy specialist after extensive training and using detailed procedural guidelines developed to standardize the methodology. Coders meet bi-weekly to review auditing feedback, discuss difficult coding case scenarios, and ensure consistent application of the coding guidelines, algorithms, and protocols. The taxonomy and coding processes are overseen by a Taxonomy Governance Committee that manages data integrity via algorithms, auditing, education, and ongoing review. All adjustments or updates to the taxonomy rubric (to meet evolving patient safety needs) are managed in collaboration with analytics leadership to ensure consistency and integrity of the historical and future data.

Diagnosis-related allegations are coded in a manner consistent with the NAM definition used in this study, including considering failed communication with the patient (e.g. in cancer cases). Misdiagnosis-related harms are also coded in a manner consistent with the definition used in this paper, including harms resulting from both omission and commission. Cases include the initial misdiagnosis (or, in the absence of such, the presenting symptoms and signs) as well as the final diagnosis and clinical severity of the outcome using the NAIC scale. The case reviewer captures the decision-making and care processes, noting voids, actions, or inactions that delayed the diagnostic process. Coders review the case for specific contributing factors including clinical judgment-related factors (e.g. failure to order an appropriate test or consultation), communication-related factors (e.g. failure to communicate with other providers or patient), and clinical systems factors (e.g. failure to follow up new findings on test results). Coders then write a detailed case summary narrative explaining the entire event, highlighting the risk issues, near misses and errors, and including the circumstances surrounding each of the identified contributing factors.

Throughout the course of our study, CRICO-based members of the research team who are clinical providers and claims content experts (Dana Siegal, RN and Adam C. Schaffer, MD) had ready access to all case summaries prepared by trained nurse coders. In the current study, the analysis used an existing database to identify previously coded case features. However, direct reviews of case summaries were undertaken as needed, particularly for confirmation of appropriate CCS disease grouping classifications. During his review of these case summaries for CCS groupings, Dr. Schaffer found no CBS-misclassified diagnostic error cases (i.e. that would not be considered NAM-defined diagnostic errors).

List of conditions and category groupings using the Clinical Classifications Software (CCS)

CCS is a standardized grouping of ICD diagnosis codes into a clinically sensible, comprehensive, hierarchical, multi-level classification of illnesses [14]. The CCS classification schema is sometimes referred to eponymously as the "Elixhauser" coding system. Final diagnoses for all closed claims in the CRICO CBS database are first coded according to ICD-9-CM by clinical taxonomy specialists. These ICD codes are then grouped using AHRQ's Healthcare Cost and Utilization Project (HCUP) CCS multi-level coding schema from 2015 [14], which is the most recent ICD-9-CM version issued by AHRQ. This version will probably not be updated further as a transition has been made to ICD-10-CM coding in 2015.

Established CCS Level 1 codes were used to define the Big Three categories – "Diseases of the Circulatory System" (vascular events), "Infectious and Parasitic Disease" (infections), and "Neoplasms" (cancers). All other CCS Level 1 codes were grouped together as "Other." For each of the Big Three categories, Level 3 code groupings were then used to rank diseases by the total number of claims. For this study, the standard CCS schema was adapted to match the study needs (Supplementary material A3, Table S3). The principal change was to re-group organ system-specific infections (e.g. meningitis, endocarditis, and appendicitis) with the "infections" category (in the standard CCS, they are placed with the specific organ system category). In addition, some CCS Level 3 codes were lumped or split in order to place "like with like" with respect to diagnosis. Clinically related diseases were grouped together (e.g. angina with myocardial infarction; transient ischemic attack with stroke; deep vein thrombosis with pulmonary embolus; septicemia with septic shock; rectal cancer with colon cancer). Diseases with highly variable clinical phenotypes were split [e.g. tuberculosis (Tb), where specific clinical manifestations of Tb were reclassified with similar infections – Tb pneumonia with other pneumonias, Tb meningitis with other meningitides, and Tb spinal abscess with other spinal abscesses]. These disease-level groupings were used for all subsequent analyses.

The disease-level groupings were determined iteratively by a team that included subject matter (diagnostic errors and medical malpractice claims) and methods experts (clinical research methods, large data set analysis, and biostatistics). When it was ambiguous whether code groups should be lumped together or split apart, a granular clinical review of malpractice case summaries was followed by consensus decisions. The final schema was constructed so that each malpractice case was counted once (and only once).

Statistical analysis and reporting

We performed descriptive analyses to measure case frequency, patient demographics, harm severity, contributing factors, clinical settings, and primary responsible service. Payments (total, mean, and median) were analyzed by category and disease for cases in which an indemnity payment was made (i.e. paid claims). We examined differences across the Big Three categories and the top five diseases in each category, grouping remaining diseases as "other." We calculated the death-to-disability ratio for each disease in order to determine, from an epidemiologic and public health perspective, the extent to which focusing solely on deaths from diagnostic error (e.g. via autopsy) might bias future disease-specific measurement of serious harms from diagnostic error. We also assessed for any relevant temporal trends in the malpractice data during the 10-year study period to assess any changes over time that might impact our findings. Sample sizes, totals, proportions, weighted averages, medians, 95% confidence intervals (CIs), and odds ratios (ORs) were used to describe the populations and outcomes, as appropriate. CIs, ORs, and chi square (χ^2)-tests for trend were conducted using R v3.4.4 (Vienna, Austria). This paper follows EQUATOR (STROBE) [23] reporting guidelines for observational studies.

Institutional Review Board (IRB) approval

No human subjects were involved in this study, and no IRB approval was required.

Results

Key results are described below and illustrated in Tables 1 and 2 and Figures 1–5; supporting data are provided in Supplementary material B1, Tables S4–S8. During the 10-year study window, there were 55,377 cases with closed malpractice claims, of which 21,743 involved high-severity harms (NAIC 6–9). We identified 11,592 diagnostic error cases, 7379 with high-severity harms (Table 1). Diagnostic errors represented (a) 21% (n = 11,592/55,377) of *all claims* cases [rank #3 after surgical treatment (28%) and medical treatment (23%)]; (b) 34% (n = 7379/21,743) of *high-severity claims* cases (rank #1); and (c) 28% ($2.65B of $9.39B) of *total payouts* (rank #1). Among all-severity diagnostic error cases (n = 11,592), harm severity was low (NAIC 0–2) in 4.0% (n = 459); medium (NAIC 3–5) in 32.4% (n = 3754); high/no death (NAIC 6–8) in 29.9% (n = 3467); and high/death (NAIC 9) in 33.7% (n = 3912). Harm severity distributions varied across the Big Three categories and individual diseases (Figure 1; Supplementary material B1, Table S4).

Among all-severity diagnostic error cases (n = 11,592), patients were 51.7% female with a median age of 49 [interquartile range (IQR) 36–60, range 0–99]. Among high-severity diagnostic error cases (n = 7379),

Table 1: All severity diagnostic error cases with claim details by Big Three disease category (n = 11,592).

Claim-related variable	Vascular	Infection	Cancer	Other	Aggregate[a]
Total closed claims cases, n	2019	1660	3470	4443	11,592
High-severity harms, n (% of cases)	1684 (83.4%)	992 (59.8%)	2793 (80.5%)	1910 (43.0%)	7379 (63.7%)
Disability, n (% of high-severity)	492 (29.2%)	390 (39.3%)	1643 (58.8%)	942 (49.3%)	3467 (47.0%)
Death, n (% of high-severity)	1192 (70.8%)	602 (60.7%)	1150 (41.2%)	968 (50.7%)	3912 (53.0%)
Total payouts $	$546,617,123	$458,624,368	$776,251,670	$864,374,671	$2,645,867,832
Median payout $	$73,494	$41,801	$42,047	$27,153	$38,676
Mean payout (all) $	$270,737	$276,280	$223,704	$194,548	$228,249

[a]Aggregate column contains sum, median, or weighted mean, as appropriate.

Table 2: Payments in high-severity harm malpractice claims by Big Three category and disease (n = 7379).

Big Three category and disease	Cases, n	Cases with paid[a] Claims, n (%[b])	Total payments, $	Mean per-claim payout (all), $	Mean per-claim payout (paid[a] only), $	Rank[c]
Vascular						
Stroke	424	168 (39.6%)	$134,278,886	$316,695	$799,279	#4
Myocardial infarction	367	172 (46.9%)	$120,050,269	$327,112	$697,967	#7
Venous thromboembolism	239	106 (44.4%)	$74,219,293	$310,541	$700,182	#6
Aortic aneurysm and dissection	191	91 (47.6%)	$58,621,501	$306,919	$644,192	#10
Arterial thromboembolism	123	58 (47.2%)	$32,543,403	$264,581	$561,093	#15
Other vascular	340	129 (37.9%)	$99,264,302	$291,954	$769,491	NA
All vascular	1684	724 (43.0%)	$518,977,654	$308,182	$716,820	NA
Infection						
Sepsis	179	65 (36.3%)	$51,408,008	$287,196	$790,892	#5
Meningitis and encephalitis	136	60 (44.1%)	$69,149,562	$508,453	$1,152,493	#2
Spinal abscess	131	66 (50.4%)	$95,674,011	$730,336	$1,449,606	#1
Pneumonia	97	33 (34.0%)	$29,108,022	$300,083	$882,061	#3
Endocarditis	57	21 (36.8%)	$12,580,534	$220,711	$599,073	#11
Other infection	392	156 (39.8%)	$144,053,537	$367,484	$923,420	NA
All infection	992	401 (40.4%)	$401,973,674	$405,215	$1,002,428	NA
Cancer						
Lung cancer	472	178 (37.7%)	$104,937,298	$222,325	$589,535	#12
Breast cancer	434	188 (43.3%)	$110,332,523	$254,222	$586,875	#13
Colorectal cancer	334	128 (38.3%)	$85,208,445	$255,115	$665,691	#9
Prostate cancer	147	71 (48.3%)	$41,365,887	$281,401	$582,618	#14
Melanoma	142	66 (46.5%)	$45,704,472	$321,862	$692,492	#8
Other cancer	1264	452 (35.8%)	$328,025,177	$259,514	$725,719	NA
All cancer	2793	1083 (38.8%)	$715,573,803	$256,203	$660,733	NA
Aggregate						
Big Three total	5469	2208 (40.4%)	$1,636,525,131	$299,237	$741,180	NA
Total "Top 5" for Big Three[d]	3473	1471 (42.4%)	$1,065,182,114	$306,704	$724,121	NA
Total of "Other" for Big Three[d]	1996	737 (36.9%)	$571,343,016	$286,244	$775,228	NA
Non-Big Three (all others)	1910	818 (42.8%)	$682,853,806	$357,515	$834,785	NA
Grand total	7379	3026 (41.0%)	$2,319,378,937	$314,322	$766,483	NA

[a]Paid refers to those cases with a non-zero indemnity payment compensating harms, whether this occurred through settlement, arbitration, or legal judgment. Unpaid claims are those with $0 indemnity payment, though there may still be an "expenses" payment associated with the claim (i.e. case management cost, including legal fees, which can be substantial, especially for cases going to trial and ending in a defense verdict). [b]Percentage represents the fraction of all claims for that row that were "paid" as per the definition listed in footnote "a". [c]Rank represents the rank order from the highest (#1) to the lowest (#15) by mean per-claim payment (paid claims only) for specific diseases. "Other" and total rows are not considered in the payment ranking. [d]The "Top 5" diseases represent the five most common from each Big Three category (vascular events, infections, and cancers). The "Other" diseases represent the aggregation of all other cases within that Big Three category. Together, "Top 5" and "Other" sum to the Big Three total. NA, not applicable.

patients were 50.2% female with a median age of 51 (IQR 39–61, range 0–99). Although high-severity cases skewed slightly older, they were distributed across all age deciles, with more cases age ≤30 (n = 1951) than >70 (n = 1029) (Figure 2; Supplementary material B1, Table S5). Infections outnumbered vascular events and cancers among those aged ≤20 and >90; cancers were more common among those aged 21–80; vascular events were more common among those aged 81–90 (Figure 2; Supplementary material B1, Table S5).

The Big Three diseases accounted for 61.7% (n = 7149/11,592) of all diagnostic error claims and 67.3% ($1.8B of $2.6B) of all diagnostic error payouts (Table 1). Among low- or medium-severity cases (n = 4213), the Big Three accounted for 39.9% (n = 1680), broken down as vascular events (8.0%, n = 335), infections (15.9%, n = 668), and cancers (16.1%, n = 677). Among high-severity cases (n = 7379), the Big Three accounted for 74.1% (n = 5469), broken down as vascular events (22.8%, n = 1684), infections (13.5%, n = 992), and cancers (37.8%, n = 2793). The distribution of high-severity
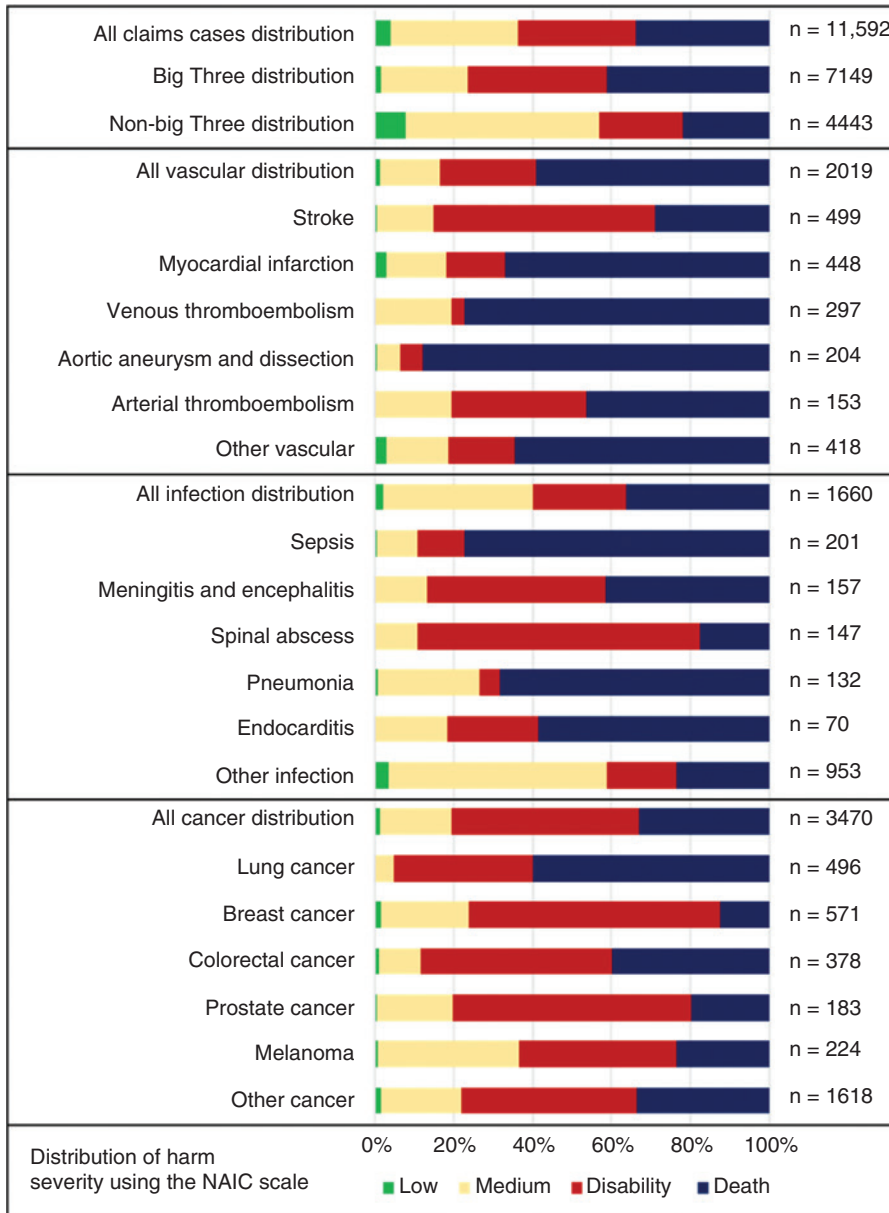
Figure 1: Distribution of harm severity by disease category and disease (n = 11,592).
The stacked bar graph shows the distribution of harms for each category or disease group using the NAIC Severity Scale. High-severity harm cases (n = 7379) include both "disability" (permanent, significant/major/grave) and "death" (including fetal death). NAIC, National Association of Insurance Commissioners.

cases across the Big Three categories was stable over the study period, despite growth in total claims in the CBS database – vascular mean 23% (range 19–24%); infection mean 13% (range 11–16%); cancer mean 38% (range 34–40%); and other (non-Big Three) mean 26% (range 23–30%) (Supplementary material B2, Tables S9 and S10, Figure S4).

The "Top 5" diseases for each of the Big Three categories are shown in Figure 3. Collectively, these 15 specific conditions accounted for 47.1% (n = 3473/7379) of all high-severity misdiagnosis-related harm cases and 63.5% (n = 3473/5469) of all Big Three high-severity cases. The top vascular disease was stroke; the top infection was sepsis; and the top cancer was lung cancer. Big Three diseases were more likely to be associated with high-severity misdiagnosis-related harms than non-Big Three diseases [76.5% (n = 5469/7149) vs. 43.0% (1910/4443), OR 4.32 (95% CI 3.98–4.68)]; this was also true for each individual Big Three category and disease (Figure 1). The overall death-to-disability ratio was 1.1; among vascular events, it was 2.4 (disease-specific ratio range 0.5 to 22.9); among infections, it was 1.5 (range 0.2–12.9); among cancers, it was 0.7 (range 0.2–1.7) (Supplementary material B1, Table S4).
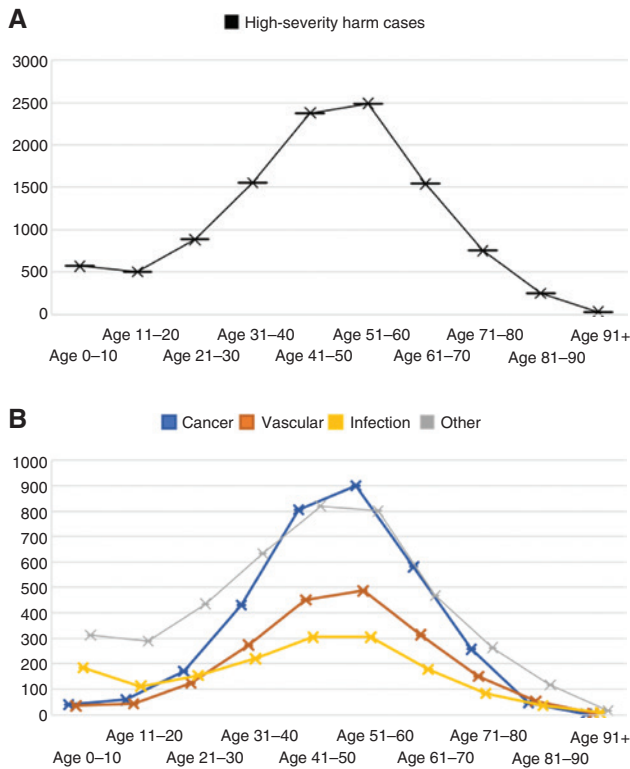
A

B

Figure 2: Age distribution of high-severity harm cases, 2006–2015. (A) All high-severity harm cases; (B) high-severity harm cases stratified by Big Three category (cancer, vascular, and infection) and other.

Mean per-claim payments were greatest for infections, intermediate for vascular events, and least for cancers; the differences were more pronounced when averaging only cases with a non-zero indemnity payment (Table 2). Payments varied by disease, ranging from $0.6M to $1.4M per claim. Neurologic diseases occupied three of the top five spots [spinal abscess $1.4M (#1); meningitis and encephalitis $1.2M (#2); stroke $0.8M (#4)] and non-neurologic infections the other two [pneumonia $0.9M (#3); sepsis $0.8M (#5)]. The remaining 10 diseases were clustered between $0.6M and $0.7M per case.

Misdiagnosis-related harms from diseases in each of the Big Three categories were not evenly distributed across clinical settings or clinical providers. The majority of diagnostic errors, including those resulting in serious harms, occurred in ambulatory settings – 71.2% (n = 5257/7379) of high-severity cases occurred in emergency departments (EDs) or outpatient clinics. High-severity cases from EDs and inpatient settings were disproportionately from vascular events and infections, while those in non-ED ambulatory care settings were disproportionately from cancers (Figure 4), except in pediatric care, where infections and other non-Big Three diseases dominated (Supplementary material B1, Table

S6). The primary responsible service corresponded to these differences in diseases and settings. More than half of the high-severity harm cases [52.0% (n = 3836/7379)] involved claims against four disciplines – internal medicine (n = 1047, cancer > vascular > infection), emergency medicine (n = 1025; vascular > infection >> cancer), family medicine (n = 938; cancer > vascular > infection), and radiology (n = 826; cancer >> vascular > infection). Vascular events were the leading claims in emergency medicine, cardiology, hospital medicine, and neurology. Infections were the leading claims in pediatric care. Cancers were the leading claims in radiology, internal medicine, family medicine, pathology, gynecology, gastroenterology, urology, general surgery, dermatology, otolaryngology, pulmonology, and oncology.

Overall, roughly half of the cases involved general care clinicians. Among high-severity cases (n = 7379), 51.0% (n = 3763) involved either the Centers for Disease Control and Prevention (CDC)-defined primary care disciplines [24] [34.0% (n = 2507) – internal medicine > family medicine >> pediatrics > gynecology > obstetrics >> geriatrics] or generalist acute care providers [17.0% (n = 1256) – emergency medicine >> hospitalist]. The other primary responsible services were divided almost evenly among medical specialties [16.3% (n = 1205) – cardiology > gastroenterology > neurology > pulmonology > oncology > others]; surgical disciplines [15.4% (n = 1135) – general surgery > orthopedics > urology > ophthalmology > neurosurgery > otolaryngology > dermatology > others]; and diagnostic service providers [14.8% (n = 1090) – radiology >> pathology >> others]. A small number of cases involved non-physicians as the primary responsible service (2.4%, n = 174) and a few were unclassified (0.2%, n = 12).

Among high-severity harm cases, there were 26,506 specific causal factors identified; only 2.5% (n = 186/7379) of cases had no risk management issues identified. Among cases with at least one cause identified (n = 7193), an average of 3.7 specific instances per case were found. When aggregated in the 11 top-level categories (Figure 5), clinical judgment factors dominated [present in 85.7% (n = 6165/7193) of cases with causes, range 82.0–88.8% across the Big Three disease categories; 60.6% (n = 16,068/26,506) of all specific causes] (Supplementary material B1, Table S7). Overall, the five most common specific causal factors were all clinical judgment factors, representing 39.0% (n = 10,329/26,506) of all specific causal factors and 64.3% (n = 10,329/16,068) of clinical judgment factors: (1) failure or delay in ordering a diagnostic test; (2) narrow diagnostic focus with failure to establish a differential diagnosis; (3) failure to appreciate and reconcile relevant symptoms, signs, or test results; (4) failure or delay in obtaining

"Big Three" breakdown of high-severity harms among diagnostic error claim cases CRICO 2006–2015 (n = 7379)

High-severity harms include permanent, serious disability or death, including fetal death (NAIC scale 6–9)
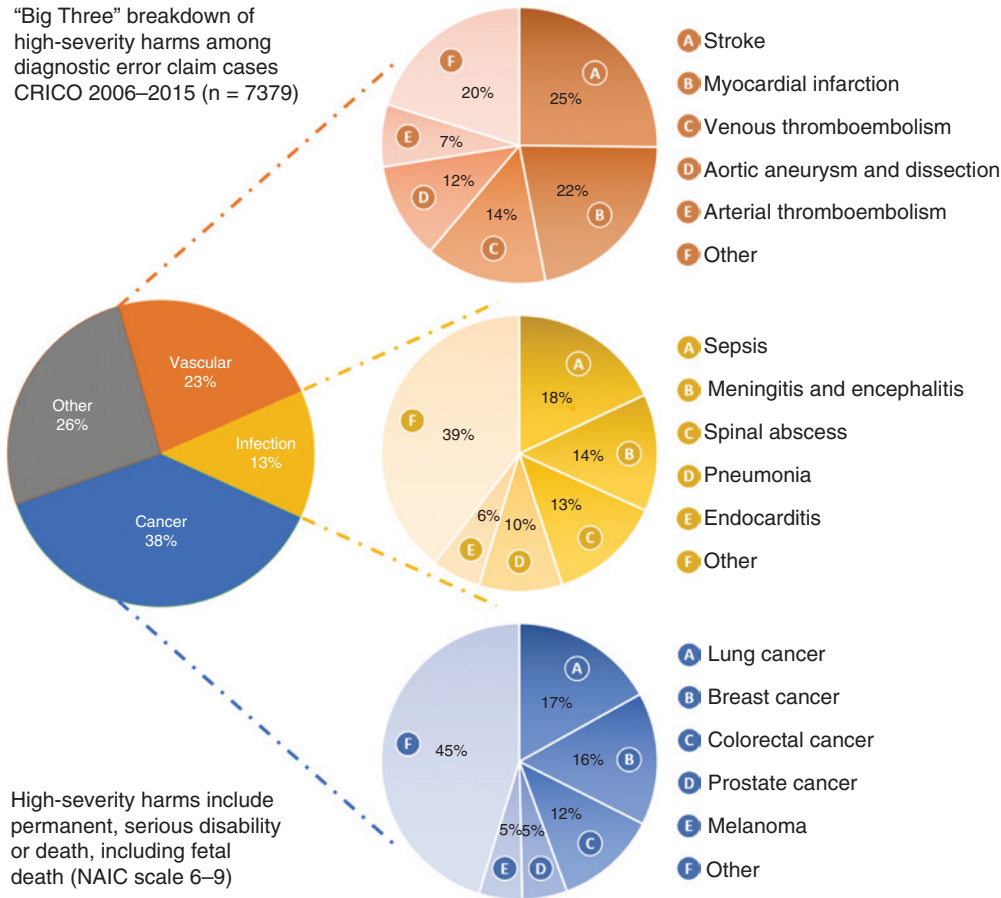
Figure 3: Big Three category and disease proportions for malpractice claim cases associated with serious misdiagnosis-related harms. The category-level pie chart (at left) reflects all 7379 high-severity claim cases, while the three disease-level pie charts (at right) show the top five diseases in each Big Three category. Together these 15 diseases account for 63.5% of Big Three cases and 47.1% of all high-severity cases. CRICO, Controlled Risk Insurance Company, Ltd. Comparative Benchmarking System; NAIC, National Association of Insurance Commissioners Severity of Injury Scale.

consultation or referral; and (5) misinterpretation of diagnostic studies (imaging, pathology, etc.). A distant second and third were communication factors [seen in 34.9% (n = 2509/7193) of cases with causes, range 32.9–39.0%; 10.1% (n = 2678/26,506) of causes] and clinical systems factors [seen in 22.0% (n = 1584/7193) of cases with causes, range 16.7–28.6%; 6.5% (n = 1722/26,506) of causes], respectively. Only 4.1% (n = 302/7379) of cases had communication factors without clinical judgment factors; only 3.8% (n = 279/7379) of cases had clinical systems factors without clinical judgment factors; and only 5.4% (n = 402/7379) had one or the other without clinical judgment factors. The breakdown of top identified specific causes within each of these high-level categories is shown in Supplementary material B1, Table S8.

Causal factors were similar across disease categories (Figure 5; Supplementary material B1, Tables S7 and S8). Two high-level causal factor categories (clinical

environment and clinical systems) showed an absolute difference in frequency of ~12% across disease categories; the other causal categories showed maximum absolute differences of 0.8–7.4% (Supplementary material B1, Table S7). The main disparity for "clinical environment" was that weekend/night-shift care was not an issue for cancer cases (1.0% cancer vs. 9.4–10.7% other categories). The main disparity for "clinical systems" was that cancer cases had a higher rate of (a) patients not receiving test results (10.1% cancer vs. 1.5–3.6% other categories) or (b) failure to follow up a new finding (9.2% cancer vs. 1.8–3.8% other categories). These two failures in "closing the loop" on diagnostic test results (n = 757) accounted for just 2.8% of specific identified causes and were present in only 8.5% of high-severity cases, but they were unevenly distributed across the Big Three categories (15.6% of cancers, 6.3% of infections, 3.8% of vascular events, and 3.2% of others). Most of these cases had more than one
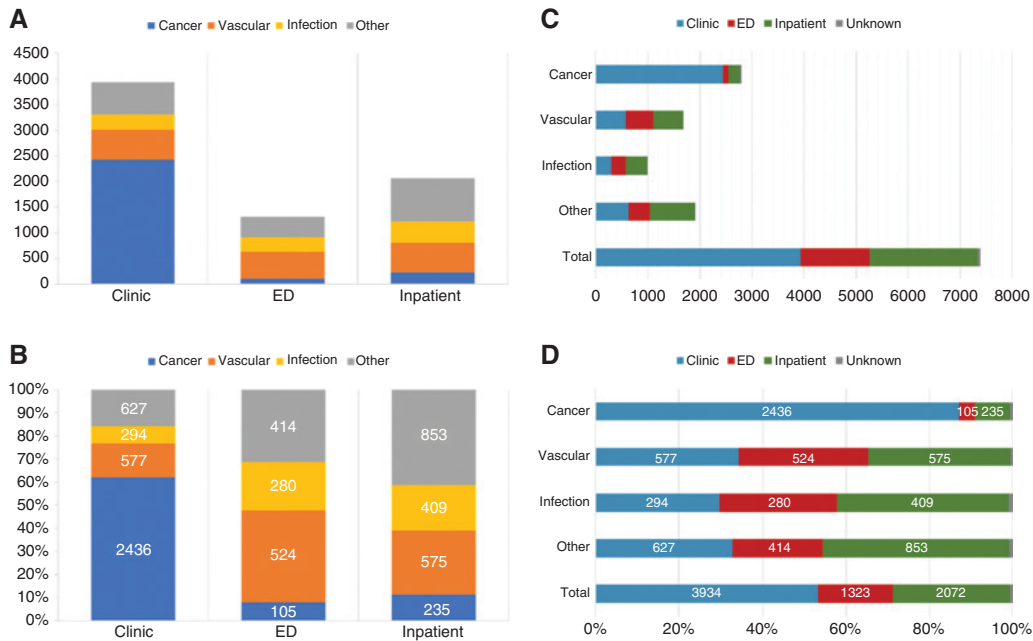
Figure 4: Relationship between Big Three serious misdiagnosis-related harms and clinical settings.
(A/B) Stacked columns show the distribution of the Big Three disease categories (bottom to top – cancer, vascular, infection, other) stratified by clinical setting. (C/D) Stacked bars show the distribution of clinical settings (left to right – clinic, ED, inpatient, unknown) stratified by the Big Three disease categories. Panels A and C show raw totals (n). Panels B and D show the same proportions (%) normalized to 100%. Although numbers of cases are only displayed in panels B and D, the same numbers are represented in panels A and C (but not shown for graphical simplicity). ED, emergency department.

identified cause (typically clinical judgment) – only 2.2% (n = 160/7379) had one of these two closing-the-loop failures without clinical judgment factors (4.7% of cancers, 1.1% of infections, 0.7% of vascular events, and 0.3% of others). Overall, high-severity cases with at least one communication or clinical systems factors without any clinical judgment factors (n = 402) were uncommon across categories (8.1% of cancers, 4.9% of infections, 3.1% of vascular events, and 3.9% of others).

Discussion

This study confirms that diagnostic errors remain the most common, most catastrophic, and most costly of serious medical errors in closed malpractice claims [3]. We found that nearly three-fourths of serious misdiagnosis-related harms are attributable to diseases in just three major categories – vascular events, infections, and cancers (the "Big Three"). Perhaps more importantly, we found that nearly half of the serious harms from diagnostic error are attributable to one of just 15 disease states (aggregating the top five diseases from each category). Not surprisingly, the Big Three diseases were unevenly distributed across

settings, with vascular events and infections dominating for inpatient and ED care, and cancers dominating for ambulatory clinics (except pediatrics, where infections dominated over the other two categories). Causes were remarkably uniform, with clinical judgment failures responsible in >85% of cases. These results suggest considerable progress could be made toward reducing overall serious misdiagnosis-related harms by improving diagnostic decision-making for a relatively small number of high-risk conditions in just a few clinical settings.

In our study, roughly half of the high-severity harms represented death and half serious, permanent disability. This accords with prior nationally representative claims-based work indicating that death and disability occur in similar proportions among diagnostic errors [3]. However, this ratio was not distributed evenly across the Big Three disease categories or individual diseases. Among vascular events, death was 2.4-fold more likely than disability, but this was reversed for stroke (1.9-fold more disability), which was the most common disease in this category. Among infections, death was 1.5-fold more likely, but this was reversed for spinal abscess (4.0-fold more disability) and meningitis/encephalitis (1.1-fold more disability), which were the second and third most common diseases in this category, respectively. These differences are
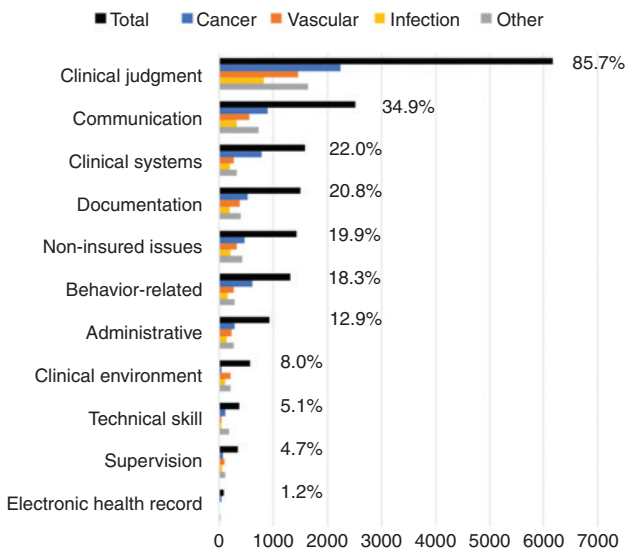
Figure 5: Contributing causal factors for high-severity harm cases. The 11 categories shown represent high-level aggregations of more granular causal factor codes. Overall, 2.5% (n = 186/7379) of cases had no risk management issues identified. The x-axis reflects the number of cases with one or more causes in that group. The total is shown in black, and the percentage at right represents the percentage of cases with at least one cause identified in that group (of 7193 with at least one cause identified overall). The percentages add up to more than 100% because cases were frequently attributed to multiple causes. The breakdown by Big Three category is represented in the colored, clustered bars below the total bar. The bars in each causal category are listed (top to bottom) as total, cancer, vascular, infection, and other.

not surprising, as dangerous neurologic diseases often produce severe disability, but not necessarily death [3, 25] (as shown before using NPDB data [3], the highest indemnity payments are for non-lethal, severe neurologic outcomes). By contrast, among cancers, disability was 1.4-fold more likely than death, although this was reversed for lung cancer (1.7-fold more death). In this case, the difference may reflect the relative paucity of life-saving treatments for late-stage lung cancer. These findings make it clear that measures of misdiagnosis-related harms must consider both morbidity and mortality.

Failures in clinical judgment were, by far, the leading identified cause of serious misdiagnosis-related harms. This result accords with prior work indicating that the vast majority of diagnostic process failures happen in bedside assessment and clinical reasoning [16, 26–28] (many of which appear to derive from knowledge gaps [27, 29, 30]) and points to a need for solutions that support better bedside clinical decision-making. This might include not only computer-based tools (e.g. device-based decision support [31] or automated image interpretation [32]) but also enhancements to diagnostic education (e.g.

simulation-based training [33]), diagnostic performance feedback (e.g. dashboards showing adverse diagnostic events after discharge [34]), and clinical teamwork in diagnosis (e.g. more immediate access to specialists [35] or more effective engagement of patients [36], nurses [37], and allied health professionals [38]).

The noted differences in causes for cancer cases, with higher rates of failing to "close the loop" and lower rates of night or weekend care, match both disease-specific biology and setting-specific workflow. Unlike vascular events and infections, which usually unfold over hours to days, dangerous cancers usually evolve over weeks to months. As a result, cancers typically present to ambulatory care clinics, where diagnostic evaluations proceed in a discontinuous fashion over a series of outpatient visits during business hours, rather than as part of a single hospital visit involving a discrete bolus of diagnostic testing that may occur off hours. This ambulatory care discontinuity offers many more opportunities for lost test results or failure to follow up new findings than during an ED evaluation or inpatient stay. Despite this, the absolute frequency of failures in closed-loop testing, even for cancers (~15%), was still dwarfed by failures in clinical judgment (>80% for missed cancers). While efforts to increase closed-loop test results reporting are sensible, they are unlikely to eliminate more than ~2–4% of serious misdiagnosis-related harms, absent additional tools that also support improved clinical reasoning.

It is an open question to what extent malpractice claims mirror serious misdiagnosis-related harms in actual clinical practice. Malpractice claims are certainly not representative of the full spectrum of diagnostic errors. It is estimated that only about 1.5% of medically negligent care events result in a malpractice claim [39], so non-representativeness is a legitimate concern. There is a clear bias in claims data toward more severe harms, as the likelihood of a claim is linked tightly to injury severity rather than the presence of care process failures (i.e. errors of planning or execution) [40]. Nevertheless, claims might still be representative of diseases causing serious misdiagnosis-related harms in practice. The proportion of Big Three diseases found in our study is nearly identical to that found in physician surveys [26, 28], triggered chart reviews [4, 41], physician voluntary reports [42], and autopsies [7, 8, 12, 13]. Vascular events, infections, and cancers accounted for 61.7% of all-severity cases in our study and 58.5% in prior clinical practice studies; 74.1% of high-severity cases in our study and 76.8% in the same clinical practice studies; and 75.3% of deaths in our study and 81.2% in prior autopsy studies (Table 3; Supplementary material C1/C2, Tables S11 and S12).

However, the distribution of Big Three diseases in claims appears to be skewed toward cancer cases, which account for a greater fraction of claims cases than vascular events and infections combined, despite the fact that they account for a smaller fraction in both clinical practice-derived and autopsy-derived diagnostic error series (Table 3). Furthermore, cancers have a substantially (~3.2 to 4.5-fold) lower annual US incidence, making it improbable that an unbiased sample of serious misdiagnosis-related harms would find a greater frequency of cancer cases (Supplementary material C3, Table S13). This probable bias toward cancer claims in malpractice data could be because it is easier to demonstrate negligence from a medicolegal perspective, either because there are more missed opportunities (as the disease process unfolds over a longer time period) or because there are more likely to be tangible artifacts from encounters that legally "prove" the prior detectable presence of the disease (e.g. radiographs showing a missed lesion). In addition to being face valid from a clinical perspective, this theory aligns with our study findings that cancers were the leading claims in both primary care settings and radiology.

Limitations

Our diagnostic error definition did not demand a care process failure, so differed from definitions that have been used by some well-respected researchers [26, 43]; however, as only 2.5% of our cases had no risk management concerns identified (i.e. would not have been considered errors using these definitions), this is unlikely to have impacted our results meaningfully. We did not review original medical records, and case summaries may have been incomplete. Some claims could have been miscoded, mischaracterized, or misclassified. Code groups were modified from the CCS original by team discussion and consensus; had they been constructed differently, the frequency order among individual disease groups might have differed. CRICO CBS data may not be representative of all malpractice claims in the US. Claims in the CBS database (or claims in general) may

Table 3: Big Three disease proportions in the current claims-based study vs. prior non-claims-based studies.[a]

Big Three category/subcategory	Malpractice claims cases (current study, CRICO CBS 2006–2015, n=11,592)	Non-claims sources (physician surveys, primary care and ED triggered chart reviews, ED physician voluntary reports, general autopsies, ICU autopsies)
Big Three % of any severity (n)	61.7% (n=7149/11,592)	58.5% (n=717/1226)[b]
Vascular	17.4% (n=2019)	23.4% (n=287)
Infection	14.3% (n=1660)	19.7% (n=241)
Cancer	29.9% (n=3470)	15.4% (n=189)
Other	38.3% (n=4443)	41.5% (n=509)
Big Three % of high severity (n)	74.1% (n=5469/7379)	76.8% (n=159/207)[c]
Vascular	22.8% (n=1684)	32.9% (n=68)
Infection	13.4% (n=992)	17.4% (n=36)
Cancer	37.9% (n=2793)	25.6% (n=53)
Other	25.9% (n=1910)	24.2% (n=50)
Big Three % of deaths (n)	75.3% (n=2944/3912)	81.2% (n=604/744)[d]
Vascular	30.5% (n=1192)	27.2% (n=202)
Infection	15.4% (n=602)	45.8% (n=341)
Cancer	29.4% (n=1150)	8.2% (n=61)
Other	24.7% (n=968)	18.8% (n=140)

[a]Additional details may be found in Supplementary materials C1/C2, including Tables S11 and S12. [b]Studies from frontline clinical practice (five studies 2009–2016, n=1226 errors) [4, 26, 28, 41, 42]. [c]Studies from frontline clinical practice (two studies 2009–2012, n=207 errors) [26, 28]. Only two of five studies reported harm severity in a way that permitted calculation. [d]Autopsy studies (45 general and ICU-based studies 1947–2011, n=8377 autopsies; n=744 errors) [7, 8, 12, 13]. There were two multi-year general autopsy studies representing 2144 autopsies and 43 ICU-based studies (30 adult, seven pediatric, six neonatal) representing 6271 autopsies. Only 24 of 30 adult studies reported on the disease breakdown, so the totals here reflect only 39 autopsy studies. General autopsies and ICU autopsies are fairly homogeneous with respect to missed Big Three diseases as causes of death (Table S12 – general 73–82%; ICU 78–82%), so are combined here. However, cancers are far less frequent among ICU deaths relative to all (general autopsy) deaths (see Table S12 for details). As a result, the data shown here in Table 3 (which over-represent ICU data relative to general autopsy data) are skewed towards underrepresenting missed cancers as causes of death. Similarly, they slightly over-represent vascular events and infections. CRICO CBS, Controlled Risk Insurance Company Comparative Benchmarking System; ED, emergency department; ICU, intensive care unit.

not be fully representative of the distribution of diseases in real-world clinical practice. It is uncertain whether serious misdiagnosis-related harms would have been prevented by prompt, correct diagnosis.

Conclusions

In malpractice claims, the Big Three diseases (vascular events, infections, and cancers) account for nearly 75% and 15 specific diseases for nearly 50% of serious misdiagnosis-related harms. Serious morbidity is about as common as mortality, so diagnostic safety and quality measures must take this into account. Mortality more often reflects missed non-neurologic vascular events/infections, while severe morbidity more often reflects missed neurologic vascular events/infections or cancer delays. Future research should seek to better clarify the relationship between harms in malpractice claims and harms in real-world clinical practice, particularly with respect to cancer, which appears to be over-represented in claims. Serious harms are disproportionately due to failures in clinical judgment, rather than problems with communication or closing the loop on test results; this suggests it will be necessary to develop systems solutions to solve cognitive problems [17] (e.g. device-based decision-support [31], simulation to improve medical education [33], diagnostic performance dashboards [34], or access to specialists via tele-consultation [35]). Research and quality improvement initiatives should target interventions that improve clinical diagnosis for high-harm diseases in specific practice settings such as stroke in the ED, sepsis in the hospital, and lung cancer in primary care.

Author contributions: *David E. Newman-Toker:* I declare that I designed the study; had primary oversight over the data analysis; designed the figures; authored the primary manuscript draft and all major revisions; and that I have seen and approved the final version. I serve as President of the Board of Directors of the Society to Improve Diagnosis in Medicine (unpaid). I act as a medico-legal consultant for both plaintiff and defense in cases related to diagnostic error. I have no other relevant conflicts of interest. *Adam C. Schaffer:* I declare that I assisted in the study design; assisted in the analysis of malpractice data, including case reviews; edited the manuscript for scientific content; and that I have seen and approved the final version. I have no conflicts of interest. *C. Winnie Yu-Moe:* I declare that I conducted the data analysis of malpractice data; edited the manuscript for scientific content; and that I have seen and approved the final version. I have no conflicts of interest. *Najlla Nassery:* I declare that I assisted in the study design; edited the manuscript for scientific content; and that I have seen and approved the final version. I have no conflicts of interest. *Ali S. Saber Tehrani:* I declare that I assisted in the study conduct; edited the manuscript for scientific content; and that I have seen and approved the final version. I have no conflicts of interest. *Gwendolyn D. Clemens:* I declare that I assisted in the study conduct; edited the manuscript for scientific content; and that I have seen and approved the final version. I have no conflicts of interest. *Zheyu Wang:* I declare that I led all statistical analyses; edited the manuscript for scientific content; and that I have seen and approved the final version. I have no conflicts of interest. *Yuxin Zhu:* I declare that I assisted in the design and conduct of statistical analyses; edited the manuscript for scientific content; and that I have seen and approved the final version. I have no conflicts of interest. *Mehdi Fanai:* I declare that I assisted in data analysis; edited the manuscript for scientific content; and that I have seen and approved the final version. I have no conflicts of interest. *Dana Siegal:* I declare that I assisted in the study design; oversaw the analysis of CRICO malpractice data; edited the manuscript for scientific content; and that I have seen and approved the final version. I serve as a member of the Board of Directors of the Society to Improve Diagnosis in Medicine (unpaid). David E. Newman-Toker served as the principal investigator, designed the study, and was the primary author of the work, including figures. Dana Siegal led the analysis of closed malpractice claims at CRICO. All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Employment or leadership: Dr. Newman-Toker conducts research related to diagnostic error, including serving as the principal investigator for grants on this topic. He currently serves as President of the Board of Directors of the Society to Improve Diagnosis in Medicine (unpaid). He acts as a medico-legal consultant for both plaintiff and defense in cases related to diagnostic error. Dana Siegal serves as a member of the Board of Directors of the Society to Improve Diagnosis in Medicine (unpaid). There are no other conflicts of interest. None of the authors have any financial or personal relationships with other people or organizations that could inappropriately influence (bias) their work.
Honorarium: None declared.
Competing interests: The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

References

1. Shojania KG, Burton EC, McDonald KM, Goldman L. Changes in rates of autopsy-detected diagnostic errors over time: a systematic review. J Am Med Assoc 2003;289:2849–56.
2. Leape LL, Berwick DM, Bates DW. Counting deaths due to medical errors (in reply). J Am Med Assoc 2002;288:2404–5.
3. Saber Tehrani AS, Lee H, Mathews SC, Shore A, Makary MA, Pronovost PJ, 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.
4. Singh H, Giardina TD, Meyer AN, Forjuoh SN, Reis MD, Thomas EJ. Types and origins of diagnostic errors in primary care settings. JAMA Int Med 2013;173:418–25.
5. 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;23:727–31.
6. Deaths and Mortality. Centers for Disease Control and Prevention, National Center for Health Statistics, 2017. Available at: http://www.cdc.gov/nchs/fastats/deaths.htm. Accessed 22 April 2019.
7. Winters B, Custer J, Galvagno Jr SM, Colantuoni E, Kapoor SG, Lee H, et al. Diagnostic errors in the intensive care unit: a systematic review of autopsy studies. BMJ Qual Saf 2012;21:894–902.
8. Custer JW, Winters BD, Goode V, Robinson KA, Yang T, Pronovost PJ, et al. Diagnostic errors in the pediatric and neonatal ICU: a systematic review. Pediatr Crit Care Med 2015;16:29–36.
9. Newman-Toker DE, Tucker L, on behalf of the Society to Improve Diagnosis in Medicine Policy Committee. Roadmap for Research to Improve Diagnosis, Part 1: Converting National Academy of Medicine Recommendations into Policy Action: Society to Improve Diagnosis in Medicine; 2018. Available at: https://www.improvediagnosis.org/roadmap/. Accessed 22 April 2019.
10. Troxel DB. Diagnostic Error in Medical Practice by Specialty. The Doctor's Advocate 2014:2,5. Available at: https://www.

thedoctors.com/the-doctors-advocate/third-quarter-2014/diagnostic-error-in-medical-practice-by-specialty/. Accessed: 22 April 2019.
11. Hanscom R, Small M, Lambrecht A. Diagnostic Accuracy: Room for Improvement: Coverys; 2018. Available at: https://coverys.com/PDFs/Coverys_Diagnostic_Accuracy_Report.aspx. Accessed: 22 April 2019.
12. Gruver RH, Freis ED. A study of diagnostic errors. Ann Intern Med 1957;47:108–20.
13. Sarode VR, Datta BN, Banerjee AK, Banerjee CK, Joshi K, Bhusnurmath B, et al. Autopsy findings and clinical diagnoses: a review of 1000 cases. Hum Pathol 1993;24:194–8.
14. Elixhauser A, Steiner C, Palmer L. Clinical Classifications Software (CCS) 2015. US Agency for Healthcare Research and Quality; 2015. Available at: http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp;http://www.hcup-us.ahrq.gov/toolssoftware/ccs/CCSUsersGuide.pdf. Accessed: 22 April 2019.
15. Improving Diagnosis in Healthcare. Institute of Medicine, 2015. Available at: http://www.nationalacademies.org/hmd/Reports/2015/Improving-Diagnosis-in-Healthcare.aspx. Accessed: 22 April 2019.
16. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. Arch Intern Med 2005;165:1493–9.
17. Newman-Toker DE, Pronovost PJ. Diagnostic errors – the next frontier for patient safety. J Am Med Assoc 2009;301:1060–2.
18. Newman-Toker DE. A unified conceptual model for diagnostic errors: underdiagnosis, overdiagnosis, and misdiagnosis. Diagnosis (Berl) 2014;1:43–8.
19. NAIC Malpractice Claims, Final Compilation. Brookfield, WI: National Association of Insurance Commissioners; 1980. Available at: https://www.naic.org/documents/prod_serv_special_med_lb.pdf. Accessed: 22 April 2019.
20. Guideline for Implementation of Medical Professional Liability Closed Claim Reporting (GDL-1077). National Association of Insurance Commissioners; 2010:p. 38. Available at: http://www.naic.org/store/free/GDL-1077.pdf. Accessed: 22 April 2019.
21. Comparative Benchmarking System. CRICO Strategies. Available at: https://www.rmf.harvard.edu/Products-and-Services/CRICO-Strategies-Products-and-Services/CBS. Accessed: 22 April 2019.
22. National Practitioner Data Bank Public Use Data File. U.S. Department of Health and Human Services, Health Resources and Services Administration, Bureau of Health Professions, Division of Practitioner Data Banks, 2019. Available at: https://www.npdb.hrsa.gov/resources/publicData.jsp. Accessed: 22 April 2019.
23. von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med 2007;147:573–7.
24. 2016 NAMCS Micro-Data File Documentation. Center for Disease Control (CDC), National Center for Health Statistics (NCHS); 2016. Available at: ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NAMCS/doc2016.pdf. Accessed: 22 Apr 2019.
25. Glick TH. The neurologist and patient safety. Neurologist 2005;11:140–9.
26. Schiff GD, Hasan O, Kim S, Abrams R, Cosby K, Lambert BL, et al. Diagnostic error in medicine: analysis of 583 physician-reported errors. Arch Intern Med 2009;169:1881–7.
27. Zwaan L, de Bruijne M, Wagner C, Thijs A, Smits M, van der Wal G, et al. Patient record review of the incidence, consequences,

and causes of diagnostic adverse events. Arch Intern Med 2010;170:1015–21.

28. Ely JW, Kaldjian LC, D'Alessandro DM. Diagnostic errors in primary care: lessons learned. J Am Board Fam Med 2012;25:87–97.

29. Sherbino J, Norman GR. Reframing diagnostic error: maybe it's content, and not process, that leads to error. Acad Emerg Med 2014;21:931–3.

30. Kerber KA, Newman-Toker DE. Misdiagnosing dizzy patients: common pitfalls in clinical practice. Neurol Clin 2015;33:565–75.

31. Newman-Toker DE, Curthoys IS, Halmagyi GM. Diagnosing stroke in acute vertigo: the HINTS family of eye movement tests and the future of the "eye ECG". Semin Neurol 2015;35:506–21.

32. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542:115–8.

33. Omron R, Kotwal S, Garibaldi BT, Newman-Toker DE. The diagnostic performance feedback "calibration gap": why clinical experience alone is not enough to prevent serious diagnostic errors. AEM Educ Train 2018;2:339–42.

34. Mane KK, Rubenstein KB, Nassery N, Sharp AL, Shamim EA, Sangha NS, et al. Diagnostic performance dashboards: tracking diagnostic errors using big data. BMJ Qual Saf 2018;27:567–70.

35. Gold D, Tourkevich R, Peterson S, Bosely J, Maliszewski B, Fanai M, et al. A novel Tele-Dizzy consultation program in the emergency department using portable video-oculography to improve peripheral vestibular and stroke diagnosis. In: Diagnostic Error in Medicine 11th Annual Conference (New Orleans, LA), November 4–6, 2018.

36. Berger ZD, Brito JP, Ospina NS, Kannan S, Hinson JS, Hess EP, et al. Patient centred diagnosis: sharing diagnostic decisions with patients in clinical practice. Br Med J 2017;359:j4218.

37. Gleason KT, Davidson PM, Tanner EK, Baptiste D, Rushton C, Day J, et al. Defining the critical role of nurses in diagnostic error prevention: a conceptual framework and a call to action. Diagnosis (Berl) 2017;4:201–10.

38. Thomas DB, Newman-Toker DE. Diagnosis is a team sport – partnering with allied health professionals to reduce diagnostic errors. Diagnosis (Berl) 2016;3:49–59.

39. Localio AR, Lawthers AG, Brennan TA, Laird NM, Hebert LE, Peterson LM, 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;325:245–51.

40. Studdert DM, Mello MM, Gawande AA, Gandhi TK, Kachalia A, Yoon C, et al. Claims, errors, and compensation payments in medical malpractice litigation. N Engl J Med 2006;354:2024–33.

41. Hudspeth J, El-Kareh R, Schiff G. Use of an expedited review tool to screen for prior diagnostic error in emergency department patients. Appl Clin Inform 2015;6:619–28.

42. Okafor N, Payne VL, Chathampally Y, Miller S, Doshi P, Singh H. Using voluntary reports from physicians to learn from diagnostic errors in emergency medicine. Emerg Med J 2016;33:245–52.

43. Singh H. Editorial: Helping health care organizations to define diagnostic errors as missed opportunities in diagnosis. Jt Comm J Qual Patient Saf 2014;40:99–101.