Public Health and Informatics
J. Mantas et al. (Eds.)

doi:10.3233/SHTI210115

Can FHIR Support Standardization in Post-Market Safety Surveillance?

Xingtong WANG [a], Harold LEHMANN [a] and Taxiarchis BOTSIS [a,1]
[a] *Johns Hopkins University School of Medicine, Baltimore, MD, United States*

Abstract. The Fast Healthcare Interoperability Resources (FHIR) contain multiple data-exchange standards that aim at optimizing healthcare information exchange. One of them, the FHIR AdverseEvent, may support post-market safety surveillance. We examined its readiness using the Food and Drug Administration's (FDA) Adverse Event Reporting System (FAERS). Our methodology focused on mapping the public FAERS data fields to the FHIR AdverseEvent Resource elements and developing a software tool to automate this process. We mapped directly nine and indirectly two of the twenty-six FAERS elements, while six elements were assigned default values. This exploration further revealed opportunities for adding extra elements to the FHIR standard, based on critical FAERS pieces of information reviewed at the FDA. The existing version of the FHIR AdverseEvent Resource may standardize some of the FAERS information but has to be modified and extended to maximize its value in post-market safety surveillance.

Keywords. Post-market Safety Surveillance, Adverse Drug Event, FHIR

1. Introduction

The US Food and Drug Administration (FDA) and other regulatory agencies worldwide monitor the safety of medical products in the clinical trial phase and after their release to the market. The post-market evaluation mainly relies on the reports submitted by physicians, manufacturers, and consumers to spontaneous reporting systems. The FDA's Adverse Event Reporting System (FAERS) is one of these repositories and receives millions of post-market reports for drugs and biologics every year [1]. FAERS reports contain many structured and unstructured data fields reviewed by the FDA's Safety Evaluators (SEs) to identify and evaluate potential safety signals. These are thorough and time-consuming processes that often require examining external knowledge and requesting additional information from the reporters. The increasing number of FAERS reports necessitates automated approaches to support their assessment. Computer-based systems are essential to one of the significant challenges in post-market surveillance, the standardization of the information collected by the reporters and submitted to the FDA.

One of the first attempts to address this challenge has been the Observational Medical Outcomes Common Data Model (OMOP CDM) [2]. The OMOP model, developed in the Observational Health Data Sciences and Informatics (OHDSI) program, transforms data within clinical databases into a CDM and prepares it for analysis. Such analyses may be conducted using an OHDSI library of standard analytic routines or other

[1] Corresponding Author: Taxiarchis Botsis, Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, MD, United States. E-mail: tbotsis1@jhmi.edu.

statistical packages. A recent study indicated that the OMOP model is likely not adequate for post-market purposes yet and discussed using the more complete Fast Healthcare Interoperability Resources (FHIR) [3]. FHIR contains multiple data elements and formats that may support healthcare information exchange. The Health Level 7 organization has developed FHIR to increase the utilization of the Electronic Health Record (EHR) data for primary (clinical) purposes through structuring and standardization. A uniform data standard like FHIR may support the exchange and transfer between different facilities and systems, such as a clinical institution's EHR and FAERS. In particular, the FHIR AdverseEvent (hereafter, AER) was formed to communicate events during medical care or medical research that may impact human subjects [4]. Our study examines the AER's potential use for secondary use in post-market safety surveillance. We also analyze its completeness based on actual FAERS data and discuss opportunities and next steps.

2. Methods

The FDA maintains an internal FAERS database and releases a portion of every Quarter's post-market information [5]. The public release is fully de-identified, contains no free-text narratives, and can be accessed via the openFDA and the FAERS online dashboard. XML and ASCII versions of the quarterly files are also available for direct download. The ASCII version contains more data fields (in seven files) and has a more transparent structure and supporting documentation than the XML version and the openFDA. Therefore, we elected to use the demographic, drug, reaction, outcome, report sources, therapy, and indication data from the ASCII files.

We thoroughly reviewed the public FAERS files and identified the data fields that may directly and entirely (by retrieving values from homonymous or conceptually identical FAERS fields) or indirectly and conditionally (by deriving the values according to rules and specific conditions) feed the AER. Default values were assigned to the remaining elements. As shown in Figure 1, AER contains twenty parent and six child elements; some of these elements reference other FHIR Resources. We looked at these Resources and investigated whether any FAERS fields could fill any of their elements and, consequently, extend the FHIR utilization for post-market standardization. We also explored AER's extension by examining other fields in FAERS. After finalizing the mapping process, we analyzed the completeness of the selected FAERS data fields since 2012 (Q4).

In the last step of our work, we developed a software tool in Python 3.9 that processes the ASCII files and retrieves the values required to fill the elements in the extended AER described above. We measured the tool's efficiency by randomly selecting 50 reports submitted to FAERS in 2019, Q1-2.

3. Results

We thoroughly reviewed the public FAERS ASCII files to map the FAERS data fields to the AER elements. Figure 1 shows the results of the mapping process. Of the twenty-six parent and child elements in the AER (non-shaded elements in Figure 1), we were able to match nine to the FAERS repository entirely. The "EVENT_DT" FAERS field was found to support the "date" and the "detected" elements.

FHIR AER	FHIR Brief Description	Ref	FAERS File	FAERS Field
identifier	Business identifier for the event	NO	Demographic	PRIMARYID
actuality	Actual OR potential	NO	***Default***	***Actual***
category	Potential reasons for event	NO	***Default***	***Product-problem***
event	Event coded to SNOMED CT	NO	Reaction	PT
subject	Subject impacted by event	YES	***Default***	***Patient***
age	Age in Years		Demographic	AGE
encounter	Encounter created as part of	YES	***N/A***	***N/A***
date	When the event occurred	NO	Demographic	EVENT_DT
detected	When the event was detected	NO	Demographic	EVENT_DT
recordedDate	When the event was recorded	NO	Demographic	INT_FDA_DT
resultingCondition	Effect on the subject	YES	***N/A***	***N/A***
location	Location where event occurred	YES	Demographic	OCCR_COUNTRY
seriousness	Seriousness of the event	NO	Outcome	OUTC_COD
severity	Mild OR moderate OR severe	NO	***Outcome***	***OUTC_COD (Der)***
outcome	Recovered or not	NO	***Outcome***	***OUTC_COD (Der)***
recorder	Who recorded the event	YES	Demographic	OCCP_COD
contributor	Who was involved in the event	NO	***Default***	***Practitioner***
suspectEntity	Suspected agent causing the event	NO	Drug	DRUGNAME (Con)
instance	The entity that caused the event	YES	***Default***	***Medication***
causality	Possible cause of the event	NO	***N/A***	***N/A***
assessment	Has the entity caused the event?	NO	***N/A***	***N/A***
productRelatedness	Description ofcausality	NO	***N/A***	***N/A***
author	Who made the assessment	YES	***N/A***	***N/A***
method	ProbabilityScale, Bayesian, Checklist	NO	***N/A***	***N/A***
otherMedication	Drug names		Drug	DRUGNAME (Con)
subjectMedicalHistory	Medical History	YES	***N/A***	***N/A***
referenceDocument	Document describing the event	YES	***Default***	***FAERS report***
study	Reference to a research study	YES	***N/A***	***N/A***

Figure 1: FHIR AER to FAERS Mapping. The elements shaded in blue could enrich AER. Ref: Referenced; N/A: Not applicable; Der: Derived Based on Certain Rules; Con: Conditional, relies on other field values.

Notably, the "suspectEntity" element could entirely match the "DRUGNAME" field, if the value in the "ROLE_COD" field (also in the "Drug" table) were set to "PS" (Primary Suspect Drug). It was also found that the values for the "severity" and the "outcome" elements could be (partially) derived from the "OUTC_COD" FAERS field. The latter contains seven codes to describe serious patient outcomes (death, life-threatening, hospitalization, disability, congenital anomaly, required intervention to prevent permanent impairment/damage, and other serious outcomes). If the "OUTC_COD" element were completed with any of these values, the event would be "severe" (for "severity") and the subject would "not recover" (for "outcome") in the short-term. Six of the AER elements were assigned a default value based on facts related to the explicit use of the FAERS repository use and not, for example, of a clinical database. FAERS reports contain the descriptions for "actual" reports related to a "product-problem"; the subject experiencing the event is a "patient," and a "practitioner" is generally involved in the reporting process; the instance is always a "medication", not an "immunization" or any of the potential values shown in the AER; and, the reference is the FAERS report.

It was also found that nine of the AER elements could not be completed (marked as N/A in Figure 1). This observation should be attributed to two main reasons. First, the corresponding information (for the "resultingCondition" and the "subjectMedicalHistory") is only included in additional structured data fields and the

free-text narrative of the non-public version of the FAERS database. Second, to complete the "causality" branch, we would need to determine the causality association between the event and the product in a thorough medical review of all the available information stored in the FDA's internal FAERS database. Even if we had this level of access, it would be difficult to assign any value to the "encounter" element that primarily refers to clinical services provided to patients.

Our close examination of the referenced FHIR Resources and the FAERS repository revealed some opportunities to add more elements to the AER by either creating new elements or looking into other Resources. Figure 1 shows two examples of new elements (shaded in blue), "age" and "otherMedication". The former could be calculated from the "birthdate" element in the Patient Resource, and the latter, representing concomitant medications, would refer to the Medication Resource. Other elements not included in the AER but exist in other resources could further enrich the standardized description of an adverse event. For example, the date of administration for the "suspectEntity" element is part of the Medication Resource and could be mapped to the corresponding FAERS field.

We subsequently evaluated the FAERS repository's ability to support the completion of the AER (existing and proposed) elements mapped to the FAERS fields (mappings shown in Figure 1) using the FAERS data since 2012 (Q4). Elements with default and derived values or unmapped elements were excluded from these calculations. Some of the FAERS data fields' missingness was worrisome (49% and 42% of the values were missing in EVENT_DT and AGE, respectively) but might be corrected by retrieving data from FAERS narratives. Missingness was less of an issue for OCCR_COUNTRY (3%), OCCP_COD (3%), and START_DT (5%) and was not observed in the remaining fields.

Our tool offers single and batch file processing options to generate output stored to a JSON a CSV file, respectively. The run processing time in MS Windows to process the 50 FAERS reports and generate the JSON (one for each report) and the CSV file (one for all reports) was 0.1346 and 0.0149 seconds, respectively. The code is publicly available at https://github.com/xingtongJHU/FAERS-to-FHIR-Automated-Mapping.

4. Discussion

We successfully mapped, either directly or indirectly, eleven AER elements to FAERS data fields and assigned default values to six elements. The public FAERS could not support the remaining nine elements. Filling these missing elements would require access to the non-public FAERS database and the outcome of the safety review conducted by FDA's SEs that determines the causality association between an adverse event and a drug product. Interestingly, we found many other FAERS data fields that could support an AER extension and provided a few representative examples accordingly.

Our study has three limitations. First, the public FAERS does not offer access to the free-text narratives that describe the adverse event and contain critical details not always included in the structured data fields, such as "subjectMedicalHistory" and "date". The natural language processing (NLP) of these narratives would allow us to complete these values. Our team has multi-year experience with the NLP of FAERS and other post-market reports and could accomplish this task [6]. Second, the derived values for "severity" and "outcome" cannot always be completed based on what is included in the mapped FAERS fields. If an event is not "severe", we will not know whether the "mild" or "moderate" value should be assigned. It might also be argued that a patient could have a serious "outcome" but "recover" quickly. This criticism is valid, and the related matters

have to be revisited in a future study. Third, the default values assigned to six of the AER elements is a reasonable compromise for the FAERS data but might not work well in a different environment and setup, such as a clinical institution. The potential application of the AER to a different setting would require a new round of dedicated analysis.

Our proposal for the potential extension of the FHIR AER is in line with the post-market tasks and the critical elements reviewed by the SEs at the FDA [7]. It is also very timely because this Resource is still in the Trial Use phase, allowing time for changes. Some of the additional elements may be found in the other referenced FHIR Resources, such as the "Patient" Resource. Other elements may not explicitly exist in FHIR but are significant and should be included in a future AER release. Any alterations should be carefully examined and discussed with the FDA stakeholders and other experts interest in standardizing post-market information. We also foresee the need to revisit other aspects in the existing AER version. For example, the "event" values should be coded to the Medical Dictionary for Regulatory Activities terms, widely used in post-market surveillance, rather than SNOMED CT terms. Some NLP tools, like MetaMap, can support such conversions.

Our study examined the post-market information standardization for drug products but could inspire similar work for medical devices and vaccines. The structuring, normalization, and standardization of post-market information for all medical products may best support FDA's analyses, facilitate the development of automated approaches, and likely improve the quality of submitted data. It may further empower the communication with the clinical institutions, upon any request for additional information, and set the path for the automated transmission of adverse event data to the FDA.

Acknowledgments

This work was supported in part by a Center of Excellence in Regulatory Science and Innovation (CERSI) grant to Johns Hopkins University from the US Food & Drug Administration (FDA) U01FD005942. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the HHS or FDA.

References

[1] FDA. FDA Adverse Events Reporting System (FAERS) Public Dashboard. Available at: https://fis.fda.gov/sense/app/d10be6bb-494e-4cd2-82e4-0135608ddc13/sheet/7a47a261 -d58b-4203-a8aa-6d3021737452/state/analysis. Accesed: Jan 13, 2021

[2] Knowledge Base workgroup of the Observational Health Data, S. and c. Informatics, Large-scale adverse effects related to treatment evidence standardization (LAERTES): an open scalable system for linking pharmacovigilance evidence sources with clinical data. J Biomed Semantics, 2017;8(1):11.

[3] Choi M, Starr R, Braunstein M, Duke J. OHDSI on FHIR platform development with OMOP CDM mapping to FHIR Resources. in OHDSI Symposium, Observational Health Data Sciences and Informatics, Washington, DC. 2016.

[4] HL7, FHIR Resource AdverseEvent - Content. Official HL7 online resource, Available at: https://www.hl7.org/fhir/adverseevent.html. Accessed Jan 13, 2020.

[5] FDA. FDA Adverse Event Reporting System (FAERS) Quarterly Data Extract Files, Available at: https://fis.fda.gov/extensions/FPD-QDE-FAERS/FPD-QDE-FAERS.html, Accesed Jan 13, 2021.

[6] Botsis T, Foster M, Arya N, Kreimeyer K, Pandey A, Arya D. Application of Natural Language Processing and Network Analysis Techniques to Post-market Reports for the Evaluation of Dose-related Anti-Thymocyte Globulin Safety Patterns. Appl Clin Inform. 2017 Apr 26;8(2):396-411.