Using Community Detection Techniques to Identify Themes in COVID-19–Related Patient Safety Event Reports

Christian Boxley, BS,* Seth Krevat, MD,* Srijan Sengupta, PhD,† Raj Ratwani, PhD,* and Allan Fong, MS*

Objectives: The COVID-19 pandemic has transformed how healthcare is delivered to patients. As the pandemic progresses and healthcare systems continue to adapt, it is important to understand how these changes in care have changed patient care. This study aims to use community detection techniques to identify and facilitate analysis of themes in patient safety event (PSE) reports to better understand COVID-19 pandemic’s impact on patient safety. With this approach, we also seek to understand how community detection techniques can be used to better identify themes and extract information from PSE reports.

Methods: We used community detection techniques to group 2082 PSE reports from January 1, 2020, to January 31, 2021, that mentioned COVID-19 into 65 communities. We then grouped these communities into 8 clinically relevant themes for analysis.

Results: We found the COVID-19 pandemic is associated with the following clinically relevant themes: (1) errors due to new and unknown COVID-19 protocols/workflows; (2) COVID-19 patients developing pressure ulcers; (3) unsuccessful/incomplete COVID-19 testing; (4) inadequate isolation of COVID-19 patients; (5) inappropriate/inadequate care for COVID-19 patients; (6) COVID-19 patient falls; (7) delays or errors communicating COVID-19 test results; and (8) COVID-19 patients developing venous thromboembolism.

Conclusions: Our study begins the long process of understanding new challenges created by the pandemic and highlights how machine learning methods can be used to understand these and similar challenges. Using community detection techniques to analyze PSE reports and identify themes within them can help give healthcare systems the necessary information to improve patient safety and the quality of care they deliver.

Key Words: patient safety, coronavirus, COVID-19, natural language processing

Changing healthcare facility processes in response to COVID-19 may have exacerbated existing patient safety hazards and may have given rise to new patient safety hazards that did not exist before COVID-19. Identifying the nature of patient safety hazards associated with COVID-19 and determining whether these hazards are being addressed or persist over time is critical to improving patient safety.

Patient Safety Event Reports

Patient safety event (PSE) reports, which are descriptions of patient safety issues reported by frontline healthcare workers, offer a unique opportunity to better understand COVID-19–related safety hazards. Many healthcare systems have a PSE reporting system in place where staff members can report unsafe conditions, near-miss safety events, and patient harm events.8,9 Patient safety event reports are one lens for understanding hazards in healthcare and can indicate where systems may be failing and potentially putting current or future patients at risk of harm.10

Patient safety event reports generally contain structured data and unstructured free text. The structured data in PSE systems tend to include fields such as the general event type, specific event type, the site where the event occurred, the severity of the event, and patient demographics. In the unstructured free text, the reporter can provide details on the event’s narrative, outcome, and resolution.11

Data from these reporting systems have the potential to guide efforts toward providing safer care for patients. Patient safety event reports can be analyzed to understand trends and patterns in patient safety12,13; however, working with these data present several challenges.

Challenges With Analyzing PSE Reports

The first challenge revolves around the unstructured free-text portions of the PSE reports. These descriptions, though varied, can contain rich information as the reporter can describe the event in as much detail as they would like. However, the lack of formatting and structure along with site-specific language and frequent medical abbreviations make information challenging to abstract.14,15

The second challenge is the volume of PSE reports. Although studies have found safety events to be underreported,16 large healthcare systems still generate thousands of PSE reports each month. The unstructured free text of these PSE reports is essential in understanding the safety event. However, grouping thousands of events by their complex narratives to extract clinically relevant themes is near impossible through manual reviews.17

The third challenge is reporters are often required to tag reports with a general and specific event type category (e.g., “wrong drug,” “drug confusion”) using the taxonomy integrated into their reporting system.18-20 However, the default categories in reporting systems are challenging to interpret and differentiate—especially for complex events. In addition, frontline staff who select the event type categories may have limited understanding of the category nuances because these categories are often not defined for the reporter.21,22

These factors contribute to a large spurious tagging of categories resulting in the need for manual review.23 This process requires an extensive amount of effort by patient safety analysts...
and takes away from the analysts' time to identify safety patterns and develop solutions.²³

Objectives
These challenges highlight an opportunity for machine learning to support analysis of PSE reports. In this study, we use community detection techniques on PSE reports to better understand COVID-19 pandemic's impact on patient safety. In addition, we seek to understand how community detection techniques can be used to process and extract information from PSE reports.

METHODS

Data Source
We analyzed PSE data reported between January 1, 2020, and January 31, 2021, from 10 hospitals and various ambulatory care centers in the Mid-Atlantic region of the United States. The hospitals range from large, academic hospitals found in urban centers to smaller community hospitals in rural settings. The patient population is diverse in terms of race/ethnicity, sex, age, and health condition resulting in a data set that is more generalizable.

A list of COVID-19–related key words was used to search the free-text descriptions fields of each PSE report. We used the same key word search criteria as those in a related study—a inclusion filter of coronavir or corona vir or covid or cov-2 or cov2 or sars and an exclusion filter of covidienn or covi-den or covidean. We found these search terms captured many PSE reports while avoiding reports not related to COVID-19 (i.e., false positives). The resulting PSE reports mentioning COVID-19 were aggregated for analysis. The preprocessing and modeling method are summarized in Figure 1. This study was approved by the institutional review board at MedStar Health Research Institute.

Community Detection Techniques
Each PSE report mentioning COVID-19 was stemmed, lower cased, preprocessed to remove numbers and punctuations, and tokenized into unigrams. Term frequency-inverse document frequency was then used to create a feature vector for each of these PSE reports. This technique uses the Bag-of-Words model to create a term-document matrix where rows represent the terms and columns represent the PSE reports. The resulting term-document matrix describes what terms (i.e., features) are present throughout PSE reports. In addition, the matrix describes how often a feature appears in each PSE report relative to other reports (i.e., term frequency-inverse document frequency).

Latent semantic analysis was then used as a dimension reduction technique and performed on the term-document matrix.²⁴ We found 100 topics to be sufficient for this task. This method offers a good combination of accuracy, speed, and ease of use. It is also similar to previous approaches to model and categorize these types of data.²⁵,²⁶ A correlation matrix was then created using the latent semantic analysis transformed document topic matrix to describe how similar each PSE report was to every other PSE report. This correlation matrix was transformed into an edge list, a data structure that represents a graph through a list of its edges. Edges with correlations less than 0.65 were first removed. The remaining edge list was used to create our network.

We applied community detection techniques using the Louvain algorithm to partition the edge list into individual communities for downstream analysis. Community detection is a well-known statistical technique for clustering a network graph into dense, homogeneous clusters.²⁷–³⁰ We choose to cluster reports as we found that groups of reports were easier and more intuitive to interpret compared with groups of words and phrases (i.e., the interpretability of topic model results can be challenging).³¹ The Louvain algorithm is a greedy optimization method used in community detection techniques during network analyses. Unlike traditional clustering algorithms like k means, Louvain seeks to optimize an objective function given by the modularity of the network and has a runtime of $O(n\log n)$ where $n$ is the number of nodes.³² The Louvain algorithm was chosen for this analysis because of its benchmark performances and comparable results with other similar approaches.³³

Lastly, communities were reviewed together by a research analyst with more than 3 years of safety research experience and a clinical safety expert with more than 10 years of experience to identify clinically relevant themes.³⁴ A minimum of 10% of reports from each community were manually reviewed and summarized. These summaries were grouped into clinically relevant themes by the research analyst and clinical safety expert. Themes were then analyzed to explore potential temporal patterns.

RESULTS

Overview
There were 2353 PSE reports mentioning COVID-19 of a total of 32,524 reports. Sixty-five communities accounting for 2082 PSE reports were grouped into 8 clinically relevant themes. These summaries were grouped into clinically relevant themes by the research analyst and clinical safety expert. Themes were then analyzed to explore potential temporal patterns.

![FIGURE 1. Flow diagram for how 32,524 patient safety event reports were grouped into 8 clinically relevant themes.](image-url)
Clinically Relevant Themes

The most frequently seen theme was errors due to new or unknown COVID-19 protocols/workflows accounting for 16 communities and 497 PSE reports, followed by pressure ulcers (accounting for 6 communities and 402 PSE reports). Errors due to new or unknown COVID-19 protocols/workflows made up of the most communities while other themes were made up of as few as one community (hospital-acquired venous thromboembolism [VTE], 16 PSE reports). The 8 clinically relevant themes (excluding other) we identified are reported below.

Errors Due to New or Unknown COVID-19 Protocols/Workflows

Patient safety event reports indicated that new COVID-19 protocols and workflows throughout hospitals occasionally led to either errors in care or inadequate care for patients. There were PSE reports of patients being inappropriately kept on COVID-19 precautions after testing negative or assigned to the incorrect provider. Other patients were incorrectly placed in COVID-19 units or incorrectly placed in non–COVID-19 units after being admitted. This theme peaked in April and stayed relatively high through our study period (Fig. 2B).

Table 1: Clinically Relevant Theme, the Number of PSE Reports Within the Theme, the Number of Communities Within the Theme, and Sample Free Text From 2 PSE Reports Within the Theme

| Clinically Relevant Theme                                      | No. PSE Reports | No. Communities | Examples                                                                                                                                                                                                 |
|---------------------------------------------------------------|-----------------|-----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Error due to new or unknown COVID-19 protocols or workflows   | 497             | 16              | “Patient was admitted to semiprivate room…. COVID-19 test should be done prior to coming semiprivate room.” “COVID-19 patient arrived to the unit around and was left in the hallway unattended. There was a bed in the room however, patient was not transferred into the bed.” “HAPI found on right facial cheek—COVID-19–positive patient has had several exacerbations and reintubations… placed in prone position and changed to supine position.” “COVID-19–positive patient on airborne and contact isolation. Patient has wound to L cheek, above the hydrocolloid of the Hollister ET tube holder. Per description appears to have been a blister, stage 2 pressure injury.” “Doctor approved COVID-19 test for patient; however, there was no communication between staff regarding completing the test.” “Initial COVID-19 testing resulted in ‘quantity not sufficient’. This was not reported in handoff… the swab was not reordered or completed as needed.” “COVID-19–positive should have been on droplet and contact isolation precautions on arrival to unit… writer exposed to patient during transfer process due to lack of communication.” “Patient admitted from ED had order for Droplet isolation. Patient was brought to the floor without having mask on. The nurse and the family who came with the patient did not have masks on.” |
| COVID-19–positive patient developing pressure ulcers           | 402             | 6               | “Patient was given discharge instructions for herself and baby however patient did not sign the identifying footprint sheet for the baby… due to the COVID-19 pandemic the patient was not allowed to come back and sign the paper.” “Patient left against medical advice because of concern about COVID-19.” “Nurse failed to repeat a critical bedside glucose result for COVID-19–positive patient in timely manner.” “Patient has vital signs ordered q4h. Patient found to not have documented respiratory rate for 12 h.” “COVID-19–positive patient found sitting up on the floor… confused and impulsive… not able to follow commands. Was on complete isolation precautions with door closed.” |
| Unsuccessful/incomplete COVID-19 testing                       | 317             | 7               | “Patient was admitted to semiprivate room…. COVID-19 test should be done prior to coming semiprivate room.” “COVID-19 patient arrived to the unit around and was left in the hallway unattended. There was a bed in the room however, patient was not transferred into the bed.” “HAPI found on right facial cheek—COVID-19–positive patient has had several exacerbations and reintubations… placed in prone position and changed to supine position.” “COVID-19–positive patient on airborne and contact isolation. Patient has wound to L cheek, above the hydrocolloid of the Hollister ET tube holder. Per description appears to have been a blister, stage 2 pressure injury.” “Doctor approved COVID-19 test for patient; however, there was no communication between staff regarding completing the test.” “Initial COVID-19 testing resulted in ‘quantity not sufficient’. This was not reported in handoff… the swab was not reordered or completed as needed.” “COVID-19–positive should have been on droplet and contact isolation precautions on arrival to unit… writer exposed to patient during transfer process due to lack of communication.” “Patient admitted from ED had order for Droplet isolation. Patient was brought to the floor without having mask on. The nurse and the family who came with the patient did not have masks on.” |
| Inadequate isolation of COVID-19 patients                      | 256             | 8               | “Patient was given discharge instructions for herself and baby however patient did not sign the identifying footprint sheet for the baby… due to the COVID-19 pandemic the patient was not allowed to come back and sign the paper.” “Patient left against medical advice because of concern about COVID-19.” “Nurse failed to repeat a critical bedside glucose result for COVID-19–positive patient in timely manner.” “Patient has vital signs ordered q4h. Patient found to not have documented respiratory rate for 12 h.” “COVID-19–positive patient found sitting up on the floor… confused and impulsive… not able to follow commands. Was on complete isolation precautions with door closed.” |
| Inappropriate/ inadequate care for COVID-19 patients            | 164             | 8               | “Patient was admitted to semiprivate room…. COVID-19 test should be done prior to coming semiprivate room.” “COVID-19 patient arrived to the unit around and was left in the hallway unattended. There was a bed in the room however, patient was not transferred into the bed.” “HAPI found on right facial cheek—COVID-19–positive patient has had several exacerbations and reintubations… placed in prone position and changed to supine position.” “COVID-19–positive patient on airborne and contact isolation. Patient has wound to L cheek, above the hydrocolloid of the Hollister ET tube holder. Per description appears to have been a blister, stage 2 pressure injury.” “Doctor approved COVID-19 test for patient; however, there was no communication between staff regarding completing the test.” “Initial COVID-19 testing resulted in ‘quantity not sufficient’. This was not reported in handoff… the swab was not reordered or completed as needed.” “COVID-19–positive should have been on droplet and contact isolation precautions on arrival to unit… writer exposed to patient during transfer process due to lack of communication.” “Patient admitted from ED had order for Droplet isolation. Patient was brought to the floor without having mask on. The nurse and the family who came with the patient did not have masks on.” |
| Patient falls                                                  | 147             | 2               | “Patient was admitted to semiprivate room…. COVID-19 test should be done prior to coming semiprivate room.” “COVID-19 patient arrived to the unit around and was left in the hallway unattended. There was a bed in the room however, patient was not transferred into the bed.” “HAPI found on right facial cheek—COVID-19–positive patient has had several exacerbations and reintubations… placed in prone position and changed to supine position.” “COVID-19–positive patient on airborne and contact isolation. Patient has wound to L cheek, above the hydrocolloid of the Hollister ET tube holder. Per description appears to have been a blister, stage 2 pressure injury.” “Doctor approved COVID-19 test for patient; however, there was no communication between staff regarding completing the test.” “Initial COVID-19 testing resulted in ‘quantity not sufficient’. This was not reported in handoff… the swab was not reordered or completed as needed.” “COVID-19–positive should have been on droplet and contact isolation precautions on arrival to unit… writer exposed to patient during transfer process due to lack of communication.” “Patient admitted from ED had order for Droplet isolation. Patient was brought to the floor without having mask on. The nurse and the family who came with the patient did not have masks on.” |
| Delay or error communicating COVID-19 test result             | 57              | 2               | “The provider read the patient's rapid COVID-19 test as negative when it was positive and told the patient it was negative.” “COVID-19 test resulted as positive, patient was told she was positive, however result was entered as negative in system” |
| Hospital-acquired VTE                                         | 17              | 1               | “COVID-19–positive patient acquired VTE. Necessary treatments ordered.” “COVID-19 (asymptomatic) patient developed RLE DVT.” |

The free text was deidentified and edited for length and clarity.

HAPI, A hospital-acquired pressure injury (formerly known as a pressure ulcer); q4h, taken/received every four hours; RLE DVT, right lower extremity deep vein thrombosis.
TABLE 2. The Breakdown of PSE Reports by Hospital/Care Center

| Setting     | Hospital             | No. Reports |
|-------------|----------------------|-------------|
| Urban       | Urban hospital 1     | 559 (26.8%) |
|             | Urban hospital 2     | 255 (12.2%) |
|             | Urban hospital 3     | 234 (11.2%) |
|             | Urban hospital 4     | 96 (4.6%)   |
|             | Urban hospital 5     | 94 (4.5%)   |
|             | Urban hospital 6     | 57 (2.7%)   |
|             | Urban hospital 7     | 48 (2.3%)   |
| Suburban    | Suburban hospital 1  | 287 (13.8%) |
|             | Suburban hospital 2  | 177 (8.5%)  |
| Rural       | Rural hospital 1     | 115 (5.5%)  |
| Various     | Various ambulatory care centers | 160 (7.7%) |
| Total       |                      | 2082        |

COVID-19 Patients Developing Pressure Ulcers

Many of the identified safety events in this theme were a result of pronation—a position COVID-19–positive patients have been placed in to improve oxygenation. These safety events often saw pressure ulcers form on patients’ faces. There was also a large number of safety events where pressure ulcers developed on patients’ backsides. Pressure ulcers quickly increased in volume starting in April and peaking in May. Patient safety event report volume decreased just as quickly with only a few PSE reports involving pressure ulcers seen in July, August, and September (Fig. 2C).

Unsuccessful/Incomplete COVID-19 Testing

COVID-19 test results could be delayed for several different reasons. Reports in this theme noted breakdowns in communication between providers ordering the test and laboratory staff completing the test. Unlabeled, incorrect, broken, or missing swabs could be sent to testing facilities leading to unsuccessful testing as well (Fig. 2D).

Inadequate Isolation of COVID-19 Patients

Patient safety event reports in this theme indicated delays in properly isolating patients suspected of having COVID-19 in both inpatient and outpatient settings. There were also instances of known COVID-19–positive patients not being properly isolated on arrival to ambulatory sites and hospitals as well as transfers within hospitals. These events led to safety events where other patients, staff, and providers were unintentionally exposed to COVID-19—often without wearing personal protection equipment. The volume of these types of PSE reports increased quickly in March and plateaued until August (Fig. 2E).

Inappropriate/Inadequate Care for COVID-19 Patients

There were safety events where respiratory rates and blood glucose levels of COVID-19–positive patients were inadequately monitored. This theme peaked in May and started declining immediately. There was another smaller peak later the same year in December (Fig. 2F).

Patient Falls

Safety events documented delayed responses for patients who had fallen. This theme declined after peaking in April. This theme again grew later the same year increasing from November 2020 through January 2021 (Fig. 2G).

Delay or Error Communicating COVID-19 Test Result

Incorrect COVID-19 test results could be entered into electronic health records or communicated directly to patients, or COVID-19 test results could be available but not promptly communicated to patients. In addition, PSE reports of deviations from COVID-19 testing protocols were seen in COVID-19–positive preoperative patients leading to delays in treatment. This theme peaked in the summer and plateaued over the following months (Fig. 2H).

Hospital-Acquired VTEs

COVID-19–positive patients were at an increased risk of developing VTEs because of immobilization. Venous thromboembolisms saw a much smaller peak in April and May before decreasing to zero by August (Fig. 2I).

DISCUSSION

Clinically Relevant Themes and Changing Hazards

As the COVID-19 pandemic progresses, better understanding how the pandemic has affected patient safety over time can help ensure healthcare systems provide quality and safe care moving forward. This work expands on previous studies that analyzed COVID-19–related PSE reports by using community detection techniques to better and more quickly identify clinically relevant themes found in PSE reports and track these themes over time.

Looking through the “pressure ulcer” and “VTE” themes, the increase in pressure ulcers could indicate a relationship between the COVID-19 initial surge in patient volume and these safety events, as hospital volume of COVID-19–positive patients was greatest in April and May. Strict workflows nurses followed before the COVID-19 pandemic suggest new workflows where COVID-19–positive patients are flipped between prone and supine positions could contribute to an increased number of pressure ulcers. Hospital-acquired pressure ulcers and VTEs are both important quality metrics used by hospitals to evaluate their level of care; therefore, learning when and why the number of these reports are increasing and how to minimize these safety events are incredibly valuable.

Trends were found when reading through the “unsuccessful/incomplete COVID-19 testing” and the “delay or error communicating COVID-19 test result” themes. As the pandemic began, it was more common to see PSE reports describing test swabs lost in transit to the laboratory as well as incorrect, broken, unsealed, and unlabeled test swabs sent to laboratories. As the pandemic continued and testing became more frequent, it was more common to see tests completed on the incorrect patient and test results delayed because of breakdowns in communication. This suggests as the pandemic progressed swabbing was more likely to be done correctly. However, these improvements were occasionally offset by communications in ordering tests and delivering test results.

Reports grouped into the “improper isolation” theme declined 6 months after their initial peak suggesting new protocols were either initially difficult to implement or ineffective. However, providers were better able to isolate patients after given time to adjust. These reports did not decline as sharply as other types of PSE reports, which might point to how great of a challenge it is to properly isolate patients, even when effective protocols are put in place.

Due to the novelty of COVID-19 treatment and care, continual learning through the analysis of safety events facilitated via community detection techniques, similar to the 3 examples above, can
help us understand new risks and hopefully mitigate them before they cause patient harm. The temporal analysis presented in our study is exploratory in nature, describing the behavior of different themes over time. A more detailed and rigorous temporal analysis could provide further inferential insights by analyzing PSE reports in conjunction with external factors. For example, the temporal dynamics might be related to restrictions on medical procedures because of COVID-19 or time constraints on frontline staff who report a majority of safety event.

**Improvement Over Manual Review**

Compared with a previous study that manually reviewed every PSE report that mentioned COVID-19, we found that using...
community detection techniques and reading through 10% of the reports in a community was sufficient for us to determine the theme. The structured categories currently available to reporters are limited, and reporters do not always select the most appropriate categories. By using community detection techniques, we can bypass the unreliability of the structured categories and groups reports by their free text.

In addition, compared with manually reading through each report, our method of using community detection to group reports saves time by requiring manual review of only a fraction of the reports and by making clinically relevant themes easier to uncover. This ensures that analysts have more time to identify safety patterns and develop solutions.

Operational Relevance

The COVID-19 pandemic created new risks and the need for new protocols. Identifying patient safety themes associated with the COVID-19 pandemic is an important first step in creating concrete outputs necessary to understand how the pandemic impacted healthcare delivery. The identified patient safety themes do not fit into any structured data fields, so we can use community detection techniques to analyze the free text and better uncover patient safety themes.

To increase operational relevance, it is important to understand when and how these themes presented themselves during the COVID-19 pandemic rather than analyzing these reports as one batch. It is important to find which themes saw sharp relative increases in the volume of their PSE reports and when during the pandemic those increases took place. Likewise, it is important to know which themes had a consistent volume of PSE reports as the pandemic progressed.

Understanding changing risks and hazards and tracking them over time can help healthcare systems: (1) identify safety patterns and (2) guide decisions aimed at making care safer and more effective as the COVID-19 pandemic progresses. Future work should track similar risks over time based on harm score or potential to cause severe harm. Similar methodologies could be used to develop protocols to combat patient safety risks, track the impact of the quality interventions based on the relative volume change of reports, and compare the impact of quality interventions between sites in a healthcare system.

Challenges and Limitations

Working With PSE Reports

The voluntary nature of PSE reports often led to under reporting. This can be especially true when healthcare systems are strained like during the COVID-19 pandemic. In addition, PSE reports can be used to identify general themes not as often to conclude causal relationships. For example, medical conditions are often mentioned in PSE reports, and compare the impact of quality interventions between sites in a healthcare system.

Operational Relevance

The COVID-19 pandemic created new risks and the need for new protocols. Identifying patient safety themes associated with the COVID-19 pandemic is an important first step in creating concrete outputs necessary to understand how the pandemic impacted healthcare delivery. The identified patient safety themes do not fit into any structured data fields, so we can use community detection techniques to analyze the free text and better uncover patient safety themes.

To increase operational relevance, it is important to understand when and how these themes presented themselves during the COVID-19 pandemic rather than analyzing these reports as one batch. It is important to find which themes saw sharp relative increases in the volume of their PSE reports and when during the pandemic those increases took place. Likewise, it is important to know which themes had a consistent volume of PSE reports as the pandemic progressed.

Understanding changing risks and hazards and tracking them over time can help healthcare systems: (1) identify safety patterns and (2) guide decisions aimed at making care safer and more effective as the COVID-19 pandemic progresses. Future work should track similar risks over time based on harm score or potential to cause severe harm. Similar methodologies could be used to develop protocols to combat patient safety risks, track the impact of the quality interventions based on the relative volume change of reports, and compare the impact of quality interventions between sites in a healthcare system.

Challenges and Limitations

Working With PSE Reports

The voluntary nature of PSE reports often led to under reporting. This can be especially true when healthcare systems are strained like during the COVID-19 pandemic. In addition, PSE reports can be used to identify general themes not as often to conclude causality. For example, medical conditions are often mentioned in PSE reports, such as VTE for COVID-19-positive patients. However, unless the report specifically states that the COVID-19-positive patient developed VTE because of his or her COVID-19 care process, it is not possible to conclude a causal relationship. As a result, we are limited in our ability to look at medical conditions caused by COVID-19 and instead we focused on safety events surrounding COVID-19.

Informatics Techniques to Identify COVID-19-Related PSE Reports

In addition, we used a key word search of the free text to find PSE reports related to the COVID-19 pandemic. This ad hoc approach is limited, because PSE reports could be related to the pandemic without using one of the key words in our list or words could be misspelled in the free-text narrative. Furthermore, we used community detection techniques, which are unsupervised machine learning methods for graph clustering, to group similar PSE reports. There could be other approaches, such as supervised machine learning methods like classification or unsupervised machine learning methods such as topic modeling, which are worth exploring.

Lastly, this work is not a formal causal analysis based on a case-control study. Rather, this article should be interpreted as early, exploratory work using a novel method and statistical analysis based on word association, where we applied an ad hoc strategy based on key words to identify PSE reports that are associated with COVID-19. In other words, this serves as an application of community detection to explore PSE COVID-19 reports. Future work should look to compare the themes uncovered in this analysis to themes found in a baseline data set from previous years for a formal causal analysis of the effects of the pandemic.

CONCLUSIONS

Patient safety event reports, which are collected by most healthcare facilities, can provide key information on where there are hazards within healthcare systems that could lead to harm events. Because of the nature and volume of PSE reports, it can be difficult to find clinically relevant themes within PSE reports and even more difficult to understand how these themes change over time. Our approach involved using community detection techniques to identify clinically relevant themes in PSE reports that mention COVID-19 to identify how the COVID-19 pandemic has impacted patient safety over time. We found that the COVID-19 pandemic has directly and indirectly created challenges around the diagnosis, treatment, and proper isolation of patients. Furthermore, compared with previous methods, these themes were easier and quicker to discover because of the community detection techniques used.

Because our analysis is based on data from 10 different hospitals and various ambulatory care centers, ranging from urban to rural areas and serving a diverse patient population, our study begins to develop a general understanding of the new challenges brought forth by the pandemic. The identification of specific PSE reports and the themes arising from them using community detection techniques offers an opportunity to reduce the risk of future patient harm.

REFERENCES

1. Chopra V, Toner E, Waldhorn R, et al. How should U.S. hospitals prepare for coronavirus disease 2019 (COVID-19)? Ann Intern Med. 2020;172: 621–622.
2. Griffin KM, Karas MG, Ivascu NS, et al. Hospital preparedness for COVID-19: a practical guide from a critical care perspective. Am J Respir Crit Care Med. 2020;201:1337–1344.
3. Brat GA, Hersey S, Chhabra K, et al. Protecting surgical teams during the COVID-19 outbreak: a narrative review and clinical considerations. Ann Surg. 2020. doi:10.1097/SLA.0000000000003926.
4. Kompas M. Coronavirus disease 2019 (COVID-19): protecting hospitals from the invisible. Ann Intern Med. 2020;172:619–620.
5. Søreide K, Hallet J, Matthews JB, et al. Immediate and long-term impact of the COVID-19 pandemic on delivery of surgical services. Br J Surg. 2020; 107:1250–1261.
6. Norgan AP, Simon KE, Feehan BA, et al. Radio-frequency identification specimen tracking to improve quality in anatomic pathology. Arch Pathol Lab Med. 2020;144:189–195. doi:https://urldefense.proofpoint.com/v2/url?u=https-3A__dx.doi.org_10.5858_arpa.2019-2D0011-2DOA&d=DwIG-g&c=RxBXp2Kc-n3n5tr3n00QK zL4whPmdVgs8-bJ04M4rz&k=PAGW4ZKvZTbE3WZFmYkAnEAyJ9LG-A-63oCwH7c4u-sLhVpZkLs7r7pcn-m= 54smgZ00Zpko-eHDPs86p63cZ7T05Q_x1vdNxa28s&=.
