Harnessing Event Report Data to Identify Diagnostic Error During the COVID-19 Pandemic

Lin Shen, MD, MBI; Alexandra Levie, MPH; Hardeep Singh, MD, MPH; Kristen Murray, MPH; Sonali Desai, MD, MPH

Introduction: COVID-19 exposed systemic gaps with increased potential for diagnostic error. This project implemented a new approach leveraging electronic safety reporting to identify and categorize diagnostic errors during the pandemic.

Methods: All safety event reports from March 1, 2020, to February 28, 2021, at an academic medical center were evaluated using two complementary pathways (Pathway 1: all reports with explicit mention of COVID-19; Pathway 2: all reports without explicit mention of COVID-19 where natural language processing [NLP] plus logic-based stratification was applied to identify potential cases). Cases were evaluated by manual review to identify diagnostic error/delay and categorize error type using a recently proposed classification framework of eight categories of pandemic-related diagnostic errors.

Results: A total of 14,230 reports were included, with 95 (0.7%) identified as cases of diagnostic error/delay. Pathway 1 ($n = 1,780$ eligible reports) yielded 45 reports with diagnostic error/delay (positive predictive value [PPV] = 2.5%), of which 35.6% (16/45) were attributed to pandemic-related strain. In Pathway 2, the NLP–based algorithm flagged 110 safety reports for manual review from 12,450 eligible reports. Of these, 50 reports had diagnostic error/delay (PPV = 45.5%); 94.0% (47/50) were related to strain. Errors from all eight categories of the taxonomy were found on analysis.

Conclusion: An event reporting–based strategy including use of simple-NLP–identified COVID-19–related diagnostic errors/delays uncovered several safety concerns related to COVID-19. An NLP–based approach can complement traditional reporting and be used as a just-in-time monitoring system to enable early detection of emerging risks from large volumes of safety reports.

Diagnostic errors are receiving intense investigation in the safety community due to their high prevalence and harmful impact on patients.[1–3] Diagnostic errors affect 12 million US adult patients per year in the outpatient setting.[3] and at least 0.7% of adult admissions involve a harmful diagnostic error.[4] The COVID-19 pandemic has further strained the health care system, resulting in cognitive errors, burnout, challenges with hospital resources, and a rapid shift in operational workflows that may contribute to missed and delayed diagnoses.[5–8] Due to the novel characteristics of COVID-19–related disease, as well as its impact on hospital capacity, staffing shortages, and burnout, Gandhi and Singh proposed that the COVID-19 pandemic would exacerbate diagnostic errors.[7] They developed a taxonomy to define eight types of diagnostic errors that could be expected in the pandemic: Classic, Anomalous, Anchor, Secondary, Acute Collateral, Chronic Collateral, Strain, and Unintended. *Classic* and *Anomalous* refer to missed or delayed COVID-19 diagnoses, whereas the other six categories pertain to missed or delayed non-COVID-19 diagnoses that may result from factors related to the COVID-19 pandemic. The classification definitions as well as examples are described in Table 1. This approach accounted for diagnostic errors or delays based on possible disruptions COVID-19 may have on health care providers and the health care system, and identification of these can be used in specific mitigation strategies discussed in the original article.[7] Examples include cognitive errors, such as various forms of availability and anchoring bias; care deferment; and effects of rapidly expanding care delivery changes, such as use of telemedicine and personal protective equipment (PPE).

At our institution, the ability to recognize COVID-19–related diagnostic errors was an important part of the COVID-19 response. This included an increased emphasis on the use of data and transparency for more rapid strategic response.[9] We thus embarked on a project to characterize diagnostic errors at our institution using the Gandhi and Singh taxonomy by leveraging our incident reporting systems. Although simplified clinician reporting mechanisms have been recently developed to improve reporting of diagnostic error, incident reporting has not yet been widely used to study diagnostic error.[10,11] There are a number of criticisms related to incident reporting systems, particularly the voluntary nature of reporting, which may lead to reporting bias and hindsight bias.[12–14] However, we believed that the data would be a readily accessible and valuable source of information about events pertaining to diagnostic errors during the pandemic.

Table 1. Diagnostic Error Categories*

Nomenclature	Description	Example
Classic	Missed or delayed COVID-19 diagnosis in a patient who presents with respiratory symptoms	Patient with shortness of breath, body aches, and productive cough, tested for COVID-19 and found negative. Admitted to non-COVID-19 unit without COVID-19 precautions and subsequently found to have COVID-19 on repeat testing.
Atypical COVID-19 (Anomalous)	Missed or delayed COVID-19 diagnosis in a patient who presents with nonrespiratory symptoms	Patient presented to ED for acute hematemesis in the setting of nausea/vomiting. Admitted to ED observation unit for monitoring and serial CBCs before ultimately being admitted to hospital. Just prior to the admission, patient was noted to have fever of 100.7°F and O_2 saturation of 92% on room air. Subsequently found to be COVID-19 positive.
Anchor	Missed or delayed non-COVID-19 diagnosis because it was assumed to be COVID-19	Patient presenting with altered mental status and limited history of present illness early in the pandemic. Presumed possible COVID-19 and diagnostic evaluation for other etiologies, including primary neurologic/ischemic stroke, initially delayed due to unclear precaution protocols at the time.
Secondary	Missed or delayed non-COVID-19 diagnosis or secondary diagnosis in a patient being treated with known COVID-19 disease	Patient with history of congestive heart failure admitted for COVID-19, with worsening shortness of breath. Respiratory symptoms attributed to COVID-19 rather than to heart failure.
Delayed Presentation—Acute Condition (Acute Collateral)	Delayed diagnosis of acute non-COVID-19 diagnoses because patients are not coming in for evaluation due to infection risk	Patient had foot wound in February 2020 and was referred by primary care physician for specialty appointment. However, due to COVID-19, appointment was cancelled and rescheduling attempts not successful. Three months later, patient presented to urgent care and was found to have osteomyelitis, requiring hospitalization and IV antibiotics.
Delayed Presentation—Chronic Condition (Chronic Collateral)	Delayed diagnosis of ambulatory conditions when appointments or elective procedures are canceled	Patient with metastatic cancer with known bone metastases undergoing outpatient evaluation and systemic therapies. Significant difficulties in arranging diagnostic imaging visits due to COVID-19. Communication challenges and appointment difficulties delayed staging scans, which ultimately showed progression of disease and new onset brain metastases.
Strain	Missed or delayed non-COVID-19 diagnosis in non-COVID-19 patient because of heightened state of attention to other COVID-19 patients in an overwhelmed system	Patient admitted for GI bleeding requiring nuclear medicine bleeding scan. Primary team inadvertently ordered study on the wrong patient; however, error was caught before scan completed. Led to delay in scan for correct patient.
Unintended	Any missed or delayed diagnosis because of less direct interactions, including rapid increase of telemedicine and PPE	Patient on psychiatric medication with recent dose adjustment during hospitalization. Due to COVID-19, follow-up visits were conducted via telemedicine (no access to video visits). Patient had developed symptoms of medication adverse reaction, but subtle tremor was not noted during telemedicine visit.

ED, emergency department; CBC, complete blood count; IV, intravenous; GI, gastrointestinal; PPE, personal protective equipment

* Adapted with permission from Gandhi TK, Singh H. Reducing the risk of diagnostic error in the COVID-19 era. J Hosp Med. 2020;15:363–666. Examples are taken directly from real cases identified from our institution as part of this project.

As the project evolved, it became clear that the task of identifying COVID-19–related diagnostic errors included the need to analyze a high volume of safety event reports to discern whether a diagnostic error or delay occurred. Manual chart review for the volume of reports was not feasible given the multiple demands on our patient safety and risk management team during the pandemic; thus, we developed an informatics-based approach with the capability to preprocess large numbers of safety reports. To our knowledge, such a natural language processing (NLP)– and logic-based cohort enrichment have not been applied to safety event reports to facilitate identification of COVID-19–related diagnostic errors.

In this article we describe results of our study that aimed to identify and analyze diagnostic errors at a large US health care system to identify patient safety risks during the COVID-19 pandemic. To achieve the study aims, we sought to (1) rapidly develop a safety reporting–based workflow to identify sources of diagnostic error in our institution, particularly in the context of a novel pandemic; and (2) examine application of Gandhi and Singh's classification framework in the real world and describe categories of potential diagnostic errors that were found.

METHODS

Setting

We conducted the study at an academic tertiary care referral center in Northeastern United States with 753 inpatient beds and more than 135 ambulatory practices. The institution uses an electronic vendor-based safety reporting system (RL Solutions; RLDatix, London) capturing both inpatient and ambulatory safety events. Approximately 10% of our safety reports are from the ambulatory setting, with the remainder from the inpatient setting. Project managers within the Department of Quality and Safety reviewed all safety events related to COVID-19. We created customized fields within RL Solutions corresponding to the eight classes of COVID-19–related diagnostic errors and delays as proposed by Gandhi and Singh. These custom fields were available to patient safety and risk management specialists who routinely review safety reports, but not to frontline staff filing the initial safety report.

Safety Reporting: Pathway 1 (Original Workflow)

Early in the pandemic, we developed a workflow using safety reports that either were manually flagged as COVID-19 related or explicitly mentioned COVID-19 or coronavirus.[15] All safety reports containing the keywords "COVID" or "coronavirus" were extracted from the safety report database. These reports were then manually reviewed for potential diagnostic error using the classification for COVID-19–related diagnostic error developed by Gandhi and Singh. Chart reviews were performed in instances in which the safety reports had insufficient detail (Figure 1). Because the resources needed were substantial, it was not feasible to review all safety reports to look for diagnosis-related signals.

Safety Reporting: Pathway 2 (Complementary Workflow)

We developed a second pathway later in the pandemic response to complement Pathway 1, which looked specifically at safety reports excluded in Pathway 1 and reduced the number of manual chart reviews through the application of an NLP approach. This second pathway included cases that may not have been explicitly linked to COVID-19 by the staff member filing the safety report. We developed the software algorithm iteratively, following methods used successfully to develop other health informatics innovations.[16] This included forming a working group with both clinical and informatics expertise, establishing design requirements, and using an agile approach for iterative development to develop a rapidly deployable tool that could be integrated into an existing workflow.

We considered two main requirements for the NLP algorithm in the design phase. First, the algorithm must be able to process a high volume of safety reports. The high volume precluded manual review of each report to assess for COVID-19–related diagnostic errors. Second, the algorithm must be able to rank cases to optimize the efficiency of human review. In other words, safety reports should be rank-ordered such that study personnel can review a high-yield enriched cohort of cases to discover COVID-19–related diagnostic errors.

Pathway 2 included the following steps: (1) extraction of case-related details of safety reports from RL Solutions, (2) automated processing of safety report free text to categorize concepts it contained, (3) creation of a ranked list of safety reports based on number of concept categories flagged, and (4) manual case review of the enriched cohort. We wrote the algorithm in R programming language (R Foundation for Statistical Computing, Vienna) and developed a custom NLP approach using heuristic keyword checking against a custom lexicon. In essence, we parsed the free-text report narrative for the presence of specific keywords in specific concept categories. The list of keywords was derived from working group consensus. Case-insensitive string matching for keywords, partial word fragments, common abbreviations, and misspellings were used. We used 11 concept categories based on an association with the diagnostic process—"COVID," "Communication," "Testing," "Orders," "Precautions," "Workflow gap," "Patient condition," "Symptoms," "PPE," "Diagnostic," and "Care plan"—and multiple keywords for each. For example, the "Communication" concept category included the following keywords: "call," "video," "virtual," "VV," "misunderst," "telemedicine," "hear," "ipad," "communic," and "phone." A full list of terms used for heuristic

Cases of Diagnostic Error or Delay Identified Through Safety Report Review

14,230 total safety reports submitted March 1, 2020-February 28, 2021

Pathway 1

Pathway 2

1,780 safety reports tagged as COVID-related

12,450 safety reports not tagged as COVID-related

Of these, 110 reports identified as *potentially* diagnostic error/delay related by algorithm

Manual review of cases to identify diagnostic error/delay
25 safety reports required electronic health record review for more details

45 safety reports identified as related to Dx Error/Delay from COVID-tagged reports

50 safety reports identified as related to Dx Error/Delay from algorithm
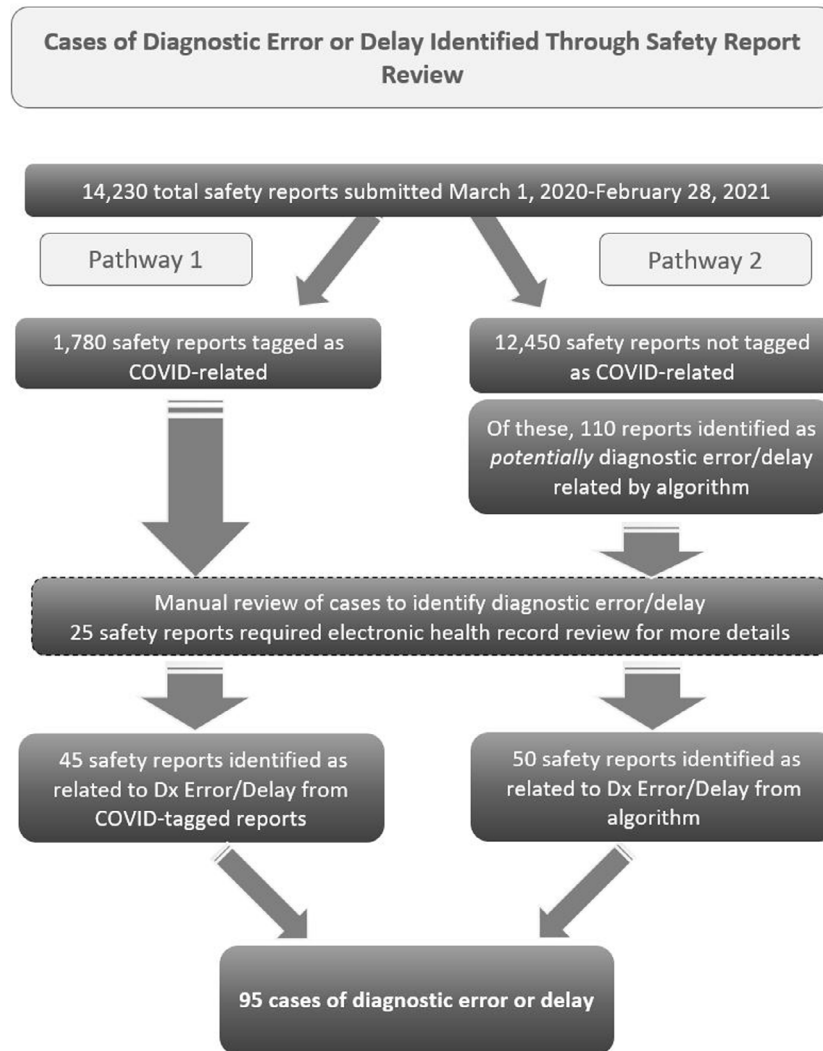
95 cases of diagnostic error or delay

Figure 1: This flowchart shows that a total of 14,230 safety reports were filed between March 1, 2020, and February 28, 2021. These were processed through two pathways. Pathway 1 (1,780 reports) contained all reports with explicit mention of COVID-19, whereas Pathway 2 (12,450 reports) used automated natural language processing to highlight specific cases for manual review. Manual review was performed for 1,780 reports in Pathway 1 and 110 safety reports in Pathway 2. A total of 95 cases of diagnostic error or delay were identified.

matching, as well as the R code for our algorithm, can be found in the GitHub repository (distributed under GPL v3 license).[17] The COVID-19 category was used specifically to exclude reports processed through Pathway 1 to avoid redundancy.

A safety report was deemed to have a match for a particular concept category if any matching keywords were found in the safety report narrative. A single safety report could contain multiple concept categories. For each safety report, categories found were tallied; reports were then sorted in descending order of number of concept categories. We hypothesized that cases involving more concept categories were more likely to be high yield for human review. Safety reports meeting the threshold of 6 or more flagged categories were manually reviewed by project managers to assess for presence of diagnostic error or delay and subcategorization using concept analysis. A threshold of 6 was

chosen from a practical perspective, as it generated a cohort that was of reasonable size for our team to manually review.

Review and Categorization for Pathways 1 and 2

Our working group, which consisted of two project managers, a clinician quality and safety expert, and a clinician-informatician, as well as an external clinician expert leader on quality and safety, was tasked with design of the pathway and rigorous review of potential diagnostic error/delay cases using the Gandhi and Singh framework. Team members included both clinical (physicians, nurses) and nonclinical staff. Primary reviewers of cases were nonclinical project managers (MPHs) who were extensively trained in the Gandhi and Singh taxonomy at the start of the project. In addition, reviewers were provided with an infographic with definitions and examples of each category. A report
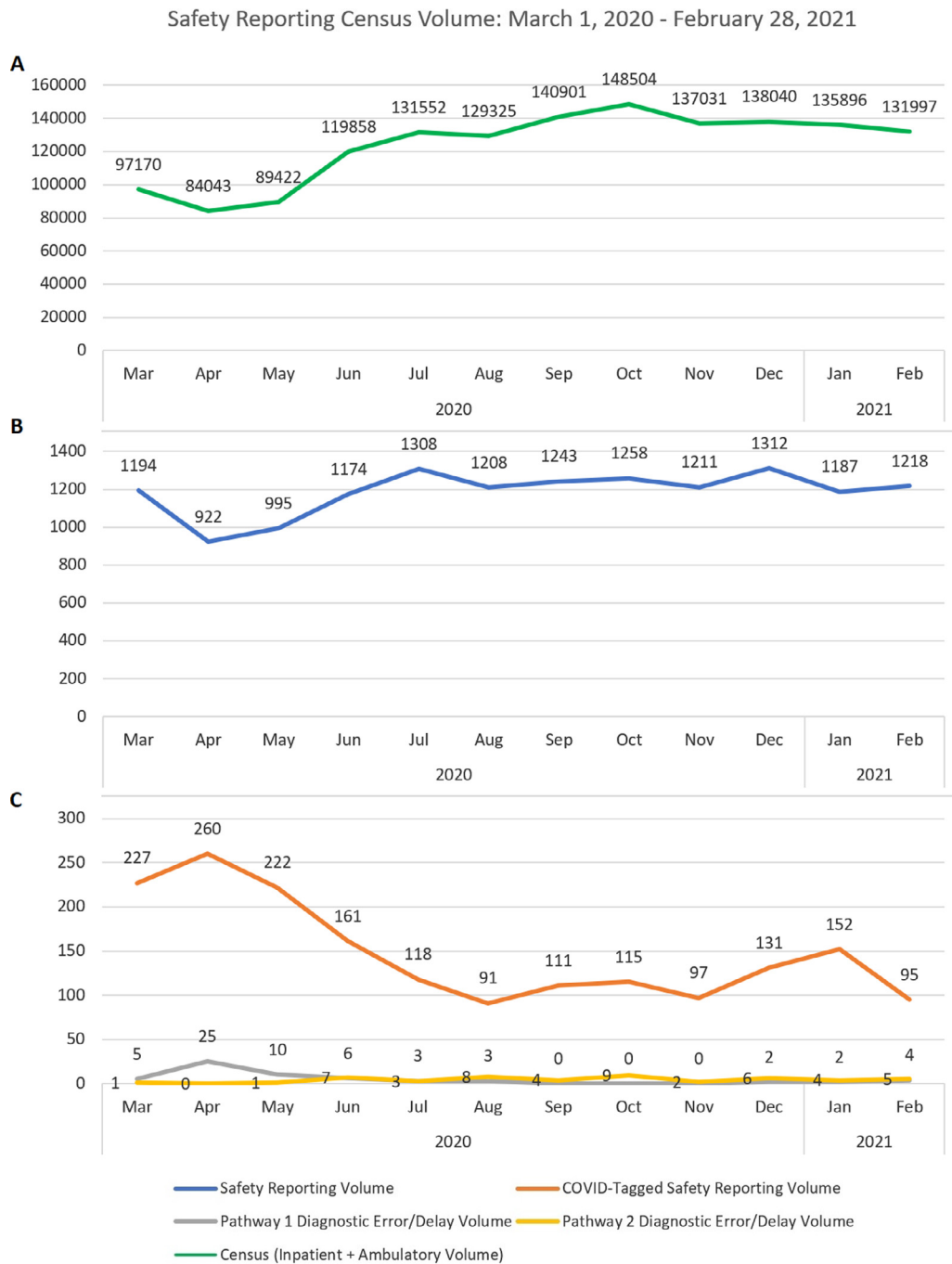
Safety Reporting Census Volume: March 1, 2020 - February 28, 2021

Figure 2: Panel A: total hospital census (ambulatory and inpatient); Panel B: safety report volume; Panel C: volume for Pathway 1 (COVID-19–tagged safety reports), Pathway 2 (natural language processing–based reports), and diagnostic errors or delays found using each pathway. Safety reporting volume roughly mirrored total hospital volume. Pathway 1 volume peaked in April 2020 and declined by midsummer. Pathway 2–identified cases increased in the summer and remained steady thereafter. Important hospital policy events during this period include deferment of nonurgent care (March 13, 2021), resumption of nonurgent care (July 20, 2021), and second deferment of nonurgent care (December 26, 2020).

was classified as having a diagnostic error or delay when it met the definition of one of the eight categories in the Gandhi and Singh taxonomy. When the safety report itself did not contain sufficient information for categorization, additional chart review in the electronic health record (EHR) was performed to find contextual details. Any reports in which either the presence of diagnostic error/delay or the most appropriate classification was unclear during the project manager review were subsequently reviewed by a physician team member of the working group for a clinical assessment. Unclear cases were subsequently presented in a group setting by the physician team member who performed the secondary review for discussion until consensus was reached.
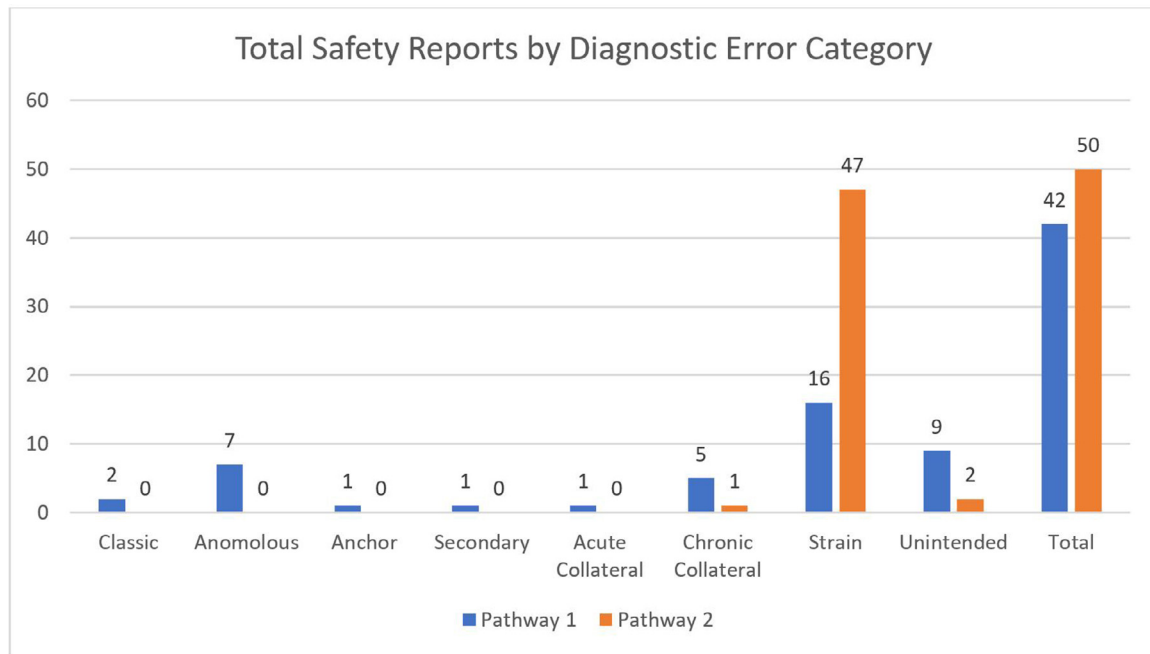
Figure 3: This chart illustrates that COVID-19–tagged safety reports (Pathway 1) had all types of errors represented, as compared to natural language processing–based reports (Pathway 2). Pathway 2 was most sensitive for detecting strain-type diagnostic errors.

RESULTS

A total of 14,230 safety reports were included in the study period of March 1, 2020, to February 28, 2021. COVID-19–related diagnostic error or delay was identified in 95 (0.7%) reports, and 1,780 reports were tagged as COVID-19 related (Pathway 1). The remaining 12,450 were not explicitly COVID-19 related (Pathway 2). Manual report review was conducted on 1,890 safety reports (Pathway 1 = 1,780; Pathway 2 = 110).

Pathway 1

Manual review of 1,780 COVID-19–explicit reports in Pathway 1 revealed 45 reports with diagnostic error/delay, a positive predictive value (PPV) of 2.5% (45/1780). Ten safety reports in Pathway 1 required additional chart review by a physician for classification. Safety reports explicitly mentioning COVID-19 peaked in April with 260 reports, before gradually declining to a steady average of around 100 reports per month (Figure 2). Pathway 1 had highest yield in April and May 2020 with 10 diagnostic error/delay–related safety reports, with subsequently months having fewer cases. Of the cases identified in Pathway 1, the most common error type was "Strain" (n = 16, 35.6%), followed by "Unintended" (n = 9, 20.0%) (Figure 3).

Pathway 2

Of the 12,450 reports in Pathway 2, 110 were highlighted for manual review by the NLP–based tool using the threshold of 6 or more category flags. Fifteen of the 110 reports required additional content review, including accessing the EHR for clinical context. Fifty of the 110 reports were found to have a diagnostic error/delay with a PPV of 45.5 % (50/110). Pathway 2 yielded an average of approximately 4.2 cases per month (Figure 2). Of the 50 diagnostic errors/delays found in Pathway 2, the predominant type was "Strain" (n = 47, 94.0%), with "Unintended" and "Chronic Collateral" making up the remainder (Figure 3).

Due to the disproportionate representation of "Strain" categorization in Pathway 2, our team opted to conduct additional qualitative content review of these safety reports. We identified three major drivers and one minor driver as primary contributors of safety events in general: supply vs. demand imbalance, patient handoff, care provider fatigue and burden, and COVID-19 status uncertainty, respectively. Across the 110 manually reviewed safety reports in Pathway 2, 39 reports involved supply vs. demand imbalance (hospital resources or services insufficient to meet prompt clinical demand), 33 reports involved patient handoff (communication challenge or disagreement between providers, teams, or services during patient transfer or shift change), 25 reports involved care provider fatigue and burden (decision-making error by staff), and 5 related to COVID-19 status uncertainty (unclear patient COVID-19 infection status). The remaining 8 reports had other miscellaneous causes unrelated to the above four categories (Table 2).

DISCUSSION

We found that COVID-19 diagnostic errors accounted for 0.7% of all safety reporting volume during a one-year period during the pandemic (95 out of 14,230 safety reports).

Table 2. Pathway 2 Diagnostic Error Domains		
Safety Driver Category	Examples	Number of Safety Reports
Supply vs. Demand Imbalance (major)	Patient required radiology study, which was delayed because of backlog in patient transport queue. Supply vs. demand imbalance of patient transport capacity.	39
Patient Handoff (major)	Patient with renal disease on dialysis admitted for unrelated postsurgical issue noted to have abnormal electrolyte lab value. Lab value not communicated in a timely fashion between unit caring for patient and admitting team, leading to delay in timing for urgent after-hours dialysis.	33
Care Provider Fatigue and Burden (major)	Patient with evolving sepsis while in emergency department. Due to concurrent competing clinical demands, signs and symptoms of worsening sepsis were not recognized until the patient was ready to be transported to the floor, leading to re-triage to the ICU.	25
COVID-19 Status Uncertainty (minor)	Patient had chronic cough unrelated to COVID-19. Chronic cough not communicated to a covering clinician during handoff. Covering clinician activates a COVID-19 response with infection control upon seeing patient with cough, causing delays for care until COVID-19 status is clarified.	5
Other		8
Grand Total		**110**

We developed two complementary pathways to enrich the cohort of reports for manual review: a manual review pathway for COVID-19–explicit reports (PPV = 2.5%) and an NLP prescreened manual review pathway for the remainder of reports (PPV = 45.5%). In addition, qualitative review of Pathway 2 safety reports revealed three major drivers and one minor COVID-19–specific driver that contributed to safety events (major: supply vs. demand imbalance, patient handoff, and care provider fatigue and burden; minor: COVID-19 status uncertainty). Safety reporting during the COVID-19 pandemic can serve as an important tool to recognize potential gaps that lead to diagnostic errors or delays in care, with the potential to serve as an early monitoring system.

NLP–Assisted Cohort Enrichment

As our project evolved, we quickly realized that the volume of safety reports made it impossible to manually review them all for COVID-19–related diagnostic errors. As such, we needed a way to extract potential reports of interest from the larger pool. We first developed a workflow reviewing only COVID-19–explicit reports. For the remainder, which constituted the majority of the reports, we developed a rapidly deployable method of processing large volumes of safety reports to identify potential diagnostic errors. By focusing on a simple, yet practical NLP approach that drove a logic-based ranking algorithm, we were able to better focus human resources to find COVID-19 diagnostic errors (Pathway 2 PPV = 45.5%). Although the intent was to find diagnostic errors, we found that our algorithm logic was sensitive to strain-related safety events.

A wide range of machine learning–based NLP with various degrees of complexity has been used for safety report-

ing in the past,[18–21] but we opted to use a simple heuristic keyword approach for the practicality of rapid deployment without the lengthy machine learning–specific validation needed in NLP approaches. We have successfully used the keyword heuristic approach in other informatics innovations at our institution.[16] Methods we developed may serve as potential resources that can be adapted and implemented in other health systems. We chose to use R programming language because it is freely available, has easily interpretable (noncompiled) code, and is widely used in the biostatistics and informatics community, lowering the barrier to entry. In addition, by avoiding a machine learning NLP approach, retraining of a machine learning model for localization would not be necessary to deploy to other sites—a task that would likely require expertise and resources more commonly found only in large academic medical centers.

Detection of early signals of trends or systemic patterns is increasingly important in the age of large volumes of data.[22] Our approach is generalizable outside of the COVID-19 pandemic because the keywords used are COVID-19 agnostic. For the specific tool developed in Pathway 2, the sensitivity and specificity can be easily adjusted by changing the threshold of categories flagged. An advantage of using a logical framework as the underpinning of the tool is that it can easily be adapted to search for other types of safety reports. For example, if an institution was interested in having monitoring for signals related to COVID-19 testing in safety reports, one could create a query to identify safety reports containing COVID-19 and testing concept categories. In addition, new features could be built on top of this foundation, including more complex tasks such as aggregating data for summary; searching for more specific concepts for more

in-depth secondary analysis for quality improvement, such as whether handoff errors were more common in specific locations; or feeding traditional and nontraditional dashboards, such as word clouds.

Limitations

One limitation of our tool is that because manual review for diagnostic error was not done for all 14,230 safety reports, we are unable to estimate the sensitivity for all diagnostic errors using our approach. This is mainly because the tool was developed out of necessity as the project evolved and not planned a priori. As such, our project was not scoped to have the resources to conduct a rigorous evaluation of the tool. However, given that the tool was designed to generate a cohort that was practical for our study staff, we think that this approach is still useful in a real-world setting. Second, the lexicon we developed for keyword heuristic checking may not be all inclusive. Although the number of diagnostic errors found was low, the figure in our study was quite similar to the 0.7% pooled rate of errors for hospitalized patients in a recent meta-analysis.[4] Future work could include further refinement and expansion of the heuristic categories used in this project.

Safety Reporting Trend Shift During the Pandemic Response

Overall safety report volume was similar in 2020 as compared to 2019, except for a notable decline from March to May. We saw a shift in the type of COVID-19 diagnostic errors as the pandemic went on. Initially, diagnostic error types varied, with the Strain, Unintended, Anomalous, and Chronic Collateral categories being the most frequent (Figure 3). Over the first four months, we saw both a decline in diagnostic errors found in COVID-19–labeled reports and a decrease in non-strain-related diagnostic errors. By June 2020 most diagnostic errors were predominantly strain related and were found using Pathway 2 (Figure 2). Our observations reflect that certain types of diagnostic errors may have been more frequent early in the pandemic, when knowledge of the disease and its management was still evolving. Specifically, certain factors were more prominent at the start of the pandemic, including (1) development, implementation, and repeated revisions of regulations, policies that affect care delivery[23–25]; (2) scientific and epidemiologic knowledge gaps on novel infectious diseases[26]; (3) need for education for both health care workers and patients[27]; and (4) development of individual attitudes and emotional response toward the pandemic.[28,29] Recognizing that different types of errors may occur in different stages of a pandemic or with a novel disease may be useful to assess future safety risks.

Major Themes in Pandemic-Strained Health Care Setting

COVID-19 emphasized the need for real-time data mining for detection of early risk signals to target.[9,22] By coupling safety reports with an NLP–based approach, human resources can be directed to safety events that may be indicative of larger systemic problems while the number of cases is still small. At our institution, safety reports drew attention to various drivers of safety events. Some of these drivers, such as communication challenges related to patient handoff, have long been recognized as an important source of medical error.[30,31] Others, such as supply vs. demand imbalance, took on new meaning during periods of extreme clinical demands on the hospital system. From a capacity and surge response perspective, significant attention has been paid to considerations such as number of beds and durable equipment such as ventilators and PPE.[32–34] However, Pathway 2–flagged safety reports repeatedly highlighted cases of supply vs. demand imbalance in less visible areas, such as patient transport and phlebotomy services. Early recognition of these signals can help inform a more robust response to ensure that additional resources are not overlooked in key low-visibility areas.

Clinician fatigue and burden is often a difficult phenomenon to identify. Diagnostic errors, particularly certain categories such as Strain, Anomalous, or Anchor errors,[7] may suggest increasing clinician exhaustion. Finding a general signal of strain may be useful as an institutional barometer as well as for identifying systemic areas to reinforce, as an overburdened clinician is an ineffective safety net and a potential source of error. The use of an NLP–based tool to assess for diagnostic error and strain may provide critical information for health system leaders. Within our institution, we saw an increase in strain-related reporting several months after the start of the pandemic that persisted for the duration of the study period, regardless of the number of active COVID-19 patients. The NLP–based tool can be applied to identify diagnostic errors using safety reporting data. Although our organization is not currently leveraging this approach due to limited bandwidth on our team, it is our goal to employ this approach in the near future.

CONCLUSION

During a one-year period in the COVID-19 pandemic, our organization developed a new safety report–based workflow to identify diagnostic errors and delays related to COVID-19. A strategy using two complementary approaches (traditional reporting and a simple NLP– and logic-based algorithm) was effective in discovering diagnostic errors and could be a useful early signal detector for trends in safety reports. This strategy significantly reduced the number of manual reviews needed to find a true diagnostic error and can be readily adapted and applied to other settings and situations. All eight categories of diagnostic errors previously

described by Gandhi and Singh were found at our institution, highlighting the need to address each of them through multifaceted interventions.

Lin Shen, MD, MBI, is Director of Clinical Informatics, Division of Gastroenterology, Hepatology and Endoscopy, Department of Medicine, Brigham and Women's Hospital (BWH), Boston, and Instructor in Medicine, Harvard Medical School, Boston. **Alexandra Levie, MPH**, formerly Project Manager, Ambulatory Safety Nets, Department of Medicine, BWH, is Project Manager, Office of Strategic Alignment, Johns Hopkins Home Care Group, Baltimore. **Hardeep Singh, MD, MPH**, is Chief, Health Policy, Quality & Informatics Program, Center for Innovations in Quality, Effectiveness and Safety, Michael E. DeBakey VA [US Department of Veterans Affairs] Medical Center, Houston, and Professor of Medicine, Baylor College of Medicine, Houston. **Kristen Murray, MPH**, formerly Program Coordinator, Department of Medicine, BWH, is Integration Design Associate, athenahealth, Watertown, Massachusetts. **Sonali Desai, MD, MPH**, is Director of Quality, Department of Medicine, BWH, and Assistant Professor of Medicine, Harvard Medical School. Please address correspondence to Sonali Desai, sdesai5@bwh.harvard.edu.

REFERENCES

1. Giardina TD, et al. Learning from patients' experiences related to diagnostic errors is essential for progress in patient safety. Health Aff (Millwood). 2018;37:1821–1827.
2. Singh H, Graber ML, Hofer TP. Measures to improve diagnostic safety in clinical practice. J Patient Saf. 2019;15:311–316.
3. Singh H, Meyer AND, Thomas EJ. The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations. BMJ Qual Saf. 2014;23:727–731.
4. Gunderson CG, et al. Prevalence of harmful diagnostic errors in hospitalised adults: a systematic review and meta–analysis. BMJ Qual Saf. 2020;29:1008–1018.
5. Jha SS, et al. The effect of COVID-19 on interventional pain management practices: a physician burnout survey. Pain Physician. 2020;23:S271–S282.
6. Cohen J, Rodgers YVM. Contributing factors to personal protective equipment shortages during the COVID-19 pandemic. Prev Med. 2020;141:106263.
7. Gandhi TK, Singh H. Reducing the risk of diagnostic error in the COVID-19 era. J Hosp Med. 2020;15:363–666.
8. Patrinley JR Jr, et al. Lessons from operations management to combat the COVID-19 pandemic. J Med Syst. 2020 Jun 10;44:129.
9. Singh H, Sittig DF, Gandhi TK. Fighting a common enemy: a catalyst to close intractable safety gaps. BMJ Qual Saf. 2021;30:141–145.
10. Marshall TL, et al. Increasing physician reporting of diagnostic learning opportunities. Pediatrics. 2021;147:e20192400.
11. Okafor N, et al. Using voluntary reports from physicians to learn from diagnostic errors in emergency medicine. Emerg Med J. 2016;33:245–252.
12. Thomas EJ, Petersen LA. Measuring errors and adverse events in health care. J Gen Intern Med. 2003;18:61–67.
13. Weissman JS, et al. Error reporting and disclosure systems: views from hospital leaders. JAMA. 2005 Mar 16;293:1359–1366.
14. Cullen DJ, et al. The incident reporting system does not detect adverse drug events: a problem for quality improvement. Jt Comm J Qual Improv. 1995;21:541–548.
15. Desai S, et al. Rapid-cycle improvement during the COVID-19 pandemic: using safety reports to inform incident command. Jt Comm J Qual Patient Saf. 2020;46:715–718.
16. Shen L, et al. Clinical decision support system, using expert consensus-derived logic and natural language processing, decreased sedation-type order errors for patients undergoing endoscopy. J Am Med Inform Assoc. 2021 Jan 15;28:95–103.
17. GitHub. BWH_safetyreport. Shen L. Apr 26, 2021. Accessed Nov 2, 2021. https://github.com/LinShenMD/BWH_safetyreport.
18. Brunson JC, Laubenbacher RC. Applications of network analysis to routinely collected health care data: a systematic review. J Am Med Inform Assoc. 2018 Feb 1;25:210–221.
19. Erhardt RAA, Schneider R, Blaschke C. Status of text-mining techniques applied to biomedical text. Drug Discov Today. 2006;11:315–325.
20. Denecke K. Automatic analysis of critical incident reports: requirements and use cases. Stud Health Technol Inform. 2016;223:85–92.
21. Young IJB, Luz S, Lone N. A systematic review of natural language processing for classification tasks in the field of incident reporting and adverse event analysis. Int J Med Inform. 2019;132:103971.
22. Adelman JS, Gandhi TK. COVID-19 and patient safety: time to tap into our investment in high reliability. J Patient Saf. 2021 Jun 1;17:331–333.
23. DeVoe JE, Cheng A, Krist A. Regional strategies for academic health centers to support primary care during the COVID-19 pandemic: a plea from the front lines. JAMA Health Forum. 2020;1:e200423.
24. Lau J, et al. Staying connected in the COVID-19 pandemic: telehealth at the largest safety-net system in the United States. Health Aff (Millwood). 2020;39:1437–1442.
25. Arora KS, Mauch JT, Gibson KS. Labor and delivery visitor policies during the COVID-19 pandemic: balancing risks and benefits. JAMA. 2020 Jun 23;323:2468–2469.
26. Qu G, et al. An imperative need for research on the role of environmental factors in transmission of novel coronavirus (COVID-19). Environ Sci Technol. 2020 Apr 7;54:3730–3732.
27. Loeb AE, et al. Departmental experience and lessons learned with accelerated introduction of telemedicine during the COVID-19 crisis. J Am Acad Orthop Surg. 2020 Jun 1;28:e469–e476.
28. Akgün KM, et al. Communication strategies to mitigate fear and suffering among COVID-19 patients isolated in the ICU and their families. Heart Lung. 2020;49:344–345.
29. da Silva FCT, Neto MLR. Psychological effects caused by the COVID-19 pandemic in health professionals: a systematic review with meta-analysis. Prog Neuropsychopharmacol Biol Psychiatry. 2021 Jan 10;104:110062.
30. Müller M, et al. Impact of the communication and patient hand-off tool SBAR on patient safety: a systematic review. BMJ Open. 2018 Aug 23;8:e022202.
31. Mardis M, et al. Shift-to-shift handoff effects on patient safety and outcomes: a systematic review. Am J Med Qual. 2017;32:34–42.
32. Griffin KM, et al. Hospital preparedness for COVID-19: a

practical guide from a critical care perspective. Am J Respir Crit Care Med. 2020 Jun 1;201:1337–1344.

33. Carenzo L, et al. Hospital surge capacity in a tertiary emergency referral centre during the COVID-19 outbreak in Italy. Anaesthesia. 2020;75:928–934.

34. Weissman GE, et al. Locally informed simulation to predict hospital capacity needs during the COVID-19 pandemic. Ann Intern Med. 2020 Jul 7;173:21–28.