

Development of a Taxonomy for Medication-Related Patient Safety Events Related to Health Information Technology in Pediatrics

Kirk D. Wyatt¹ Tyler J. Benning² Timothy I. Morgenthaler³ Grace M. Arteaga⁴

¹Division of Pediatric Hematology/Oncology, Department of Pediatric and Adolescent Medicine, Mayo Clinic, Rochester, Minnesota, United States

²Mayo Clinic Alix School of Medicine, Mayo Clinic, Rochester, Minnesota, United States

³Department of Pulmonary and Critical Care Medicine, Mayo Clinic, Rochester, Minnesota, United States

⁴Division of Pediatric Critical Care Medicine, Department of Pediatric and Adolescent Medicine, Mayo Clinic, Rochester, Minnesota, United States

Address for correspondence Grace M. Arteaga, MD, Division of Pediatric Critical Care Medicine, Department of Pediatric and Adolescent Medicine, Mayo Clinic, 200 First Street Southwest, Rochester, MN 55905, United States (e-mail: Arteaga.Grace@mayo.edu).

Appl Clin Inform 2020;11:714–724.

Abstract

Background Although electronic health records (EHRs) are designed to improve patient safety, they have been associated with serious patient harm. An agreed-upon and standard taxonomy for classifying health information technology (HIT) related patient safety events does not exist.

Objectives We aimed to develop and evaluate a taxonomy for medication-related patient safety events associated with HIT and validate it using a set of events involving pediatric patients.

Methods We performed a literature search to identify existing classifications for HIT-related safety events, which were assessed using real-world pediatric medication-related patient safety events extracted from two sources: patient safety event reporting system (ERS) reports and information technology help desk (HD) tickets. A team of clinical and patient safety experts used iterative tests of change and consensus building to converge on a single taxonomy. The final devised taxonomy was applied to pediatric medication-related events assess its characteristics, including interrater reliability and agreement.

Results Literature review identified four existing classifications for HIT-related patient safety events, and one was iteratively adapted to converge on a singular taxonomy. Safety events relating to usability accounted for a greater proportion of ERS reports, compared with HD tickets (37 vs. 20%, $p = 0.022$). Conversely, events pertaining to incorrect configuration accounted for a greater proportion of HD tickets, compared with ERS reports (63 vs. 8%, $p < 0.01$). Interrater agreement (%) and reliability (kappa) were 87.8% and 0.688 for ERS reports and 73.6% and 0.556 for HD tickets, respectively.

Discussion A standardized taxonomy for medication-related patient safety events related to HIT is presented. The taxonomy was validated using pediatric events. Further evaluation can assess whether the taxonomy is suitable for nonmedication-related events and those occurring in other patient populations.

Keywords

- ▶ patient safety
- ▶ patient harm
- ▶ safety management
- ▶ electronic health records
- ▶ medical informatics

received
June 29, 2020
accepted after revision
August 21, 2020

© 2020 Georg Thieme Verlag KG
Stuttgart · New York

DOI <https://doi.org/10.1055/s-0040-1717084>.
ISSN 1869-0327.

Conclusion Wider application of standardized taxonomies will allow for peer benchmarking and facilitate collaborative interinstitutional patient safety improvement efforts.

Background and Significance

The electronic health record (EHR) facilitates storage, retrieval, sharing, and use of health care information to support clinical care.¹ However, several challenges have been recognized, including the EHR's contribution to serious patient harm.^{2,3} One central problem is that EHRs operate within a highly complex sociotechnical system that is often highly unpredictable. This often leads to a mismatch between what the system designers envisioned for the system in principle and use of the system in practice. Other challenges include clumsy automation, poor usability, lack of interoperability, and contractual obligations that shift liability away from software vendors.³ In light of these challenges, improvement of EHRs is a major focus of contemporary patient safety and quality improvement efforts.²⁻⁹

The Institute of Medicine has identified the lack of metrics to track HIT-related patient safety events as a key barrier to patient safety and quality improvement efforts.³ A recent systematic review of studies reporting HIT-related safety flaws observed that the minority of reports on the topic were quantitative, and when quantitative data were shared, data quality was often poor.¹⁰ Specifically, 21% of the 34 studies reviewed provided quantitative data, and classification categories were often not well-differentiated.¹⁰ The authors concluded that the current evidence base fails to adequately capture the scope and severity of the problem and called for a more systematic approach to HIT-related safety events. One benefit to a standardized and quantitative approach is that it may facilitate pooling patient safety event data for peer benchmarking.¹¹⁻¹⁴

At our institution, we collectively shared concern that our efforts to track and trend HIT-related safety events were hampered by the absence of a mechanism to systematically classify and organize HIT-related events. We hypothesized that the presence of a structured taxonomy for patient safety events would facilitate consistent and reliable categorization of patient safety events with the aim of furthering patient safety and quality improvement goals.

Objectives

Our primary aim was to establish a taxonomy applicable to HIT-related patient safety events. Our secondary aims were to validate the resultant taxonomy using real-world patient safety events, including assessment of interrater agreement and reliability.

Methods

Identification of Candidate Taxonomies

We first performed a literature search using PubMed and reviewed references of references to identify existing taxon-

omies that could be adapted to meet our local needs. In addition, patient safety staff submitted a preliminary set of categories that were being considered for use after internal brainstorming sessions. Candidate taxonomies that were assessed include health information technology (HIT) safety domains,^{5,7} the sociotechnical model,⁹ Magrabi's classification,¹⁵⁻²⁰ and HIT-related safety concerns.⁶

Development of Derivation and Test Datasets

We next assembled real-world patient safety data which would be used to assess the classification schemes identified in the literature review and inform taxonomy development. We utilized two data sources. The first source was a confidential, voluntary, nonpunitive patient safety event reporting system (ERS) where all staff members are encouraged to submit patient safety event reports ("reports"). The reporting system is operated using third-party software (Midas+, Conduent, Florham Park, New Jersey, United States) which allows staff to associate events with a specific patient and prompts users to enter event-specific information into an event report form. Event report forms were generally based on the Agency for Healthcare Research and Quality Common Formats. The second source was the information technology help desk (HD) "ticket" management system which documented and managed requests by users for HIT-related technical support ("tickets"). This system is also operated using third-party software (ServiceNow, ServiceNow, Santa Clara, California, United States) and prompts frontline staff or call center representatives to enter an event narrative in addition to other event-specific information.

We restricted patient safety events to HIT-related safety events, defined as events which occur "anytime the HIT system is unavailable for use, malfunctions during use, used incorrectly, or when HIT interacts with another system component incorrectly resulting in data being lost or incorrectly entered, displayed, or transmitted."²¹ While this definition is broad in scope, its broadness underscores how even subtle system malfunctions can magnify and lead to patient harm when they occur within a complex sociotechnical system. With this in mind, operational requests based on newly emerged needs or routine system update needs were not considered safety events unless patient care was impeded in a way that is deemed unsafe. For example, a request to add a new biosimilar medication to the medication library was considered an operational request and was not classified as a patient safety event because safe patient care was not impeded. In another case, a pharmacy technician was unable to verify a medication upon scanning because the system library did not include the national drug code (NDC) of the product being used, even though it represented the correct ingredient and was stocked by the pharmacy. This incident was remedied by updating the system's library to

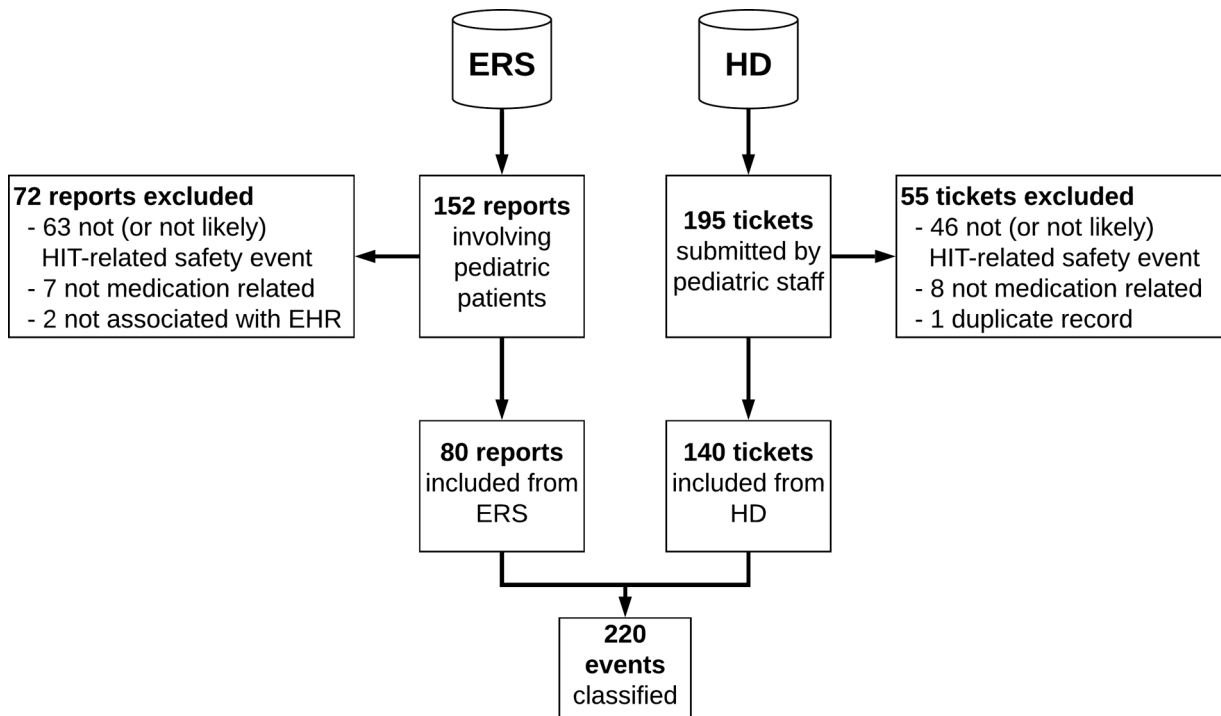


Fig. 1 Record flow for initial pilot application of classification schemes. EHR, electronic health record; ERS, event reporting system; HD, help desk; HIT, health information technology.

include the new NDC and associate it with the medication entry, but the event was considered a patient safety event because it required the pharmacy technician to override an important safety EHR feature during medication preparation.

We further limited our data to medication-related events occurring within the pediatric population (i.e., <18 years of age), as two of the authors have detailed knowledge of the clinical practice in our pediatric hospital. Furthermore, a recent study observed that clinical decision support systems within EHRs fail to prevent one in three potential dangerous medication orders, necessitating novel ways to identify and prevent unsafe situations.²²

Applying the inclusion criteria, we identified 80 ERS reports and 140 HD tickets between August 1, 2018 and December 7, 2018 which served as a derivation dataset used to develop the taxonomy (→Fig. 1). We additionally identified 310 ERS reports and 323 HD tickets between December 8, 2018 and July 31, 2019 which served as the test dataset used to determine the performance characteristics of the final taxonomy. When determining the performance characteristics of the final taxonomy, a subset of the test dataset (one quarter) was reserved for initial reviewer calibration (“calibration subset”), and the remainder was used for the primary performance evaluation (“evaluation subset”).

Evaluation of Candidate Taxonomies

A single reviewer (K.D.W.) applied each classification system to the derivation dataset and performed quantitative and qualitative assessments. Qualitative assessment included a critique of the taxonomy structure, as well as identification of practical challenges that arose in the application of the taxonomy to real-world patient safety events. Quantitative

assessment consisted of assessment of the percentage of events that could be classified using the taxonomy. Events were considered not classifiable if the reviewer could not confidently and unambiguously associate the event with a category within the classification scheme based on the information included in the report/ticket.

Taxonomy Development

A multidisciplinary Patient Safety Subcommittee reviewed the results of the initial round of application to the derivation dataset and consensus was achieved regarding which singular taxonomy to further develop. This taxonomy was revised and piloted again by the same reviewer in a second “plan-do-study-act” (PDSA) cycle with the same set of events and tickets (→Supplementary Fig. S1, available in the online version).

Taxonomy Performance Evaluation

The developed taxonomy was then assessed for interrater reliability and agreement using the test dataset. Two reviewers (K.D.W. and T.J.B.) independently reviewed a new set of ERS reports and HD tickets and applied the taxonomy. During the initial calibration phase, a subset of one quarter of events (80 ERS reports and 79 HD tickets, “calibration subset”) and tickets was reviewed in a blinded fashion and disagreements were discussed to clarify interpretation of the taxonomy categories and identify ambiguities. These ambiguities were used to iteratively update the wording of the taxonomy and facilitate consistent classification. Following this calibration phase, the reviewers evaluated the remaining events (“evaluation subset”) and tickets in a blinded fashion. Interrater reliability and agreement were assessed for inclusion (i.e., whether the ERS report or HD ticket represented a HIT-related safety event) and

for the primary classification category that applied to the event. If reviewers disagreed about event inclusion, it was not included in the interrater analyses pertaining to the primary classification category.

Statistical Analysis

Fischer's exact test was used to assess for differences in the percent of events reported to each reporting source within each category, with two-sided p -values of <0.05 considered significant. Percent agreement was measured and kappa coefficient was used to assess interrater reliability.²³ Statistical analysis was performed using JMP 14.1.0 (2018, SAS Institute, Cary, North Carolina, United States).

Human Subjects Protections

This project was reviewed by the Mayo Clinic Institutional Review Board and classified as quality improvement.

Results

Literature review identified the following four viable taxonomies for considering patient safety events:

- HIT safety domains.^{5,7}
- Sociotechnical model.⁹
- Magrabi's classification.^{15–20}
- HIT-related safety concerns.⁶

The derivation dataset was used to apply each of the taxonomies to real-world patient safety events for the quantitative and qualitative assessment.

Health information technology (HIT) safety domains were proposed as part of a framework that outlines three phases of the EHR implementation and use lifecycle.^{5,7} The first domain—safe HIT—focuses on addressing safety concerns unique to technology. Principles applicable to this domain include data availability, data integrity, and data confidentiality. The second domain—using HIT safely—focuses on optimizing the safe use of technology. Principles include complete/correct HIT use and HIT system usability. The third domain—monitoring safety—focuses on using technology to monitor and improve safety. Its principle is surveillance and optimization.

Sixty-nine of 80 (86%) ERS reports were classifiable using HIT safety domains, and 139 of 140 (99%) HD tickets were classifiable using HIT safety domains. Application of HIT safety domains to the events from both data sources was qualitatively challenging because the categories were relatively nonspecific. Therefore, certain classification categories were associated with a large proportion of events (e.g., 129 of 140 [92%] HD tickets were associated with the principle “complete/correct use of HIT”) thereby limiting the taxonomy's discriminative abilities.

The sociotechnical model for HIT^{9,24,25} emphasizes that HIT operates within the following overlapping and interacting dimensions:

- Hardware and software.
- Clinical context.
- Human–computer interface.

- People.
- Workflow and communication.
- Organizational policies and procedures.
- External rules, regulations, and pressures.
- System measurement and monitoring.

Sixty-nine of 80 (86%) ERS reports were classifiable within the sociotechnical model. All 140 HD tickets (100%) were able to be classified within the sociotechnical model. The percent of events falling within each category is shown in **►Supplementary Table S1** (available in the online version). A larger proportion of HD tickets were related to clinical context and organizational policies and procedures, compared with ERS reports which were more often pertained to workflow and communication (**►Supplementary Table S1**, available in the online version). Qualitative assessment revealed that sociotechnical domains allowed for discrimination of events, but categories did not precisely describe root causes. For example, the “people” dimension refers to the developers who design the system, support staff who implement it, clinicians who use it, and patients. Assignment of this dimension to an event does not distinguish between those various and disparate groups of people involved with implementation and use of the EHR.

The Magrabi classification is a categorization system for computer-related patient safety incidents.¹⁵ These are divided into those that occur at the computer (“machine”) level and those which occur at the computer–human interface. They are further subdivided into errors of input, transfer, output, purely technical issues, and other contributing human factors. This classification system has been used by others in published reports to classify health IT (information technology)-related safety problems,^{15–20} including one report focusing specifically on HIT-related medication errors.¹⁷ Some of these reports identified gaps in the original categories and proposed additions. Newly proposed categories from these studies^{18–20} were compiled to generate a composite version of the classification (**►Supplementary Table S2**, available in the online version). The Magrabi classification has notably been codified in an International Standards Organization (ISO) standard (ISO/TS, 20405).²⁶

Forty-nine of the 80 ERS reports (61%) could be classified using the Magrabi classification, compared with 139 of 140 (99%) HD tickets. The four most common categories from the 49 classifiable ERS reports were misreading or misinterpreting displayed information ($n = 15$, 31%), software functionality ($n = 10$, 20%), failure of computer to alert ($n = 9$, 18%), and wrong input ($n = 9$, 18%). In contrast, the four most common categories from the 139 classifiable HD tickets were software system configuration ($n = 78$, 56%), default values in system configuration ($n = 21$, 15%), staffing/training ($n = 18$, 13%), and misreading or misinterpreting displayed information ($n = 16$, 12%). The number of events classified in each category are shown in **►Supplementary Table S3** (available in the online version).

Qualitative assessment revealed that many of the events described in ERS reports could not be classified based on the user-submitted event narrative alone. In contrast, classification

of events described in HD tickets was facilitated by annotations from HD staff regarding investigation of the problem's root cause and the solution. Furthermore, the number of categories in the Magrabi classification was large (→ **Supplementary Table S2**, available in the online version), resulting in a cumbersome process of classification.

HIT-related safety concerns described by Sittig et al identified five types of HIT-related safety concerns.⁶ This classification overlapped with categories identified by the internal brainstorming by patient safety leadership at our institution (see → **Supplementary Table S4** for a comparison, available in the online version). Given this substantial overlap, the authors engaged in a consensus-building process to merge these two categorizations, aiming to increase ease of classification, and reduce ambiguity. Specifically, HIT-related safety concerns identified by Sittig et al⁶ included one category that was not part of the internally brainstormed scheme (i.e., safety features or functions were not implemented or available). Furthermore, internal brainstorming further subcategorized improper configuration or implementation into (1) faulty process design choices and (2) incorrect or inaccurate setup. The composite taxonomy included all of the categories originally proposed by Sittig et al except that improper configuration or implementation was subdivided into faulty process design choices and incorrect or inaccurate setup to incorporate the locally brainstormed categories. To reduce ambiguity and facilitate ease of classification, the categories were organized into a flowchart (→ **Supplementary Fig. S2**, available in the online version).

Twenty-nine of 80 ERS reports (36%) could be classified using the merged version of the HIT-related safety concerns, compared with all 140 HD tickets (100%, → **Supplementary Table S5**, available in the online version). When ERS reports and HD tickets were compared, a greater proportion of ERS reports pertained to usability issues in contrast to HD tickets, which more frequently pertained to incorrect setup or configuration (→ **Supplementary Table S5**, available in the online version). Qualitative assessment revealed that many of the events described in ERS reports could not be classified based on the user-submitted event narrative alone. In contrast, classification of events described in HD tickets was facilitated by annotations from HD staff regarding investigation of the problem's root cause and the solution.

Taxonomy Development

The quantitative and qualitative assessments of the candidate taxonomies were presented to the Patient Safety Subcommittee as a slideshow and oral presentation along with a proposal to adapt the composite classification that included the HIT-related safety concerns described by Sittig et al⁶ and the categories devised through internal brainstorming. Consensus was achieved among the subcommittee to adopt the recommendation and endorse pursuing the composite classification. Although this classification had the highest percentage of events that could not be classified, the committee valued several key qualitative features of the taxonomy, leading to its adoption. The first feature was its focus on “why” events happened, instead of only “what” happened. For example, it was discussed that the Magrabi classification focused heavily

on “what” occurred, when the root cause (i.e., “why”) was of most interest to patient safety officers hoping to prevent similar events from happening in the future. Moreover, the HIT safety domains were considered overly vague and conceptual for application to individual events. In this sense, although many events could be classified using the taxonomy, the classifications did not point to actionable solutions to mitigate each safety risk on an individual event basis. The third feature was that the number of categories was manageable, compared with other taxonomies such as the Magrabi classification. Finally, presentation of the categories in a flowchart facilitated placement of each event into a category.

In seeking to refine the taxonomy (→ **Supplementary Fig. S2**, available in the online version), the category of “developmental oversight” was identified as potentially problematic because of a lack of consensus regarding what features should be implemented in an EHR. This category was therefore removed from the taxonomy. Because very few events fell into the category of “data exchange failure”—and there was a desire to simplify the taxonomy for end-users—the distinction between hardware and software data exchange failure was removed. Based on qualitative challenges categorizing events and tickets during the previous event review cycle, additional clarifications, which did not alter the categories themselves, were added to the category descriptions to ensure consistency in classification. Finally, there was a desire to further elaborate on ways that usability could present challenges. We identified a manuscript describing a taxonomy for usability issues which was published after our initial literature search performed,²⁷ and the taxonomy was adapted and integrated into the modified health IT-related safety concerns taxonomy. Finally, we added a formal definition for HIT-related safety events²¹ to the flowchart. These changes yielded the revised singular taxonomy shown in → **Fig. 2**.

This modified version of the singular taxonomy was pilot tested using the same set of ERS events and HD tickets used in the initial phase (→ **Fig. 1**). Similar to the earlier version, 30 of 80 (38%) ERS reports and all 140 HD tickets were classifiable. No new qualitative concerns were raised during qualitative assessment. Therefore, we proceeded to further characterize the taxonomy through evaluation of interrater agreement and reliability.

Interrater Agreement and Reliability

The test dataset of 310 ERS reports and 323 HD tickets were used to assess interrater agreement and reliability. Following the calibration phase, we observed a kappa of 0.68 with 87.8% agreement for ERS reports and a kappa of 0.56 with 73.6% agreement for HD tickets (→ **Table 1**).

Review of contingency tables for the calibration subset revealed a large proportion of events that one reviewer classified as “setup incorrect” and other classified as “faulty process/design choices.” There were 13 instances of this discordance among the 61 HD tickets (21%) that were classified during the calibration phase. Following the calibration phase, this conflict persisted as the predominant conflict, representing 29 of 212 (14%) HD tickets.

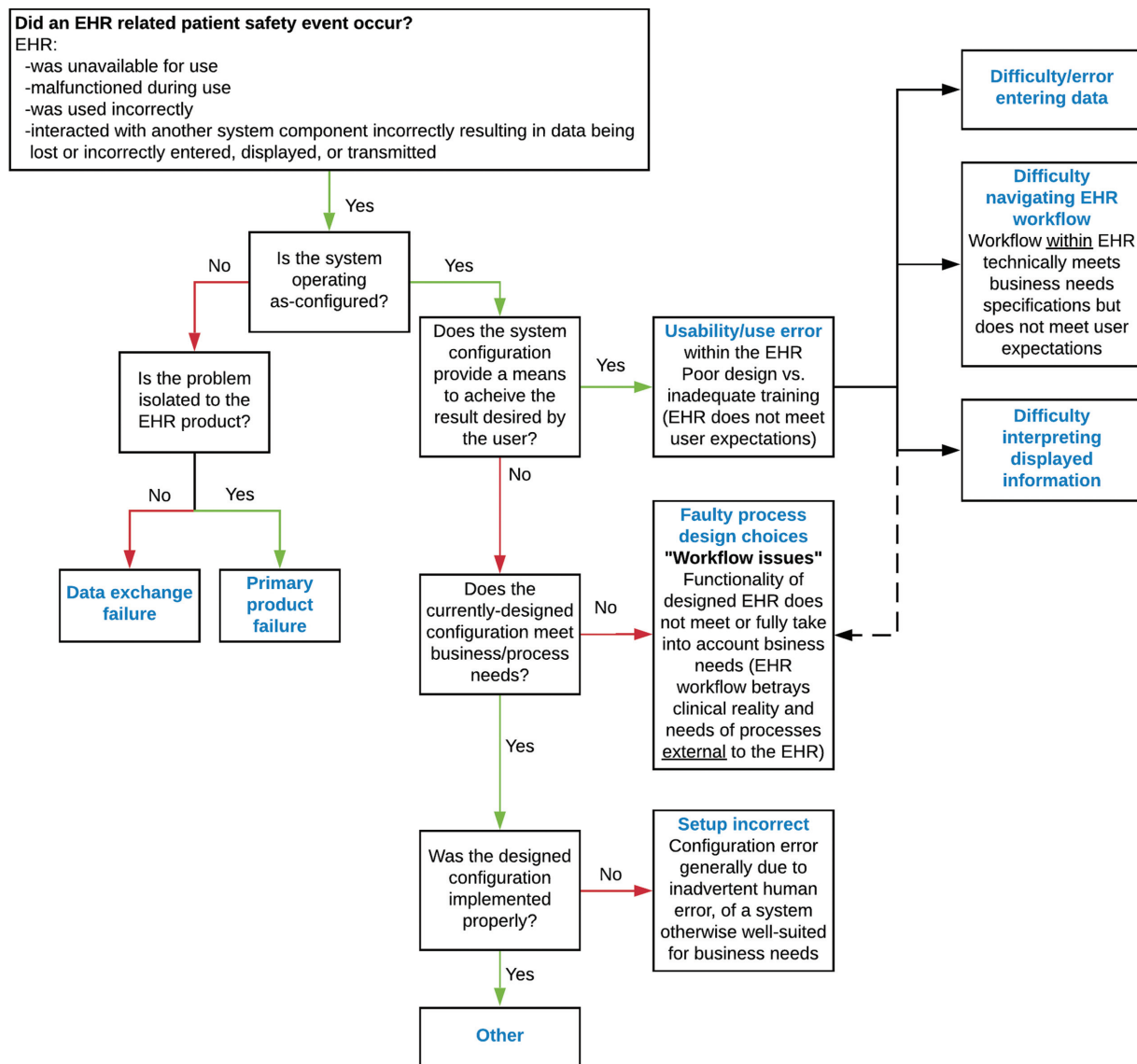


Fig. 2 Revised singular taxonomy. Dashed line to faulty process design choices signifies that usability challenges within the EHR may impair processes external to the EHR. EHR, electronic health record.

Discussion

Summary of Results

We have developed, applied, and assessed a taxonomy for classifying medication-related patient safety events related to HIT in pediatrics obtained from different sources (→ **Fig. 2**). Illustrative examples of events classified in each category are shown in → **Table 2**. The taxonomy was developed following literature review, comparison of existing taxonomies, adaptation of an existing taxonomy, and iterative improvements following pilot implementation using PDSA cycles (→ **Supplementary Fig. S1**, available in the online version). We identified differences in the types of events reported through different mechanisms. Acceptable inter-rater agreement and reliability were observed in the course of applying the taxonomy to real-world safety events. Our ability to apply the taxonomy was limited when the information provided in the narrative report was incomplete.

Comparison to Existing Literature

In addition to the seminal articles describing each of the taxonomies we analyzed, others have developed and applied taxonomies to patient safety events. The Agency for Healthcare Quality and Research (AHRQ) developed a common format for a medical devices and supplies, including HIT.²⁸ Menon et al utilized a composite taxonomy that utilized sociotechnical domains²⁹ as a backbone and also incorporated concepts from the Magrabi classification,¹⁵ the ontology included in the Health IT Hazard Manager³⁰ and the AHRQ Common Formats. They subdivided these errors according to those where EHR did not work at all, those where EHR worked incorrectly, those where EHR technology was missing or absent, and concerns related to “user errors” (nota bene: in our practice, we prefer to refer the latter as “use errors,” as the term “user error” implies the user is wholly at fault, even in cases of poor EHR design.). Similarly, Castro et al³¹ previously classified 120 HIT-related sentinel events

Table 1 Interrater reliability and agreement for singular taxonomy

Data source	Kappa	Kappa 95% confidence interval		Agreement (%)
		Lower bound	Upper bound	
ERS reports				
Calibration subset				
Inclusion (<i>n</i> = 80)	0.702	0.531	0.874	87.5
Classification (<i>n</i> = 19)	0.432	0.192	0.672	52.6
Evaluation subset				
Inclusion (<i>n</i> = 230)	0.606	0.501	0.710	82.1
Classification (<i>n</i> = 47)	0.688	0.581	0.795	87.8
HD tickets				
Calibration subset				
Inclusion (<i>n</i> = 79)	0.413	0.142	0.685	84.8
Classification (<i>n</i> = 61)	0.277	0.138	0.417	44.3
Evaluation subset				
Inclusion (<i>n</i> = 244)	0.688	0.535	0.841	94.3
Classification (<i>n</i> = 212)	0.556	0.462	0.649	73.6

Abbreviations: ERS, event reporting system; HD, help desk.

Table 2 Examples of patient safety events

Category	Example
Data exchange failure	Nurse unable to dispense a medication routed to medication dispensing machine due to discrepancy between stock listed in EHR and in medication dispensing machine database
Primary product failure	Pharmacy medication preparation report generation takes an excessive amount of time
Usability/use error (not otherwise specified)	Chemotherapy orders do not prompt for second signature because system requires user to change context for prompt to appear
Difficulty/error entering data	Provider selected default dose when ordering medication, and the dose was too high
Difficulty navigating EHR workflow	Orders were not “released” by provider within the EHR
Difficulty interpreting displayed information	Staff member did not scroll down completely to see medication administration schedule, and medication dose was missed
Faulty process design choices	Scheduled inpatient medications are not dispensed from pharmacy while patient is in operating room, leading to missed dose
Setup incorrect	Dose rounding parameters for neonates are imprecise (e.g., rounding to nearest 0.1 mL instead of 0.01 mL)

Abbreviation: EHR, electronic health record.

reported to The Joint Commission using a composite classification based on existing classifications including the socio-technical model⁹ and the Magrabi classification.¹⁵ Similarly, they concluded that both the event narrative and formal root cause analysis were important for properly classifying events, and they called for classification of events according to contributing HIT-related factors to improve patient safety.

In a report on computerized prescriber order entry (CPOE) medication safety, a taxonomy entitled “why the error occurred” was proposed with categories that were tailored and specific for medication-related events.³² Similar to our experience, they concluded that determining the root cause was often not possible through event narrative review alone, “... it was

difficult to definitely conclude why the error happened without speaking directly to the person who wrote the report.”³² Using our proposed taxonomy, only 36% of ERS reports were classifiable. In contrast, all of the events reported in HD tickets—which included additional annotations regarding what happened and why—were able to be classified using the taxonomy. This suggested that more patient safety events derived from ERS reports could be classified if additional information above and beyond the user-submitted narrative were considered (e.g., formal root-cause analysis). We therefore viewed the low percentage of ERS reports that were classifiable more as a limitation of the amount of information included in ERS reports rather than a limitation of the classification scheme per se.

Practical Challenges Encountered Applying the Taxonomy

In the course of assessing interrater agreement, the most common disagreement occurred when one reviewer classified an event as “setup incorrect,” and the other reviewer classified it as “faulty process/design choices.” During the calibration phase, the reviewers distinguished between these two event types by describing safety events as either incongruences between designers’ intentions and users’ needs/desires (faulty process/design choice) or as incongruences between designers’ intentions and actual function (setup incorrect, [▶ Supplementary Fig. S3](#), available in the online version). Faulty process/design choices represented workflow decisions that, in hindsight, were suboptimal for some situations and may have unintended adverse consequences (e.g., handling of pharmacy medication dispense logic when a patient moves between inpatient and intra-operative settings). Incorrect setup represented “correct” design choices that were incorrectly implemented, generally due to inadvertent human error during the implementation phase, and which could be easily remedied by reconfiguration (e.g., rounding doses for neonates to the nearest 0.1 mL rather than the nearest 0.01 mL).

The observation that the disagreements between “setup incorrect” and “faulty process/design choices” remained even after the calibration phase suggests that applying the distinction between them in practice may remain a challenge. Reviewer deliberation in the postcalibration phase of review revealed that disagreements often arose due to lack of consensus regarding whether the implemented configuration reflected the faithful implementation of a deliberate, agreed-upon design/configuration (i.e., faulty process/design choices) or a failure of system implementers to correctly configure the system to conform to an agreed-upon configuration design (i.e., setup incorrect). In many cases, the intended design/configuration was not explicitly outlined within the event data sources (i.e., HD tickets and ERS reports) and had to be inferred by the event reviewers. We concluded that these conflicts did not represent a limitation of the taxonomy per se, but instead a limitation of the information available to the reviewers. A higher degree of accuracy and consistency in classification of events is possible if additional information is available during event review.

Implementation Strategies

To fully implement routine usage of this taxonomy within our organization, we would need to determine when to apply it and who should apply it. The inclusion of the taxonomy within the event submission forms would integrate it into the submission workflow at the cost of increased burden on the submitter, which might reduce reporting frequency. Calibration of responses by event reviewers appears to improve consistency of applying the taxonomy, and accurate classification may depend on background knowledge regarding the system’s intended configuration. Therefore, assigning responsibility for classification to a small group of staff with expertise in the study of patient safety events may result in improved interrater reliability and agreement compared with end-user classification.

Approaches to Patient Safety Data Collection

Since a small minority of patient safety events are reported through patient safety event reporting systems,^{33–35} other data sources are needed to understand the full spectrum of HIT-related safety events. The taxonomy we developed has shown utility classifying events from two different sources (i.e., ERS and HD), which increases yield and diversity. Even when events are faithfully reported, identifying those events that are HIT-related remains a challenge.²⁴ Emerging methods for identification of patient safety events include automated data-mining mechanisms^{36–40} and simulation approaches.⁴¹ Data-mining based approaches have the benefit of utilizing data generated as a byproduct of clinical care (e.g., billing codes) to identify patient safety events in “real-time” with little human effort,³⁷ but they are susceptible to bias due to incomplete or inaccurate data.⁴² Simulation-based approaches are limited because they only assess the safety risks included in the simulation and may lack fidelity.

Moreover, simply documenting and classifying HIT-related patient safety events are not sufficient without a mechanism to follow through to improve the system. At our institution, this process is taken on by individual patient safety event staff who investigate root causes and work to implement systems-based changes following a formal analysis of each event and its causes. Innovative approaches to cataloging and prioritizing HIT-related safety risks have been developed. For example, Geisinger Health System developed a system for tracking HIT-related patient safety events under a contract from AHRQ. “Health IT Hazard Manager”³⁰ is a software tool designed to facilitate cataloging of patient safety risks (“near misses”). The software walks reporters through the process of describing an event, identifying the involved systems, describing how the risk was discovered, explaining the causal factors leading to the risk, specifying the impact of the problem, and developing a plan to mitigate the problem. The tool incorporates an ontology to describe the root causes of events, which overlaps in many ways with the classifications we reviewed in the process of developing the singular taxonomy.

Utility of Patient Safety Event Taxonomies

One advantage of classifying HIT-related patient safety events is that changes in the frequency of specific event types can be used to measure the impact of quality improvement projects. However, trends in event reporting must be interpreted with caution because 10% or fewer of events are reported through traditional reporting mechanisms and there is substantial risk of reporting bias.^{33,35} Moreover, event reporting may be susceptible to reporting bias in specific departments or by employees with certain work roles, and this may influence event reporting patterns. For example, our anecdotal experience was that a large proportion of HD tickets related to functions used by pharmacy staff, as the HD served as a formal process for documenting and requesting changes to the EHR to facilitate safe patient care. Additionally, patient safety events and near misses are frequently caught by members of the care team who are on the distal end of the medication administration process,

namely, bedside nurses. These factors likely influenced the pattern of event types observed. Collection of data on submitter department and work role may identify biases in event reporting patterns.

Recently, institutions have shown increased interest in pooling patient safety event data for peer benchmarking.^{11–14} There are several barriers to openly and transparently sharing data, including concerns regarding litigation and impact on an institution's reputation.⁴³ Patient safety experts are increasingly calling for an open and collaborative approach to reducing the incidence of patient harm, since solutions will require collaboration between end users, organizations, and EHR vendors.^{13,43} We propose that open sharing of patient safety data regarding HIT-related patient safety events—using a uniform and standards-based taxonomy—can help individual institutions learn the best practices for EHR implementation and inform EHR vendors about problematic aspects of their software. In the current state of affairs, “hold harmless” clauses—which limit the liability of EHR vendors—may disincentivize EHR vendors from engaging in collaborative patient safety efforts.⁴⁴ Furthermore, contractual prohibitions on sharing of screenshots imposed by EHR vendors may further limit public discourse about unsafe EHR features and workflows at academic conferences and in the medical literature.³

Strengths and Limitations

This study highlights the steps needed to develop a versatile taxonomy that can both meet the needs of individual centers and be used across institutions. Specifically, we performed a literature review, synthesized existing taxonomies, pilot tested using real-world data, presented the results to local content experts, developed a refined taxonomy based on event review data and consensus building among content experts, and assessed the taxonomy's performance characteristics. Strengths of this study include that we pilot tested the application of four classifications to real-world patient safety events derived from two separate sources within our institution. We refined a singular taxonomy and performed quantitative and qualitative assessments of performance, including interrater reliability and agreement.

The major limitation is that we only piloted the taxonomy to classify medication-related events in the pediatric population at a single institution. The taxonomy may not generalize to the experience at other health care institutions, and it may not translate to nonmedication-related events. Future assessments should aim to validate the use of the taxonomy to classify different types of events (e.g., laboratory- or radiology-related errors), and those occurring in different populations and at different institutions. Such validation would identify whether the taxonomy can be applied to these different types of events. This is important because certain root causes may be more common contributors in different types of events. For example, interoperability and data exchange failure issues may be more common with laboratory- and radiology-related errors.⁴⁵ If the taxonomy is not suitable for these other types of events, the validation process would ideally provide insights into how the taxonomy can be modified to the best classify those events.

Another significant limitation is that the taxonomy was piloted during an assessment of events where a causal chain of direct contribution from the EHR could be established. In many cases, the contribution of the EHR to patient safety events can be subtle and easily overlooked by the casual observer. For example, misfiring of certain clinical decision support rules can contribute to “alert fatigue,” resulting in provider burnout or habitual overriding of clinical decision support alerts without reading them, which can contribute to errors in subsequent encounters.⁴⁶ As we have highlighted, EHRs operate within a complex sociotechnical system. Errors, such as delayed diagnosis and misdiagnosis, while often attributed to cognitive biases, can be influenced by the EHR when the information necessary to make a diagnosis is not readily available to the provider or it is presented in a format that is difficult to understand. Although the taxonomy we present could be applied to these types of events, the causal link is not always evident, and methods to expose the contribution of the EHR to these types of events are needed.

Finally, due to limited resources interrater agreement and reliability were not assessed for the derivation dataset. The study would have been strengthened if these statistics could have been included in the assessment of the individual taxonomies and informed the development of a singular taxonomy.

Conclusion

The EHR is a well-recognized contributor to serious patient harm. Contributing factors include suboptimally developed workflows, poorly designed interfaces, and misconfiguration. The full spectrum of patient safety events is the best understood through assessment of data from multiple sources using a uniform taxonomy. We have proposed such a taxonomy and demonstrated its application to pediatric medication-related safety events.

Clinical Relevance Statement

Implementation of a standardized taxonomy for HIT-related patient safety events can support quality improvement efforts by facilitating identification and trending of specific event types. We present a taxonomy that has been applied to real-life HIT-related patient safety events gathered from multiple data sources.

Multiple Choice Questions

1. Which of the following advantages do automated data-mining-based approaches to health information technology related patient safety event categorization have over manual human-directed approaches?
 - a. Improved accuracy
 - b. Improved sensitivity
 - c. Improved specificity
 - d. Easier to integrate into existing routines

Correct Answer: The correct answer is option d. Easier to integrate into existing routines. Data-mining-based approaches utilize data generated in the course of routine clinical care, such as billing data or specific sequences of computer commands (e.g., immediate canceling and reordering of a new order) to identify patient safety events. One advantage is that these approaches can analyze a large amount of data to automatically identify patient safety events without the need for human intervention, such as the labor-intensive process of manually reviewing each event. However, one disadvantage is that these approaches may yield inaccurate results due to incomplete or missing data, or the data may not provide a sufficient basis upon which to conclude whether a patient safety event occurred. While these methods may be more sensitive for specific types of errors (e.g., order cancellation and reordering), this cannot be generalized to all types of patient safety events.

2. In the present study of patient safety events submitted through the information technology “help desk” (HD) and a confidential patient safety event reporting system (ERS) using a refined taxonomy, how did events submitted to the HD compare with events submitted to the ERS?
 - a. They were more likely to pertain to usability.
 - b. They were more likely to pertain to software configuration.
 - c. They were more likely to pertain to problematic process/design choices.
 - d. There were no significant differences between the types of events submitted to the two data sources.

Correct Answer: The correct answer is option b. They were more likely to pertain to software configuration. In an analysis of 140 reports submitted to the HD and 80 reports submitted to the ERS, 63% of reports to the HD pertained to incorrect setup/configuration, compared with only 8% of reports to the ERS. In contrast, 37% of reports to the ERS pertained to usability problems, compared with 20% of reports to the HD. Problematic process/design choices accounted for a similar proportion of events submitted to both sources.

Conflict of Interest

None declared.

Acknowledgments

The authors thank the Patient Safety Subcommittee and patient safety staff at our institution for providing expertise and support during development of the taxonomy. In particular, we thank Jessica Nordby, Joseph Nienow, Carina Welp, and Bridget Griffin for their insight and guidance. We also thank Ryan Pross for his assistance extracting patient safety events from the HD ticket database.

Protection of Human and Animal Subjects

This project was reviewed by the Mayo Clinic Institutional Review Board and classified as quality improvement.

References

- 1 Varghese J, Kleine M, Gessner SI, Sandmann S, Dugas M. Effects of computerized decision support system implementations on patient outcomes in inpatient care: a systematic review. *J Am Med Inform Assoc* 2018;25(05):593–602
- 2 Brown CL, Mulcaster HL, Triffitt KL, et al. A systematic review of the types and causes of prescribing errors generated from using computerized provider order entry systems in primary and secondary care. *J Am Med Inform Assoc* 2017;24(02):432–440
- 3 Institute of Medicine. Health IT and Patient Safety: Building Safer Systems for Better Care. Available at: <https://essentialhospitals.org/wp-content/uploads/2014/07/IOM-report-on-EHR-and-Safety.pdf>. Accessed September 8, 2020
- 4 Prgomet M, Li L, Niazkhani Z, Georgiou A, Westbrook JI. Impact of commercial computerized provider order entry (CPOE) and clinical decision support systems (CDSSs) on medication errors, length of stay, and mortality in intensive care units: a systematic review and meta-analysis. *J Am Med Inform Assoc* 2017;24(02):413–422
- 5 Singh H, Sittig DF. Measuring and improving patient safety through health information technology: The Health IT Safety Framework. *BMJ Qual Saf* 2016;25(04):226–232
- 6 Sittig DF, Classen DC, Singh H. Patient safety goals for the proposed Federal health information technology safety center. *J Am Med Inform Assoc* 2015;22(02):472–478
- 7 Sittig DF, Ash JS, Singh H. The SAFER guides: empowering organizations to improve the safety and effectiveness of electronic health records. *Am J Manag Care* 2014;20(05):418–423
- 8 Sittig DF, Singh H. Electronic health records and national patient-safety goals. *N Engl J Med* 2012;367(19):1854–1860
- 9 Sittig DF, Singh H. A new sociotechnical model for studying health information technology in complex adaptive healthcare systems. *Qual Saf Health Care* 2010;19(Suppl 3):i68–i74
- 10 Kim MO, Coiera E, Magrabi F. Problems with health information technology and their effects on care delivery and patient outcomes: a systematic review. *J Am Med Inform Assoc* 2017;24(02):246–250
- 11 Phipps AR, Paradis M, Peterson KA, et al. Reducing serious safety events and priority hospital-acquired conditions in a pediatric hospital with the implementation of a patient safety program. *Jt Comm J Qual Patient Saf* 2018;44(06):334–340
- 12 Lyren A, Brillli RJ, Zieker K, Marino M, Muething S, Sharek PJ. Children's hospitals' solutions for patient safety collaborative impact on hospital-acquired harm. *Pediatrics* 2017;140(03):e20163494
- 13 Lannon CM, Peterson LE. Pediatric collaborative networks for quality improvement and research. *Acad Pediatr* 2013;13(6, suppl):S69–S74
- 14 Martin BS, Arbore M. Measurement, standards, and peer benchmarking: one hospital's journey. *Pediatr Clin North Am* 2016;63(02):239–249
- 15 Magrabi F, Ong MS, Runciman W, Coiera E. An analysis of computer-related patient safety incidents to inform the development of a classification. *J Am Med Inform Assoc* 2010;17(06):663–670
- 16 Palojoki S, Mäkelä M, Lehtonen L, Saranto K. An analysis of electronic health record-related patient safety incidents. *Health Informatics J* 2017;23(02):134–145
- 17 Cheung KC, van der Veen W, Bouvy ML, Wensing M, van den Bemt PM, de Smet PA. Classification of medication incidents associated with information technology. *J Am Med Inform Assoc* 2014;21(e1):e63–e70
- 18 Warm D, Edwards P. Classifying health information technology patient safety related incidents - an approach used in Wales. *Appl Clin Inform* 2012;3(02):248–257
- 19 Magrabi F, Ong MS, Runciman W, Coiera E. Using FDA reports to inform a classification for health information technology safety problems. *J Am Med Inform Assoc* 2012;19(01):45–53

- 20 Sparnon E, Marella WM. The role of the electronic health record in patient safety events. *Pennsylvania Patient Safety Adv* 2012;9(04):113–121
- 21 Sittig DF, Singh H. Defining health information technology-related errors: new developments since to err is human. *Arch Intern Med* 2011;171(14):1281–1284
- 22 Classen DC, Holmgren AJ, Co Z, et al. National trends in the safety performance of electronic health record systems from 2009 to 2018. *JAMA Netw Open* 2020;3(05):e205547
- 23 Kottner J, Audigé L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *J Clin Epidemiol* 2011;64(01):96–106
- 24 Kang H, Wang J, Yao B, Zhou S, Gong Y. Toward safer health care: a review strategy of FDA medical device adverse event database to identify and categorize health information technology related events. *JAMIA Open* 2018;2(01):179–186
- 25 Wang J, Liang H, Kang H, Gong Y. Understanding Health Information Technology Induced Medication Safety Events by Two Conceptual Frameworks. *Appl Clin Inform* 2019;10(01):158–167
- 26 ISO/TS 20405: Health informatics—Framework of event data and reporting definitions for the safety of health software. available at: <https://standards.iteh.ai/catalog/standards/iso/68edd25c-a3b2-419d-9254-a0ff3cfb92bf/iso-ts-20405-2018>. Accessed September 8, 2020
- 27 Ratwani RM, Savage E, Will A, et al. Identifying electronic health record usability and safety challenges in pediatric settings. *Health Aff (Millwood)* 2018;37(11):1752–1759
- 28 Common Formats for Event Reporting—Hospital Version 1.2. Available at: https://www.psoppc.org/psoppc_web/publicpages/commonFormatsV1.2. Accessed September 8, 2020
- 29 Menon S, Singh H, Giardina T, et al. Safety huddles to proactively identify and address electronic health record safety. *J Am Med Inform Assoc* 2017;24(02):261–267
- 30 Walker JM, Hassol A, Bradshaw B, Rezaee ME. Final report: Health IT Hazard Manager Beta-Test: Agency for Healthcare Research and Quality. Available at: <https://digital.ahrq.gov/sites/default/files/docs/citation/HealthITHazardManagerFinalReport.pdf>. Accessed September 8, 2020
- 31 Castro GM, Buczkowski L, Hafner JM. The contribution of socio-technical factors to health information technology-related sentinel events. *Jt Comm J Qual Patient Saf* 2016;42(02):70–76
- 32 Food and Drug Administration. Computerized prescriber order entry medication safety (CPOEMS): uncovering and learning from issues and errors. Available at: <https://www.fda.gov/drugs/medication-errors-related-cder-regulated-drug-products/computerized-prescriber-order-entry-medication-safety-cpoems>. Accessed August 4, 2020
- 33 Classen DC, Resar R, Griffin F, et al. 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. *Health Aff (Millwood)* 2011;30(04):581–589
- 34 Evans SM, Berry JG, Smith BJ, et al. Attitudes and barriers to incident reporting: a collaborative hospital study. *Qual Saf Health Care* 2006;15(01):39–43
- 35 Sari AB, Sheldon TA, Cracknell A, Turnbull A. Sensitivity of routine system for reporting patient safety incidents in an NHS hospital: retrospective patient case note review. *BMJ* 2007;334(7584):79
- 36 Sammer C, Miller S, Jones C, et al. Developing and evaluating an automated all-cause harm trigger system. *Jt Comm J Qual Patient Saf* 2017;43(04):155–165
- 37 Kuklik N, Stausberg J, Amiri M, Jöckel KH. Improving drug safety in hospitals: a retrospective study on the potential of adverse drug events coded in routine data. *BMC Health Serv Res* 2019;19(01):555
- 38 Fong A, Komolafe T, Adams KT, Cohen A, Howe JL, Ratwani RM. Exploration and initial development of text classification models to identify health information technology usability-related patient safety event reports. *Appl Clin Inform* 2019;10(03):521–527
- 39 Kalendarian E, Obadan-Udoh E, Yansane A, et al. Feasibility of electronic health record-based triggers in detecting dental adverse events. *Appl Clin Inform* 2018;9(03):646–653
- 40 Feng C, Le D, McCoy AB. Using electronic health records to identify adverse drug events in ambulatory care: a systematic review. *Appl Clin Inform* 2019;10(01):123–128
- 41 Kilbridge PM, Welebob EM, Classen DC. Development of the Leapfrog methodology for evaluating hospital implemented inpatient computerized physician order entry systems. *Qual Saf Health Care* 2006;15(02):81–84
- 42 Horsky J, Drucker EA, Ramelson HZ. Accuracy and completeness of clinical coding using ICD-10 for ambulatory visits. *AMIA Annu Symp Proc* 2018;2017:912–920
- 43 Leape L, Berwick D, Clancy C, et al; Lucian Leape Institute at the National Patient Safety Foundation. Transforming healthcare: a safety imperative. *Qual Saf Health Care* 2009;18(06):424–428
- 44 Koppel R, Kreda D. Health care information technology vendors' "hold harmless" clause: implications for patients and clinicians. *JAMA* 2009;301(12):1276–1278
- 45 Adams KT, Howe JL, Fong A, et al. An analysis of patient safety incident reports associated with electronic health record interoperability. *Appl Clin Inform* 2017;8(02):593–602
- 46 Aaron S, McEvoy DS, Ray S, Hickman TT, Wright A. Cranky comments: detecting clinical decision support malfunctions through free-text override reasons. *J Am Med Inform Assoc* 2019;26(01):37–43