Evaluating Site-Level Implementations of the HL7 FHIR Standard to Support eSource Data Exchange in Clinical Research

Maryam Y. GARZA[a,1], Michael W. RUTHERFORD[a], Bhargav ADAGARLA[b], Eric EISENSTEIN[b], Karan R. KUMAR[b], Kanecia O. ZIMMERMAN[b], Umit TOPALOGLU[c] and Meredith ZOZUS[d]

[a] *University of Arkansas for Medical Sciences, Little Rock, Arkansas*
[b] *Duke Clinical Research Institute, Duke University, Durham, North Carolina*
[c] *Wake Forest University School of Medicine, Winston-Salem, North Carolina*
[d] *University of Texas Health Science Center at San Antonio, San Antonio, Texas*

Abstract. Direct extraction and use of electronic health record (EHR) data is a long-term and multifaceted endeavor that includes design, development, implementation and evaluation of methods and tools for semi-automating tasks in the research data collection process, including, but not limited to, medical record abstraction (MRA). A systematic mapping of study data elements was used to measure the coverage of the Health Level Seven (HL7®) Fast Healthcare Interoperability Resources (FHIR®) standard for a federally sponsored, pragmatic cardiovascular randomized controlled trial (RCT) targeting adults. We evaluated site-level implementations of the HL7® FHIR® standard to investigate study- and site-level differences that could affect coverage and offer insight into the feasibility of a FHIR-based eSource solution for multicenter clinical research.

Keywords. HL7 FHIR, eSource, electronic health records, clinical research, interoperability, health information exchange

1. Introduction

Real-world data (RWD) holds promise for (1) extending clinical research results to real-world evidence (RWE) and evidence-based clinical practice and (2) overcoming the threats to generalizability and effectiveness in real world-settings that can occur with tightly controlled clinical trials. At the same time, reuse of RWD, including, but not limited to electronic health record (EHR) data, may offer relief from the rising complexity and cost of multicenter clinical research studies [1–3]. The past ten years have witnessed an almost complete U.S. adoption of EHRs [4], interoperability standards for exchange and reuse of healthcare data [5], mature privacy and confidentiality rules to protect healthcare data [6, 7], and information technologies with capabilities to exchange and reuse healthcare data [8]. These advances provide key building blocks needed to re-engineer data collection and use for clinical trial design, conduct, and reporting. The now-possible end-to-end electronic data handling offers the potential of

higher data quality at lower cost and increased rigor from the source (i.e., EHR) to the regulator (i.e., electronic data capture, or EDC, system).

However, the quality of RWD data sources, including EHRs, and subsequent abstraction methods (i.e., MRA) has long been questioned [9–12]. eSource platforms that facilitate automatic extraction of patient data from EHRs-to-EDCs have the potential to reduce site burden and improve data quality in clinical research, but the scalability and generalizability of existing solutions is limited, and the impact of such solutions for multicenter clinical research remains uncertain. Evaluations of EHR data directly extracted for use in ongoing clinical studies (EHR-to-EDC eSource) have been sporadic; conducted in limited application contexts such as single studies, single-sites, single-EHRs, and single therapeutic areas; and evaluated using disparate outcome measures [8, 13]. Few studies simultaneously assessed data quality and data collection cost so that the burden on clinical sites could be evaluated [13]. Further, few studies applied rigorous data quality assessment methods, provided a gold standard comparison needed to quantify accuracy, or reported the detail needed to understand the location, extent and root cause of data errors [13] – all key factors in determining the suitability of data for generating evidence for regulatory decision-making and evidence-based practice. The results presented here attempt to address some of these information gaps by providing insight into the impact of standards in supporting data collection in clinical research.

Data standardization is necessary for computers to assist or automate the exchange and reuse of data. Data standards specify data definitions and formats to enable data exchange between computers and reuse by computers and humans. Standardization helps assure that data with the same name have the same meaning. The new and widely adopted Health Level Seven (HL7®) Fast Healthcare Interoperability Resources (FHIR®) standard goes further than predecessors toward standardizing the meaning and format of structured data collected in EHRs [5]. Developing standards, however is only the first step. The standards must facilitate desired workflows and data flows and ultimately be implemented in software toward improving human performance through mechanisms such as process automation and decision support. The results reported here brings us one step closer to understanding the current strengths and limitations of the FHIR® standard for supporting data collection and exchange in clinical research.

2. Methods

FHIR® mapping was performed to assess the level of HL7® FHIR® standard completeness (or data element coverage) in supporting data collection for Study A, a federally sponsored, pragmatic cardiovascular randomized controlled trial (RCT) targeting adults (NCT03296813). A systematic approach was developed and used to map study data elements to corresponding HL7® FHIR® standard resources as implemented across three distinct academic medical centers (Sites A, B, and C); what we refer to as the site-level mapping. The mapping approach used is a framework for evaluating HL7® FHIR® standard coverage and can be utilized by others performing similar work. The development of the framework is described in our earlier work [14] for which we mapped Study A, as well as two additional studies, directly to Release 4 (R4) of the FHIR® standard (https://hl7.org/fhir/R4/index.html); what we refer to as the standard-centric or R4 mapping. The methods and results of that the R4 mapping have been published previously [14] and will not be described here.

The case report form (CRF) and data dictionary for the parent study were utilized for the site-level FHIR® mapping. Three independent mappers across three sites mapped study data elements to corresponding data elements within their respective EHRs (CRF-to-EHR mapping). Data elements available within the EHR were then reviewed to determine the corresponding FHIR® resource(s) (if applicable) available at the site via the site-specific FHIR® servers (CRF-to-FHIR Mapping). A data element was considered "Available in FHIR" if (1) it had a corresponding data element available within the EHR *and* (2) the EHR data element was mapped to a FHIR® resource or if the data could be captured or derived using a combination of data elements across one or more FHIR® resources at the site.

Upon completion of the mapping at each site, results were reviewed and analyzed by an independent adjudicator. Discussions were conducted with the full mapping team to address questions about and/or inconsistencies between site mappings. Once the mapping was finalized, an analysis was performed to: (1) calculate FHIR® coverage at each Site, (2) compare the individual site-level mapping to the preliminary standard-centric or R4 mapping [14], and (3) compare FHIR® coverage across all three sites. FHIR® coverage was calculated as a percentage of the total data elements "Available in FHIR" compared to the total data elements. Here, we present our preliminary results.

3. Results

A total of 155 data elements were mapped to the HL7® FHIR® standard resources available at each Site. At the time of the mapping, Sites A and C were supporting the November 2019 release of the Epic EHR system, while Site B was on the earlier May 2019 version of the Epic EHR system. Accordingly, Sites A and C had access to resources from three versions of FHIR® (DSTU2, STU3, and R4); whereas, Site B was only able to access DSTU2 resources.

Individually, across the three sites, we found that 52-67% of the total study data elements (N=155) were available via FHIR® (Table 1). After excluding data elements not typically or consistently documented within EHRs, i.e., questionnaire (QA) and study administration (Admin) data, we found that close to 69-89% of the study data elements (n=117) were available via FHIR® (Table 2).

Table 1. Total Study (CRF) Data Elements Available in FHIR® across all Sites (N=155)

	Available in EHR (n(%))	Available in FHIR® (n(%))
Site A	98 (63%)	81 (52%)
Site B	104 (67%)	92 (59%)
Site C	108 (70%)	104 (67%)

Table 2. Data Elements Available in FHIR®, excluding QA and Admin Data Elements (n=117)

	Available in EHR (n(%))	Available in FHIR® (n(%))
Site A	98 (83%)	81 (69%)
Site B	104 (89%)	92 (79%)
Site C	108 (92%)	104 (89%)

Agreement rates were also calculated to compare the site mappings against one another and against the original R4 mappings (Table 3).

Table 3. FHIR® Mapping Agreement Rates, across Sites and in comparison to R4 (n=117)

	R4 (n(%))	Site A (n(%))	Site B (n(%))
Site A	84 (72%)	-	-
Site B	88 (75%)	88 (81%)	-
Site C	99 (85%)	90 (84%)	102 (90%)

Note. The denominator for Sites-to-R4 comparisons (n=117). The denominators for Site-to-Site comparisons were calculated based on the data elements "Available in EHR" for the sites being compared: A-to-B = 109; A-to-C = 107; B-to-C = 113.

4. Discussion

The results presented here quantify data availability and classify data elements to determine the level of HL7® FHIR® standard completeness (as implemented across various sites) in supporting the data collection needs across three distinct sites for a federally sponsored, pragmatic cardiovascular RCT targeting adults. On average, between 60% (N=155; full CRF) and 79% (n=117; excludes study data elements not typically collected within an EHR) of the total data elements could be directly extracted from the EHR to the EDC through the use of a FHIR-based, EHR-to-EDC eSource application. Moreover, when we looked solely at the total data elements currently available within each site's EHR, we found that, on average, 103 of 155 study data elements were available in the EHR; and, of those available in the EHR, approximately 89% were available via FHIR®. These results are promising in that they indicate the maximum potential reduction in manual abstraction across sites. Beyond that, the agreement rate across sites was 85%, on average, which indicated that (for this particular study) the mapping across sites and the availability of data via FHIR® would be relatively consistent across other sites using the same EHR vendor. We anticipate that the slight differences in mapping across these sites were likely due to differences in EHR versions. For example, Sites A and C were both using the same version of the EHR (same vendor) and had a much higher agreement rate than when compared to Sites B, which was using an older version at the time of the mapping. Inconsistencies in data availability and mapping also could be explained by differences across sites with regards to (1) what data is documented and (2) how the data is documented within the EHR.

As a whole, these results offer insight into the utility of the FHIR® standard in supporting the direct extraction of EHR data for clinical research, demonstrating the maximum potential coverage of the standard (currently) to support eSource data collection for multicenter clinical research studies. These results are comparable with the findings in our recently published work [14] in which we evaluated and compared content coverage across Study A and two additional studies to the base standard (FHIR® R4 and the US Core Implementation Guide) in an effort to improve the generalizability of the results beyond a single use case. Additional work is currently underway to evaluate how various other studies compare across distinct, site-specific implementations of FHIR®.

5. Conclusion

Although eSource solutions have relied on data exchange standards to improve data collection, none have evaluated the coverage of those standard for supporting clinical studies. The work performed here addresses this gap by evaluating the availability of

HL7® FHIR® resources to support data collection for a multicenter clinical research study. The results achieved here will help to further advance the development of the FHIR® standard in order to meet the data collection needs of multicenter clinical research studies.

Funding

References

[1]   DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: New estimates of R&D costs. J Health Econ. 2016 May;47:20-33.

[2]   Getz KA, Campo RA. Trial watch: Trends in clinical trial design complexity. Nat Rev Drug Discov. 2017May;16(5):307.

[3]   Getz KA, Campo RA. New benchmarks characterizing growth in protocol design complexity. Ther Innov Regul Sci. 2018;52(1):22-28.

[4]   ONC. Quick Stats. (Office of the National Coordinator for Health Information Technology). 2018. Available at: https://dashboard.healthit.gov/quickstats/quickstats.php

[5]   HL7. HL7 FHIR Release 4. Health Level Seven. 2019. Available at: https://hl7.org/fhir/R4/index.html

[6]   CMS. CMS Interoperability and Patient Access final rule (CMS-9115-F). 85 FR 25510(Final Rule). Centers for Medicare & Medicaid Services (CMS), HHS. 2020 May 1. Available at: https://www.federalregister.gov/documents/2020/05/01/2020-05050/medicare-and-medicaid-programs-patient-protection-and-affordable-care-act-interoperability-and

[7]   ONC. 21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program. 85 FR 25642(Final Rule). Office of the National Coordinator for Health Information Technology (ONC), Department of Health and Human Services (HHS). 2020 May 1. Available at: https://www.federalregister.gov/d/2020-07419

[8]   Garza M, Myneni S, Nordo A, Eisenstein EL, Hammond WE, Walden A, Zozus M. eSource for Standardized Health Information Exchange in Clinical Research: A Systematic Review. Stud Health Technol Inform. 2019;257:115-124.

[9]   Sung NS, Crowley WF, Genel M, Salber P, Sandy L, Sherwood LM, Johnson SB, Catanese V, Tilson H, Getz K, Larson EL, Scheinberg D, Reece EA, Slavkin H, Dobs A, Grebb J, Martinez RA, Korn A, Rimoin D. Central challenges facing the national clinical research enterprise. JAMA. 2003;52(1):1278-1287.

[10]   Embi PJ, Payne PR. Clinical research informatics: Challenges, opportunities and definitions for an emerging domain. J Am Med Inform Assoc. 2009;16(3):316-327.

[11]   Nahm M. Measuring data quality. Good Clinical Data Management Practices (GCDMP). Society for Clinical Data Management. 2012;Version 2000 - present. Available at: https://www.scdm.org

[12]   Zozus MN, Pieper C, Johnson CM, Johnson TR, Franklin A, Smith J, Zhang J. Factors affecting accuracy of data abstracted from medical records. PLoS One. 2015;10(10):e0138649.

[13]   Nordo A, Eisenstein EL, Garza M, Hammond WE, Zozus MN. Evaluative Outcomes in Direct Extraction and Use of EHR Data in Clinical Trials. Stud Health Technol Inform. 2019;257:333-340.

[14]   Garza MY, Rutherford M, Myneni S, Fenton S, Walden A, Topaloglu U, Eisenstein E, Kumar K, Zimmerman K, Rocca M, Gordon GS, Hume S, Wang Z, Zozus M. Evaluating the coverage of the HL7 FHIR standard to support eSource data exchange implementations for use in multi-site clinical research studies. AMIA Annu Symp Proc. 2020 Nov.