MEDINFO 2017: Precision Healthcare through Informatics
A.V. Gundlapalli et al. (Eds.)

doi:10.3233/978-1-61499-830-3-1048

Identifying and Synchronizing Health Information Technology (HIT) Events from FDA Medical Device Reports

Hong Kang[a], Frank Wang[b], Sicheng Zhou[a], Qi Miao[c], Yang Gong[a]

[a]*School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, USA,*
[b]*College of Natural Sciences, The University of Texas at Austin, Austin, TX, USA,*
[c]*School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, USA*

Abstract

Health information technology (HIT) events, a subtype of patient safety events, pose a major threat and barrier toward a safer healthcare system. It is crucial to gain a better understanding of the nature of the errors and adverse events caused by current HIT systems. The scarcity of HIT event-exclusive databases and event reporting systems indicates the challenge of identifying the HIT events from existing resources. FDA Manufacturer and User Facility Device Experience (MAUDE) database is a potential resource for HIT events. However, the low proportion and the rapid evolvement of HIT-related events present challenges for distinguishing them from other equipment failures and hazards. We proposed a strategy to identify and synchronize HIT events from MAUDE by using a filter based on structured features and classifiers based on unstructured features. The strategy will help us develop and grow an HIT event-exclusive database, keeping pace with updates to MAUDE toward shared learning.

Keywords:

Patient Safety; Medical Errors; Information Storage and Retrieval

Introduction

Health information technology, or Health IT (HIT) -including electronic health records (EHR), computerized provider order entry (CPOE), clinical decision support (CDS) systems, and personal health records (PHR) - has been placed in the spotlight for its bold promises to increase hospital efficiency, improve patient safety, and reduce medical errors [1]. According to the Agency for Healthcare Research and Quality (AHRQ), HIT is defined as "the use of information and communication technology in healthcare to support the delivery of patient or population care or to support patient self-management" [2]. While the use of HIT presents many new opportunities to improve patient care and safety, it can also create new hazards and opportunities for error. HIT will fulfill its potential only if the risks associated with its use are identified and a coordinated effort is developed to mitigate those risks. In fact, one in every six patient safety events (PSE) can be attributed to HIT [3], making HIT one of the top 10 technology-related hazards as identified by the Emergency Care Research Institute [4].

Event reporting is a potential approach for shared learning, which has been proven in many high-risk industries, such as the aviation, nuclear, and railroad industries. For HIT to become more widely adopted and trusted, it is crucial to gain a better understanding of the nature of the errors and adverse events caused by current HIT systems. Identifying problems of HIT systems when they occur and presenting to stakeholders is an area that has not received enough emphasis [5]. Similar to PSE reporting, an effective means suggested by the Institute of Medicine, acquiring knowledge from previous experiences to prevent the recurrence and serious consequences of similar HIT events could be a practical start [6]. PSE reporting is frequently a mainstay of frontline practitioners' efforts to detect PSEs and quality problems [7]. The reports collected from a broad range of stakeholders can generate a summary and feedback toward actionable knowledge and shared learning.

Although AHRQ has taken initial steps to standardize HIT reporting by the creation of the Common Formats [8], currently there is still no HIT exclusive reporting form/system. Part of the reason may lie in the fact that current definitions of HIT are often broad with ill-defined borders [9]. Under the current AHRQ HIT definition, almost any medical device that utilizes electronic software and hardware can be considered an HIT device. In practice, however, reporters may find a stricter definition of HIT more helpful in determining the events involving technology and should be reported under HIT. The starting point for creating an effective HIT exclusive reporting form/system lies in the creation of an HIT event database. The database will allow researchers to analyze connections among HIT events, identify common themes of technology-induced errors, synchronize HIT related reports from existing resources, and develop a classification or terminology to standardize HIT reporting.

Figure 1– Trend of new MAUDE reports per year (bars) and MAUDE related publications per year (line) since 2000. Search resources for peer reviewed publications included three databases: Ovid MEDLINE, PsycINFO, and Health and Psychological Instruments.

The scarcity of HIT event-exclusive databases and event reporting systems indicates the challenge of identifying and synchronizing the HIT events from existing resources. The FDA Manufacturer and User Facility Device Experience (MAUDE) database [10] is a rich and publicly accessible resource of HIT events. The MAUDE database contains the

reports of events involving medical devices, voluntary reports of medical device malfunction, and reports of problems leading to serious injury and death since June 1993. The MAUDE database houses medical device reporting submitted to the FDA by mandatory reporters (manufacturers, importers and device user facilities) and voluntary reporters (healthcare professionals, patients and consumers). MAUDE data are updated weekly and searchable online. As of November 2016, MAUDE had more than 5 million reports. Although the number of MAUDE-based HIT publications has increased in recent years (Figure 1), the number remains fractional given the enormous number of reports in MAUDE. Most publications were reviews about certain events, devices, or treatments based on MAUDE reports. Exploration of the utility of MAUDE for understanding HIT problems has been limited [11]. Therefore, a search and classification strategy holds potential for using the HIT reports and would result in a database to store, manage, and compare HIT reports and to identify and analyze HIT solutions [12].

In this study, we created a comprehensive search strategy that utilizes both structured (device data) and unstructured (text data) to extract HIT related events from the FDA MAUDE database. Our strategy first involved the use of a keyword filter on the structured data to screen the FDA database for reports related to HIT. Then machine learning algorithms were used to classify the selected reports based on their narrative text. Finally, reports classified with a high probability of HIT were added to our HIT event database. Using this strategy, we were able to grow an HIT event database composed of HIT related events from the FDA MAUDE database. Ultimately, the creation of this database holds promise in aiding the understanding, characterization, discovery, and reporting of HIT related events.

Methods

Construct a Filter for HIT Related Reports Based on Structured Data

The device data of FDA MAUDE consists of 45 structured fields that contain information regarding the device involved in the event. Most of the fields, however, are either left blank by reporters or are of little use for the purpose of identifying and synchronizing HIT related events. After reviewing all the fields, we found that the *generic name* and *manufacturer name* fields have the greatest potential in identifying HIT related events and thus we utilized both in the creation of our filter.

To develop the filter, we started with a set of common computer hardware and software related keywords that had been previously identified [13]. The starting keyword list was expanded by the addition of several generic terms such as "software," "program," and "hardware" and several more modern terms such as "electronic medical record" and "portal technology." Then all of the generic names from the FDA MAUDE database starting from Jan 2010 to Dec 2015 were extracted, yielding a total of 60,000 unique generic names. Next, through partial keyword matching of the generic names to the keywords list, a subset of generic names appearing HIT-related was extracted. The subset was further analyzed by domain experts to determine which generic names were actually linked to HIT reports by using a small portion of the 2015 FDA MAUDE database.

A similar approach was utilized in selecting a list of manufacturers for the filter. We started with a list of seven popular HIT manufacturers [13] and then added 347 manufacturers of HIT software. The reports from the 2015 FDA MAUDE database related to those manufacturers were extracted and checked to determine the most relevant manufacturers.

Evaluate the HIT Filter Through Expert Review

The HIT filter was first applied on the 2015 FDA MAUDE database. Then a subset (10%) of the reports screened by the filter was selected for manual review by two domain experts. The experts labeled each report with one of three labels: HIT, Not HIT, or Unsure. The reports that the reviewers disagreed on or were unsure about were resolved through group discussion.

During reviewing, we narrowed the HIT definition to identify the most clinically relevant and consequential HIT devices. Under our current understanding, an HIT device is any device that utilizes both hardware and software to facilitate health information exchange in order to aid in the diagnosis, treatment, or prevention of disease. Using this definition, priority is given to HIT systems that focus on information exchange such as electronic health records, computerized physician order entry, and picture archiving communications systems. Implantable devices, glucose monitors, defibrillators, and similar devices are excluded under this definition as they do not actively facilitate health information exchange.

Construct HIT Classifiers Based on Unstructured Data

Machine learning classifiers including logistic regression, support vector machine (SVM), naïve Bayes, and random forest were constructed using the unstructured data (narrative text) of the reviewed reports. Each report in the labeled training set was treated as a vector of words and was weighted by a term frequency-inverse document frequency (TF-IDF) schema. Each classifier was evaluated using leave-one-out cross-validation (LOOCV) and performance was weighted based on both their F1 score and ROC curve. The best classifier was selected and then attribute selection was done to further increase the performance and efficiency of the classifier.

Grow an HIT Event Database

The optimized HIT classifier was applied to previous years of the FDA Maude database starting from 2015. The reports were ranked in order of their probability of being HIT related. The probability was calculated internally by the classifier. The ranked reports were then manually reviewed until less than 90% of the reports at a given probability threshold were HIT related. The reports above the probability threshold were then added to the rest of the HIT reports previously found, forming the final HIT event database.

Results

Keywords of the HIT Filter

A subset of 336 software and 749 hardware generic names most likely to be related to HIT was extracted from the 3,634,879 reports in MAUDE during 2010-2015, which account for 72% of the total reports in MAUDE since 1993. After the expert review, 58 keywords from *generic names* (39 software and 19 hardware keywords respectively) and 16 keywords from *manufacturer names* were determined to compose the filter for HIT related reports, as shown in Table 1.

Table 1– Keywords of the HIT filter in alphabetical order

58 keywords from *generic names*
Software (39)
ADC, Alert, Automated Dispensing Cabinet, Communication Device, Communication System, CPOE, Data Backup, Database, Decision Support, Digital, Dispensing System, Dose Suggestion, Downloader, Drug Suggestion, EHR, Electronic Heath, Electronic Medical, Electronic Patient, EMR, ICT, Imaging System, Information System, Internet, Invision, LIS, Management System, Monitoring System, Network, Order Entry, PACS, Picture Archiving, Portal, Powerchart, Program, Server, Soarian, Telemetry, Trima Accel Platelet, Web
Hardware (19)
Anesthesia Monitor, Apnea Monitor, Arterial Monitor, Atlas Monitor, Blood Pressure Monitor, Central Monitor, Computer, Console Monitor, Drug Screen, Fetal Monitor, Patient Monitor, Physiological Monitor, Pressure Monitor, PT Monitor, Safety Monitor, Telemetry Monitor, Telemetry Transmitter, Vital Sign Monitor, Workstation
16 keywords from manufacturer names
Allscripts, Centricity, Cerner, Epic, Ge Healthcare, Hass, Healthtronics, Henry Schein, Homer, Isite, iSOFT, Kestral, McKesson, Medical Director, MedPro, Oasis

The filter was first applied on the 2015 FDA MAUDE database including 860,915 reports. 4871 reports (2479 software and 2392 hardware reports) were initially found. 490 reports (10%) were randomly selected according to the keyword distribution for expert review and labeling. 312 reports were identified as HIT related by experts, which means the filter can generate a report subset from original MAUDE database with about 63.7% HIT related reports. This proportion is significantly higher than 0.1%, which is the estimated proportion of HIT related reports in the entire MAUDE database [12].

HIT Classifiers

TF-IDF models were applied on the narrative data to further identify HIT events after using the filter based on the structured data. We used the manually labeled reports (312 HIT and 178 non-HIT reports) from the 2015 MAUDE database to build the training set. The same number of non-HIT-related reports was randomly selected to compose the training data with the HIT-related reports. The classifiers were trained using four methods: logistic regression, random forest, naïve Bayes, and SVM. As shown in Table 2, logistic regression and SVM exhibit better performances among the four methods.

Table 2– HIT classifier performance

Methods	Accuracy	F-Measure	ROC
Logistic Regression	88.6%	0.885	0.919
Random Forest	83.5%	0.835	0.934
Naïve Bayes	84.8%	0.846	0.946
SVM	88.6%	0.886	0.886

The Growing HIT Event Database

Due to the enormous size of MAUDE and the small percentage of HIT events (about 0.1%), directly identifying and extracting all HIT events from MAUDE are almost impossible using a straightforward strategy. The classification of data with imbalanced class distribution has encountered a significant drawback of the performance attainable by most standard classifier learning algorithms [14]. The aim of this study is to establish a model which can identify and synchronize HIT related events from MAUDE database toward an HIT exclusive database for shared learning. Therefore, we need to keep the False Positive (FP) rate within a low value to make sure the quality of the database, even additional HIT reports are missing (high False Negative (FN) rate). Random forest model can help reduce FP rate by trading in FN rate. The strategy was setting a threshold to the confidence coefficient and only taking the samples whose prediction confidence coefficients were higher than the threshold. As shown in Figure 2, 0.7 could be an appropriate threshold since more than 90% FP samples are excluded with a loss of less than a half HIT related reports.

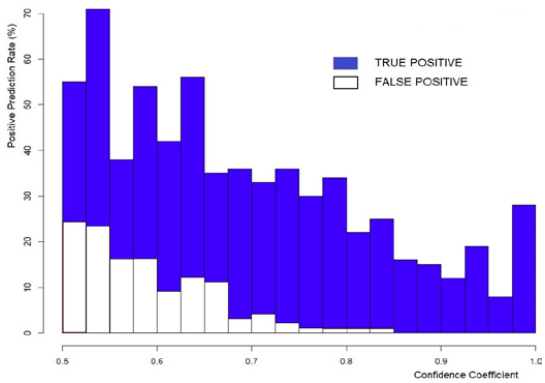

Figure 2– TP and FP distribution of the random forest model

To grow a database for HIT event reports and further improve the HIT filter and classifiers, we are applying the HIT event filter and classifiers on the 2014 dataset. Only the reports that are identified as HIT will be manually reviewed. All the reviewed reports will be added to the training set, which will consequently help build new classifiers. The same process will, in turn, be retrospectively applied on the datasets from 2013, 2012, 2011, 2010, and beyond (Figure 3).

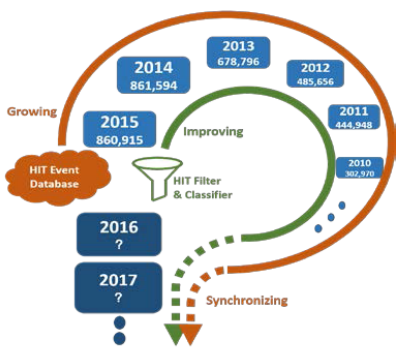

Figure 3– The strategy of improving the HIT filter and classifier and growing the HIT event database

At the end of each iteration, the four training methods will be re-evaluated based on the corresponding manual review, and the best method will be used in the classifier. The classifier is expected to be improved as more labeled reports are included. As the manually reviewed HIT events are accumulated, an HIT event database will be established and keep growing. This mechanism will keep synchronizing HIT related events from MAUDE database to the HIT database in the future.

Discussion

Offer a Broader and Organized View of HIT Events

Some HIT events may seem trivial but could represent much larger and more important problems, which is similar to a tip of a very large and dangerous iceberg. We developed a mechanism of identifying and synchronizing HIT related events from FDA MAUDE database. The outcome is a timely reflection of the evolvement of HIT events and is helpful for enriching HIT

knowledge and better using the historic reports toward an overall understanding and analysis of the characteristics, occurrence, observation, and description of HIT events. Using our proposed filter and classifiers, fragmented and isolated HIT events could be better understood when the connection with other relevant events are offered.

Challenges of Identifying HIT Events

While machine learning has been employed successfully in many contexts such as spam filtering and social media sentiment analysis, its application to HIT event identification is still challenging. Early methods of HIT event identification relied on flagging cases that contained certain technology related keywords for further human review. While these approaches were feasible on a small scale, the growing use of technology in healthcare has led to a rapid rise in technology related events. Meanwhile, the events involving technology have grown in complexity, making researchers sometimes feel difficult to decipher whether or not the technology involved was truly an integral part of the event. Although various natural language processing methods exist such as N-grams and concordance to aid in textual understanding, the classification of PSEs based solely on narrative text may still be out of the reach of current machine learning techniques.

As a result, we utilized both the structured and unstructured data of the FDA MAUDE database. In the device data reports, the two most useful fields for creating our HIT filter were generic and manufacturer name. The generic name field was especially informative as it allowed for the quick identification of device concerning each report. In some cases, the generic name field was pivotal to understanding the event as the device of interest could not be deduced from the narrative reports alone. The manufacturer name field is also proven useful as in some cases, the generic name may have referred to a new or more sophisticated product that was not yet recognized by our list of previously identified generic names. By utilizing both fields, the proposed HIT filter was able to capture more potential HIT events than it would have if only the generic name field were utilized.

Quickly Changing Nature of HIT

Each year new generic names are created and some older names fall into disuse in MAUDE. The number of generic names has shown a growing trend, with each year starting from 2010 containing approximately 1000 more terms than the previous year. This challenge was addressed during the selection of our generic name keywords by reducing the complexity of the generic names while still allowing them to retain their specific meanings. As an example, if the generic name originally found in the database was “picture archiving communications system,” the keyword added to our final list was “picture archiving.” By reducing the complexity of the keywords in our filter, the ability to capture new variations of the same generic name in the future was greatly improved. Furthermore, as previously mentioned, the manufacturer name field was used to supplement the generic names. Cases with new and unrecognized devices could still be detected by the HIT filter if they were made by a recognized HIT manufacturer.

Classifier Error Analysis and Optimization

Two domain experts independently reviewed the filtered 2015 HIT reports and had an inter agreement kappa score of 0.9, which was used as a proxy for the maximum performance that the classifier could reasonably achieve. The initial classifier tended to make more false negatives (true HIT cases labeled as non-HIT) than false positives (non-HIT cases labeled as HIT) in error analysis. In improving the classifier, greater emphasis was placed on reducing the false negatives than the false positives. This is because in practice, reports that are falsely labeled as being HIT can usually be disregarded by users of the database with little effort. Conversely, true HIT reports labeled as non-HIT may pose a much greater risk as users may be unable to learn from those events. By using the proposed approach, while our classifier’s overall accuracy may have been reduced, the classifier we built has a greater ability to capture new and unique HIT safety events.

Evaluation of Proposed HIT Extraction and Synchronization Method

The method we have developed for extracting and synchronizing HIT reports from the FDA MAUDE database shows promise in staying up to date with future changes in HIT. As new devices and technologies are emerging every year, HIT event database faces the risk of quickly becoming outdated. By utilizing an easily updatable device data filter, new types of HIT can be identified and added to the existing database, providing users with the most up-to-date information on HIT. Furthermore, the text classifier we have built allows for the timely classification of HIT related events, greatly reducing the need for future human labor to maintain the database. With an up-to-date and comprehensive HIT event database, researchers may be able to gain a deeper understanding of the nature of HIT related events and their potential consequences. As well, the database may aid in the classification of HIT related events and eventually lead to a universally accepted HIT exclusive reporting system.

Importance of the HIT Event Database

In a high stakes field such as healthcare, it is critical that HIT events are reported and that manufacturers are held accountable for their products. However, manufacturers may not have enough resources allocated towards diagnosing and fixing those issues even when the events are reported. While the staggering pace of technology has driven much innovation in healthcare, the need to take a careful look at the HIT related events has never been greater. The proposed HIT event database offers an opportunity to compare, analyze, and integrate similar HIT events, and ultimately keep HIT on the right track towards becoming a safe and integral part of our healthcare system.

Limitations

The FDA MAUDE database is currently the only publicly accessible resource to contain HIT events. The database is built on voluntary reports and utilizes only a passive surveillance system to verify reports. Consequently, the database may contain incomplete, inaccurate, or biased reports, preventing conclusions regarding the frequency or prevalence of events from being drawn. Nevertheless, the database still contains highly informative data and can be used to better understand the nature and scope of different PSEs.

Another limitation lies in the initial selection of keywords that were used to later find common generic names of HIT devices. While the initial keyword list contained a comprehensive amount of computer and software related terms, it may not have included all possible words that could be used to describe HIT related devices. As a result, few HIT events that were described in an uncommon manner may have been missed during our initial search. In the future, the proposed HIT event database may help overcome this challenge by aiding in the development of a standardized reporting terminology for HIT events. With a standardized reporting terminology, it will likely be easier for

reporters to describe HIT events in a manner that is more useful and informative to researchers.

Future work

Improve Classifiers by Using Semi-supervised Learning

Manually reviewing all cases in the FDA MAUDE database was simply infeasible and that machine learning was likely to be the only viable approach. Traditionally, the two paradigms of machine learning have been supervised (all labeled data) and unsupervised learning (all unlabeled data). However, much attention has been recently placed on semi-supervised learning for its ability to utilize only a small amount of labeled cases combined a with a large amount of unlabeled cases to improve classification accuracy. In the case of HIT, this approach seems well-suited as the cost of labeling narrative text by manual review is quite high, while the cost of obtaining unlabeled reports is minuscule in comparison. One of the simplest methods within semi-supervised learning is self training. In creating our HIT event database, we will utilize a method of iterative self-training to grow the database in a timely manner. It will become a viable method to extract and classify HIT reports from large databases such as the FDA MAUDE.

Utilize the HIT Event Database

More work needs to be done to understand, organize, and uncover the relationships among the HIT reports contained in the database. To fully utilize the database, the themes and topics of the reports must be mapped and understood in context of one and another. Doing so will allow for a better understanding of the relations among HIT events and will greatly aid in the understanding of how HIT changes over time. One potential tool for topic modeling is Latent Dirichlet Allocation (LDA). This commonly used algorithm can be used to analyze unannotated text and discover shared themes among reports. After common themes of HIT related events are found, experts may be able to focus their efforts on addressing these issues and develop actionable solutions to prevent and minimize the risks that patients face during HIT related events. Ultimately, this will advance HIT a safe and integral part of healthcare and improve patient safety.

Conclusion

We proposed a strategy to identify and synchronize HIT events from FDA medical device reports, and to grow an HIT event-exclusive database. The database provides a resource for stakeholders to analyze connections among HIT events, identify common themes of technology-induced errors, synchronize HIT related reports from existing resources, and develop a classification or terminology to standardize HIT reporting toward actionable knowledge and shared learning.

Acknowledgements

This project is supported by UTHealth Innovation for Cancer Prevention Research Training Program Post-Doctoral Fellowship (Cancer Prevention and Research Institute of Texas grant #RP160015), Agency for Healthcare Research & Quality (1R01HS022895), and University of Texas System Grants Program (#156374).

References

[1] C.R. Denham, D.C. Classen, S.J. Swenson, M.J. Henderson, T. Zeltner, and D.W. Bates, Safe use of electronic health records and health information technology systems: trust but verify, *J Patient Saf* **9** (2013), 177-189.

[2] AHRQ Announces Interest in Research on Health IT Safety (Notice Number: NOT-HS-16-009), in, Agency for Healthcare Research and Quality, March 9, 2016.

[3] K.C. Cheung, W. van der Veen, M.L. Bouvy, M. Wensing, P.M. van den Bemt, and P.A. de Smet, Classification of medication incidents associated with information technology, *J Am Med Inform Assoc* **21** (2014), e63-70.

[4] Top 10 health technology hazards for 2015, in, Emergency Care Research Institute (ECRI), November 2014.

[5] Identification and prioritization of health IT patient safety measures., in, Washington, DC, February 2016.

[6] W.E. Fassett, Patient safety and quality improvement act of 2005, *Annals of Pharmacotherapy* **40** (2006), 917-924.

[7] Voluntary patient safety event reporting (incident reporting), in, Agency for Healthcare Research and Quality, August 2014.

[8] Common Formats, in, Agency for Healthcare Research and Quality.

[9] D.F. Sittig and H. Singh, Defining health information technology-related errors: new developments since to err is human, *Arch Intern Med* **171** (2011), 1281-1284.

[10] MAUDE - Manufacturer and User Facility Device Experience, in, U.S. Food and Drug Administration.

[11] R.B. Myers, S.L. Jones, and D.F. Sittig, Review of Reported Clinical Information System Adverse Events in US Food and Drug Administration Databases, *Appl Clin Inform* **2** (2011), 63-74.

[12] F. Magrabi, M.S. Ong, W. Runciman, and E. Coiera, Using FDA reports to inform a classification for health information technology safety problems, *J Am Med Inform Assoc* **19** (2012), 45-53.

[13] F. Magrabi, M.S. Ong, W. Runciman, and E. Coiera, An analysis of computer-related patient safety incidents to inform the development of a classification, *J Am Med Inform Assoc* **17** (2010), 663-670.

[14] Y. Sun, A.K.C. Wong, and M.S. Kamel, Classification of imbalanced data: a review, *International Journal of Pattern Recognition and Artificial Intelligence* **23** (2009), 687.

Address for Correspondence

Yang Gong, MD, PhD

7000 Fannin St Suite 600, Houston, TX 77030, USA

Yang.Gong@uth.tmc.edu

Phone: +1 713 500 3547, Fax: +1 713 500 3907.