Nursing Informatics 2018
A.K. Rotegård et al. (Eds.)

doi:10.3233/978-1-61499-872-3-187

Exploring Health Information Technology Events from FDA MAUDE Database

Bin Yao, Hong Kang, Ju Wang, Sicheng Zhou, Yang Gong

School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Texas, USA

Abstract

To facilitate preventing patient safety events, researchers have been using the Federal Drug Administration (FDA) Manufacturer and User Device Experience (MAUDE) database as a publicly accessible data source for retrieving reports of medical devices including Health Information Technology (HIT) devices. A combined search strategy built with keywords in several fields such as Generic Name, Brand Name and Manufacturer is commonly used to extract reports of specific classifications of devices in the MAUDE. However, another structured field, Classification Product Code, is rarely visited. To improve the secondary usage of the MAUDE database in retrieval of HIT events, we extracted HIT reports from a nine-year MAUDE dataset by combining keywords filter and expert review, and explored Classification Product Codes. The distributions and relationships between Product Codes and keywords in Generic Name/Manufacturer in HIT reports were visualized to provide a view of the landscape. This study presents a new perspective for improving the search strategy of HIT events in MAUDE, which would facilitate the understanding of HIT events for improving patient safety.

Keywords:
Health Information Technology; Database; Patient Safety

Introduction

Health Information Technology (HIT) related adverse events

An HIT device is broadly defined by the Agency for Healthcare Research and Quality (AHRQ) as "hardware or software that is used to electronically create, maintain, analyze, store, or receive information to aid in the diagnosis, cure, mitigation, treatment, or prevention of disease and that is not an integral part of an implantable device, or an item of medical equipment" [1]. HIT is commonly used for administration/management, medication dispensing, medical imaging, clinical decision support, vital sign monitoring and adverse event reporting [2].

The increasing integration of HIT into various aspects of healthcare holds promise to improve efficiency and effectiveness of care delivery [3; 4]. However, problems come along with benefits and HIT itself can become a hazard if HIT is not properly designed, implemented, or used, which threatens the quality of care and may lead to patient safety events [4-6]. According to the Identification and Prioritization of Health IT Patient Safety Measures Final Report by National Quality Forum [2], HIT events not only impact the safety and quality but also lead to other adverse events such as medication error and medical device event. In the Eight-Dimensional Social-Technical Model proposed by Sittig and Singh [7], HIT events can be caused or contributed by several components in patient care from human factors, hardware/software and human-device interaction to management and policies [5]. The multi-dimensional nature of HIT events makes them complicated to analyze and prevent.

Urgent demand in acquiring a data source of HIT event reports

To better understand and prevent HIT events, learning from HIT reports is essential [8]. However, there is a lack of comprehensive and accessible HIT event data resources to meet the urgent demand in data aggregation, root cause analysis and signal detection for adverse events of HIT devices. HIT has been ubiquitous in nowadays practices. Thus, the patient safety events (PSE) that involved with HIT events are commonly reported as events such as adverse drug events, pressure injury, patient fall and events of medical devices. The updated version of AHRQ Common Formats does not contain the reporting form for HIT rather distributes HIT as a contributing factor across other patient safety events [1].

The Federal Drug Administration (FDA) Manufacturer and User Device Experience (MAUDE) database [9], a public resource for reports of adverse events of medical devices, enables the aggregation of reports submitted by both consumers and manufacturers. The MAUDE database contains mandatory and voluntary reports of medical device events and has been continuously updated since June 1993. Although the MAUDE database was not designed for reporting HIT events, it is still one of the most important public resources with the potential to extract HIT events. Efforts have been made to identify and classify HIT event reports from the FDA MAUDE database via approaches based on human review and machine learning methods. Nevertheless, well defined and evaluated searching strategies for HIT event reports in FDA MAUDE is still in demand.

Search HIT event reports in the MAUDE database

It is essential to include a brief review on the efforts in retrieving HIT events from the FDA MAUDE database [10-12]. In 2012, Magrabi et al. [11] developed a list of keywords in the fields of *Brand Name*, *Generic Name* and *Manufacturer* to search the FDA MAUDE database for HIT event reports and estimated 0.1% reports submitted from 2008 to 2010 in FDA MAUDE are HIT. With a broadened keyword list and revised review strategy applied on a 9-year dataset from 2008 to 2016, we updated this ratio to 0.69%. The most commonly used search strategies composed of keywords in *Generic Names*, *Brand Names*, *Manufacturers and Device Names*, which have not been standardized and may weaken generalization and consistency of the results. *Brand Names* and *Manufacturer Names* updates frequently along with tides of the market and commercial strategies, and the *Generic Names* are prone to errors such as typographical errors and mix-up between *Brand Names* and *Generic Names*. A less explored field associated with FDA MAUDE database is the *Classification Product Codes* (*Product Codes* for short) represented by a combination of three upper-case letters, which is accessible in the FDA Product Classification Database. The *Product Codes* are

developed based on the medical device product classification designated under 21 CFR Parts 862-892. A *Product Code* along with a *Device Name* is assigned to each of the reports in the MAUDE database inferring the classification of the device involved in the event, which can be a potential clue for the construction of search strategies.

The purpose of this study is to unravel distributions of HIT reports in FDA MAUDE and to explore improvement of search strategies. In this study, we reviewed the search strategies for identifying medical devices in FDA MAUDE database and specifically updated the estimated proportion of HIT reports in the database with a refined search strategy. Moreover, for further investigation and improvement of data retrieval of HIT reports from the MAUDE database, we analyzed distributions and connections of the *Classification Product Code*, *Generic Name*, *Brand Name and Manufacturer* in HIT reports and proposed a search strategy for identifying HIT reports from the MAUDE database and for enhancing our understanding of the database.

Methods

Literature review

We searched Medline, Embase and Cochrane databases by using keywords 'Federal Drug Administration' 'FDA' 'Manufacturer and User Device Experience' and 'MAUDE' for peer-reviewed articles in English language published in latest 10 years from January 1, 2007 to October 1, 2017. After removal of duplicates and screening by the titles, 61 articles were selected from 732 articles. 35 articles that performed searches with specific strategies identifying one or several categories of medical devices in the MAUDE database were selected for review. Search strategies and categories of target devices of the 35 articles were then summarized.

Data retrieval, curation and sampling

All MAUDE data submitted from 2008 to 2016 were downloaded and searched by a keyword-based filter developed in our previous work which is in preparation for a journal publication. Multiple narrative entries from the same event were merged. Duplicates were eliminated by combining records by either Health Care Facility Association (HCFA) number or distributor report number and/or manufacturer report number (HCFA/distributor/manufacturer number). Within each HCFA/distributor/manufacturer number, ten percent reports were randomly sampled for expert review. If the amount of reports with a certain HCFA number was fewer than ten, one report was randomly selected.

Expert review for HIT related reports

Review criteria for HIT reports were based on the HIT device definition by AHRQ. Reports were reviewed by three domain experts who were familiar with HIT event reporting and patient safety data. To estimate the consistency among reviewers, reports of a random year (2011) were selected to be reviewed by all three reviewers individually and the pairwise Cohen's kappa scores were calculated.

Data analysis and visualization

Frequencies of each *Product Code* and keywords in the HIT reports were calculated and ranked. Relationships between HIT *Product Codes* and keywords in the fields of *Generic Name/Manufacturer* were demonstrated in a Sankey diagram (Figure 1) plotted with rCharts version 0.4.5.

Results

Search strategies in the MAUDE database

Over the past decade, researchers have utilized the FDA MAUDE database for retrieving reports of medical devices for retrospective analysis and signal detection of medical device related adverse events. The identified 35 articles present two common scenarios when reports of specific devices were of interests. The first scenario is that the devices of a specific brand or manufacturer are requested by researchers, and the other one is that the devices under a specific classification but may cover multiple brand names and manufacturers are requested.

As is shown in Table 1, to retrieve reports for devices of specific *Brand Names* or *Manufacturer Names*, most studies performed a search strategy to search in one or a combination of manufacturer names and brand names. For this scenario, researchers can narrow the keywords in a clear and limited list to obtain highly relevant results. However, when it comes to another frequent scenario, i.e. devices under a specific classification, there are more keywords to be included in the search strategy and the accuracy of the results are unsatisfactory. Only 678 HIT related reports were identified from 1,100 reports retrieved from the dataset from 2008 to 2010 according to Magrabi et al.[11]

Table 1. Search strategies for reports of specific devices in the FDA MAUDE database

Goal of Search	Search Strategy	Number of Articles
Device of specific brand or manufacturer	Brand Name /Manufacturer Name	5
	Device Name (Product Name)	2
	Product Class	1
	Keywords in unmentioned fields	3
Device of specific classification	Brand Name /Manufacturer Name	7 (2*)
	Device Name (Product Name)	8 (2*)
	Product Code (Device Code)	6
	Narrative	2
	Product Class	1
	Keywords in unmentioned fields	11

* HIT related

HIT reports in the FDA MAUDE database

A total number of 2,234 reports from the sampled FDA MAUDE reports from 2008 to 2016 were identified as HIT reports by domain experts. The pairwise Cohen's kappa scores of the three experts on the 2011 data are 0.82, 0.85 and 0.87, indicating a high consistency.

Product Code, Generic Name and Review Panel of HIT reports in the FDA MAUDE database

The top *Product Codes* identified from HIT reports are ranked by number of corresponding HIT reports (Table 2). *Product Codes* that make up of beyond 1% HIT reports include MHX (Monitor, Physiological, Patient (With Arrhythmia Detection Or Alarms)), DQK (Computer, Diagnostic, Programmable), DSI (Detector And Alarm, Arrhythmia), LNX (Medical Computers And Software), DXG (Computer, Diagnostic, Pre-Programmed, Single-Function), NSX (Software, Transmission And Storage, Patient Data), DPS (Electrocardiograph). Specifically, 29.63% are marked by the Product Code 'MHX' whose corresponding *Device Name* is 'Monitor, Physiological, Patient (With Arrhythmia Detection Or Alarms)'. A large number of HIT reports were assigned to the *Product Code* 'LLZ', though the majority of 'LLZ' reports were identified as non-HIT. Table 3 shows the top keywords in *Generic Name* and *Manufacturer* of HIT reports. The top 6 keywords are 'Patient monitor' 'Computer' 'Program' 'Software' 'Data' and 'Telemetry', which cover 50.33% of the HIT reports. Figure 1 shows the connections between the top HIT *Product Codes* and the keywords. The *Product Codes* are labels by their Review Panels in the left column. The cumulative percentages indicate that HIT reports are more concentrating in a few *Product Codes* than in *Generic Names* (Table 2 and 3). The Sankey Diagram demonstrates the relationships between *Review-panel*, *Product Code* and keywords (Figure 1).

Table 2. Top HIT Product Codes (Percentage above 1%)

Product Code	Device Name	Filtered Reports	Sampled Reports	HIT Events	True Positive Rate (TPR)	Percentage in HIT Events (%)
MHX	Monitor, Physiological, Patient (With Arrhythmia Detection Or Alarms)	4,530	768	662	0.86	29.63
LLZ	System, Image Processing, Radiological	15,096	1,624	241	0.15	10.79
DQK	Computer, Diagnostic, Programmable	970	176	141	0.80	6.31
NSX	Software, Transmission And Storage, Patient Data	138	111	107	0.96	4.79
DSI	Detector And Alarm, Arrhythmia	715	111	98	0.88	4.39
LNX	Medical Computers And Software	207	89	82	0.92	3.67
JQP	Calculator/Data Processing Module, For Clinical Use	494	93	61	0.66	2.73
DRT	Monitor, Cardiac (Incl. Cardiotachometer & Rate Alarm)	192	62	52	0.84	2.33
NBW	System, Test, Blood Glucose, Over The Counter	775	103	40	0.39	1.79
OUG	Medical Device Data System	87	42	32	0.76	1.43
HGM	System, Monitoring, Perinatal	87	39	27	0.69	1.21
MMH	Blood Establishment Computer Software And Accessories	72	37	27	0.73	1.21
BSZ	Gas-Machine, Anesthesia	281	79	25	0.32	1.12
DXG	Computer, Diagnostic, Pre-Programmed, Single-Function	180	26	24	0.92	1.07

Table 3. Top ranking keywords (Percentage above 1%)

Generic Name	Filtered Reports	Sampled Reports	HIT Events	True Positive Rate(TPR)	Percentage in HIT Events (%)
Patient monitor	3,224	416	376	0.90	13.87
Computer	3,743	900	255	0.28	9.41
Program	5,545	620	231	0.37	8.52
Software	3,228	467	212	0.45	7.82
Data	1,215	231	154	0.67	5.68
Telemetry	517	175	136	0.78	5.02
ICT	14,660	1,513	132	0.09	4.87
Picture archiving	14,613	1,503	128	0.09	4.72
Central monitor	1,084	131	111	0.85	4.10
Monitoring system	644	164	87	0.53	3.21
CPOE	124	82	80	0.98	2.95
Management system	1,295	229	80	0.35	2.95
PACS	164	58	58	1.00	2.14
Physiological monitor	126	65	57	0.88	2.10
Automated	11,635	1,432	51	0.04	1.88
EHR	55	50	49	0.98	1.81
Network	248	51	48	0.94	1.77
Workstation	180	63	42	0.67	1.55
Telemetry monitor	196	36	30	0.83	1.11
Imaging system	259	91	29	0.32	1.07
Digital	156	89	27	0.30	1.00

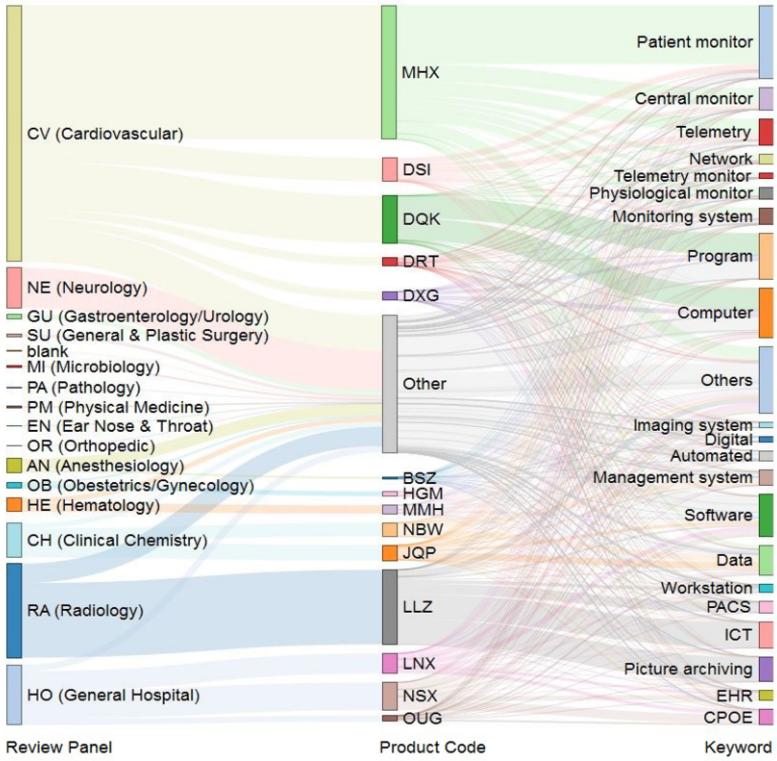

Figure 1. Sankey diagram of review-panel, Product Code and keywords in Generic Name and Manufacturer
The height of bar shows the proportion of each item, and width of the path between two bars shows the proportion of reports shared by two items. Colors of bars are applied for distinguishing different items and do not represent the relationships. The figure can be manipulated interactively in a browser by dragging the bars for demonstrating details.

Discussion

Pros and cons of search strategies

Efforts have been made to retrieve HIT reports from the FDA MAUDE database, though the database is not exclusively designed for reporting HIT events. The commonly used search strategy is based on a keyword filter in the fields of *Generic Name*, *Brand Name*, and *Manufacturer*. The keyword-based filter has its advantages. Frequently used HIT keywords such as 'patient monitor' 'computer' 'program' 'software' and 'data' are straightforward and easy to understand by users. Specific categories of devices can be easily extracted using relative keywords. However, due to the tremendous numbers of keywords and inevitable typographical errors, users have to update the filters frequently and rectify the typographical errors in these fields. *Product Code*, on the contrary, defines more general categories of devices and is highly structured. HIT reports are more concentrating in a limited number of *Product Codes* rather than in keywords. A high proportion of HIT reports are under 'MHX' 'LLZ' 'DQK' 'NSX' and 'DSI' (Table 2). However, a large number of reports caught by the top HIT *Product Codes* and keywords are excluded by the expert review criteria and identified as non-HIT reports (Table 2 and 3). To make better use of the FDA MAUDE database, standardization of the structured fields including Generic Name, Manufacturer, Brand Name and Product Code is critical for improving efficiency and effectiveness of the search strategies.

The inter-department nature of HIT events

The Sankey Diagram (Figure 1) demonstrates the relationships between *Review-panel*, *Product Code* and keywords in *Generic Name/Manufacturer*. HIT events distribute over Review-panels, which indicates multiple departments in the hospital are involved with HIT events, including 'Cardiovascular' 'Radiology' 'General Hospital' 'Neurology' 'Clinical Chemistry, etc., which demonstrates inter-department nature of HIT events. This is one of the contributing factors of HIT events because interoperability issues may occur when data are exchanged between different information systems [2].

Towards an open data source for HIT events

As HIT events are commonly associated with other types of events such as medication errors and medical device events, HIT events are usually reported along with those events, and sometimes the reporters may even do not realize HIT events are involved due to limited knowledge or lack of training. Thus, a

large proportion of invaluable HIT events are buried in patient event databases designed for other events, making it difficult of retrieve, analysis and preventing HIT events. By exploring the relationship between the *Product Code* and HIT devices, this study would shed light on the structured fields in the FDA MAUDE database that are rarely utilized towards HIT event identification.

Limitations of study

Duplicated reports still existed after applying our filter, which led to difficulties in following review and analysis. The duplicates include follow-ups of the same events, recurrences of the same type of events, or duplicated reports of the same events by multiple reporters. Besides, it is also recommended that an event involving multiple personnel be reported separately. Despite our efforts to remove the duplicates by merging reports per the HCFA/distributor/manufacturer numbers, unidentified duplicates may still compact the estimation for HIT events in the MAUDE database. Due to the limited sample size, the results may not fully describe the whole database or predict future reports. Moreover, the reports for expert review were sampled by HCFA/distributor/manufacturer numbers but not by *Product Codes*, so the proportion of HIT Events may not properly reflect the distribution in each *Product Codes*. Finally, since the definition of HIT events is evolving because of the rapid development of technologies, ambiguities and discrepancies are inevitable due to the limited domain knowledge of the reviewers.

Future work

Leveraging search strategies for HIT reports in the MAUDE database

Search strategies developed by single field or combinations of *Brand Name*, *Manufacturer*, *Generic Name* and *Product Code* will be performed to search in the MAUDE database for HIT related reports in 2017. Randomly sampled results will be reviewed by experts to filter HIT related adverse events and precision calculated for comparison.

Developing a database of HIT event reports

Although the HIT event reports are not supposed to correctly reflect the frequencies of events due to the issues such as under-reporting and duplication, the reports are invaluable in root cause analysis and case-based study to help prevent recurrence of patient safety events. Previous works unveiled the great value of the MAUDE database for aggregating HIT events. Due to an absence of a database exclusively designed for HIT events, our future work is to extract existing HIT reports in the MAUDE database and develop an exclusively HIT event database to facilitate our understanding of HIT events.

Conclusions

We analyzed the search strategies for retrieving reports of specific medical devices from the FDA MAUDE database and compared the use of *Product Codes* and keywords in the search strategies. To explore HIT reports in the MAUDE, we identified HIT reports in the nine-year MAUDE data and unraveled distributions and connections of HIT *Product Codes* and *Generic Names*. We also concluded the absence of using product code and its role in HIT reports retrieval in the MAUDE database, and the implementation of a combined search strategy.

Acknowledgements

The study is supported by a research grant (1R01HS022895) from the Agency of Healthcare Research and Quality, UTHealth Innovation for Cancer Prevention Research Training Program Post-Doctoral Fellowship (Cancer Prevention and Research Institute of Texas grant #RP160015).

References

[1] Supporting Documents - Common Formats - Hospital Version 1.2, in, Patient Safety Organization Privacy Protection Center.

[2] Identification and Prioritization of Health IT Patient Safety Measures, in, National Quality Forum, 2016.

[3] L. Kohn, J. Corrigan, and M. Donaldson, To Err Is Human: Building a Safer Health System, in, Washington, DC, 2000.

[4] B. Chaudhry, J. Wang, S. Wu, M. Maglione, W. Mojica, E. Roth, S.C. Morton, and P.G. Shekelle, Systematic review: impact of health information technology on quality, efficiency, and costs of medical care, *Ann Intern Med* **144** (2006), 742-752.

[5] G.M. Castro, L. Buczkowski, and J.M. Hafner, The Contribution of Sociotechnical Factors to Health Information Technology-Related Sentinel Events, *Jt Comm J Qual Patient Saf* **42** (2016), 70-76.

[6] Health IT and patient safety: building safer systems for safer care, in, Institute of Medicine, Washington, DC, 2012.

[7] 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.

[8] N. Rafter, A. Hickey, S. Condell, R. Conroy, P. O'Connor, D. Vaughan, and D. Williams, Adverse events in healthcare: learning from mistakes, *The Quarterly journal of medicine* **108** (2015), 273-277.

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

[10] H. Kang, F. Wang, S. Zhou, Q. Miao, and Y. Gong, Identifying and synchronizing health information technology (HIT) events from FDA medical device reports., *The 16th World Congress on Medical and Health Informatics (MedInfo 2017)* (2017).

[11] 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.

[12] H. Kang, Z. Yu, and Y. Gong, Initializing and growing a database of health information technology (HIT) events by using TF-IDF and biterm topic modeling. , *American Medical Informatics Association (AMIA) 2017 Annual Symposium.* (2017).

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