OFFICE OF THE SECRETARY
PATIENT-CENTERED OUTCOMES RESEARCH TRUST FUND

Sentinel

U.S. FOOD & DRUG ADMINISTRATION

White Paper:

Feasibility of Using the Fast Healthcare Interoperability Resources (FHIR) Specification to Enable EHR Data Linkage with Administrative Claims, with T-MSIS Data as a Motivating Example

Prepared for:

U.S. Department of Health and Human Services, Office of the Secretary for Planning and Evaluation

Prepared by:

Department of Population Health Sciences, Duke University School of Medicine:
Keith Marsolo, Kelley Jones, Bradley G. Hammill

Office of Surveillance and Epidemiology, Center for Drug Evaluation and Research, Food and Drug Administration:
Sarah Dutcher, David Moeny, Lucia Menegussi

National Library of Medicine, National Institutes of Health:
Liz Amos, James Mork, Nick Williams

Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute:
Judith Maro

July 11, 2024

Acknowledgements and Disclaimers

This work was funded by the U.S. Food and Drug Administration (FDA) through the Department of Health and Human Services (HHS) Contract number 75F40119D10037 Task Order 75F40119F19001. The Office of the Secretary Patient-Centered Outcomes Research Trust Fund (OS-PCORTF) funding was made available to FDA by the Office of the Assistant Secretary for Planning and Evaluation (ASPE) through Interagency Agreement (IAA) 75F4012135002.

The views expressed in this white paper are those of the authors and do not necessarily represent the official position of the U.S. Department of Health and Human Services (HHS) or the Food and Drug Administration (FDA).

Table of Contents

Executive Summary

With funding through the Office of the Secretary-Patient-Centered Outcomes Research Trust Fund (OS-PCORTF), the "Making Medicaid Data More Accessible Through Common Data Models and FHIR APIs" project was a collaboration between the Food and Drug Administration (FDA) and the National Institutes of Health (NIH)/National Library of Medicine (NLM). It included an activity to transform Centers for Medicare and Medicaid Services (CMS) Transformed Medicaid Statistical Information System (T-MSIS) Transformed T-MSIS Analytical extract File (TAF) Research Identifiable Files (RIF) data into two different Common Data Models (CDMs). [1] The aim of this first activity was to make the T-MSIS TAF data interoperable and therefore easily accessible to researchers and analysts by removing the need for them to reformat the data. CDMs help ensure structural and semantic consistency across data sources while also facilitating the replication or recreation of previous analyses, especially when examining whether results remain consistent in a different population or when modifying study design or analytical parameters.

In addition, the project also had an activity to conduct a feasibility assessment of using the Fast Healthcare Interoperability Resources (FHIR) specification to enable the linkage of electronic health record (EHR) data with administrative claims sources maintained by payers. Linking administrative claims with EHR data expands the type and scope of research questions that can be answered with each source individually, supplementing the "complete capture" of medically attended events found in claims with the granular information recorded within the EHR like vital signs, laboratory results, and inpatient medication administrations.

Given the overall project focus on Medicaid data, we describe an example use case for the linkage between T-MSIS TAF and EHR data to highlight key considerations. In this scenario, the linkage would occur between CMS claims, with CMS as steward of patient identifiers in the T-MSIS data, and EHR data from EHR data holders such as individual researchers or health systems. Although designed as an example, the general process outlined in this white paper could be replicated when using FHIR application programming interfaces (APIs) to enable the linkage of EHR data with data from other payers (e.g., health plan issuers).

This white paper is organized as follows: brief background providing key terms and definitions related to T-MSIS, FHIR, and legislation related to data sharing; Technical Requirements outlining considerations for the scope of additional EHR data to be included and technical aspects of obtaining and linking data, and aspects of data governance that concern the use of those data; and Other Considerations describing other factors that impact the use of a linked claims-EHR dataset.

Background

T-MSIS: T-MSIS is a compilation of data from Medicaid and Children's Health Insurance Program (CHIP) programs across U.S. states and territories and the District of Columbia (DC) (referred to as "reporting entities" throughout this document). Reporting entities send their data to T-MSIS in a format specifically defined by the CMS T-MSIS team. This format aligns definitions nationally across agencies but represents another data format that health services researchers, policymakers, or program administrators must learn to leverage in their data in analyses, hence the interest in transforming the data into other CDMs.

FHIR: The FHIR standard is published and maintained by Health Level Seven International (HL7®).[2] It allows for the exchange of healthcare-related data within a framework that defines both the structure of the data (e.g., data format, required / optional fields) as well as the underlying content (e.g., value sets, associated terminologies) to ensure that records are compatible across disparate sources. FHIR relies on the use of APIs, which are a method of exchanging small discrete pieces of information and are utilized in many web applications. One of the main drivers of its development was to serve as a standard for data exchange that was more modern and straightforward to implement than previous HL7 standards.

The FHIR standard was designed to facilitate the exchange of health information between health systems to support clinical care across a variety of settings (including decision support), as well as the exchange of health information to support other workflows, including those related to public health, research, and health plan administration. The FHIR specification was also designed to enable patient access to their healthcare data. Examples of systems that leverage FHIR APIs to enable patient access include patient portals and mobile applications that integrate data from the patient's EHR, pharmacy records, and health activity monitors.[3,4] The electronic care (eCare) plan, a mobile electronic care plan application developed for patients, was an OS-PCORTF initiative that used FHIR APIs.[5]

CMS has been a major driver in promoting the FHIR specification and other standards adopted in the Office of the National Coordinator for Health Information Technology's (ONC's) 21st Century Cures Act Final Rule as part of their Promoting Interoperability Program and their Interoperability Patient Access and Interoperability and Prior Authorization Final Rules (see below).[6] Examples of technical implementations from CMS supporting these rules include the Medicare Blue Button for enrollees (another OS-PCORTF initiative)[7] and a Beneficiary Claims Data API for Accountable Care Organizations.[8]

Fostering data sharing through the 21st Century Cures Act: The 21st Century Cures Act, passed by the United States Congress in 2016, contained several provisions that were intended to promote interoperability and foster data sharing across the healthcare industry. These provisions included the sharing of information among payers, health systems, and larger health information networks, with specific efforts to prevent information blocking (interfering with permitted exchange, access or use of information). In May 2020, CMS published the final version of the CMS Interoperability and Patient Access rule,[9] which requires or encourages payers to adopt technical standards including FHIR APIs to improve the exchange of data by sharing information with patients or exchanging data between payers at the patient's request (to better create a longitudinal record).[9] The rule applies to all payers impacted by CMS regulations, including Medicaid, CHIP, Medicare Advantage and Qualified Health Plan (QHP) issuers on the Federal-facility Exchanges. However, the mandate to share clinical data with patients only applies if the patient is a current beneficiary and the payers themselves receive EHR data from

providers. Even so, this rule will continue to promote increased sharing of administrative claims across payers and adoption of the underlying infrastructure to support FHIR API-based linkage.

ONC has defined a number of criteria that health IT systems must meet in order to obtain and maintain their ONC certification; providers who wish to maintain incentive payments under the Medicare and Medicaid EHR Incentive Programs are required by statute to use Certified EHR Technology.[10] This includes capabilities to produce the data elements specified in the United States Core Data for Interoperability (USCDI) and that these elements be available for access, use and exchange using the FHIR standard. This includes making these data available to patients. The rule has also resulted in public and private health systems gradually implementing FHIR capabilities within their health IT systems.

Technical Requirements

Motivating use case: The hypothetical use case behind the proposed linkage activity is one where a requestor, who could be an individual researcher or organization, either has or is planning to obtain EHR data with personal identifiers for a cohort of patients. The requestor wishes to obtain and link administrative claims from TAF to EHR data for this cohort.

Traditionally the EHR data holders would send a finder file to CMS/T-MSIS and receive either the crosswalk to their Beneficiary ID if they already had TAF data, or an extract of the TAF administrative claims themselves. In this revised scenario, rather than send data to CMS, or more generally to the applicable health plan, as a finder file, the EHR data holder would like to use FHIR APIs and associated workflows to identify patients and receive the underlying claims records.

For this white paper, we are limiting our discussion of the use of the FHIR specification to the initial patient matching process, rather than for the subsequent transfer/exchange of patient records (particularly given potential revisions to the CMS policy on data transfers).[11] The standard FHIR API works well to access and exchange data on a patient-by-patient or observation-by-observation basis. Large data transfers via FHIR require the use of "bulk FHIR" capabilities and would be similar to the data transfer process currently employed by payers to send or receive larger numbers of patient records. We are also assuming that data use agreements are in place to allow the transfer of identifiers to support patient matching and for the creation and subsequent use of a linked EHR-claims dataset.

***Patient matching*:** Table 1 below provides a list of data elements that are currently part of T-MSIS that could potentially be used for linkage with EHR data, along with information on whether they are required components of the T-MSIS and the US Core version of the FHIR specification.[12] Other payers likely have similar data elements within their source systems. There are only a handful of elements that are required under the FHIR US Core (e.g., patient name and gender), but many of the elements that are supported by the standard would be expected to be present within the EHR,[13,14] and would therefore be available as part of the FHIR-based data exchange. For the patient matching process, which would be conducted by CMS in this scenario, we are assuming that a deterministic method would be used (e.g., directly comparing personally identifiable information (PII) like name, address and date of birth). Probabilistic methods that use different combinations of PII to match patients or a referential approach that attempts to match patients to a reference population also may be considered. Note that this last method may not be suitable for populations with large percentages of pediatric patients depending on the frequency of updates to the reference dataset. A variety of algorithms have been evaluated for patient matching that could be used as a starting point for payers.[15,16]

***Table 1**. Current T-MSIS data elements that could potentially be used for EHR data linkage*

Data Element	Status in T-MSIS	Status in FHIR US Core Patient Profile (3.1.1)
Patient first name	Conditional	Required, Must Support*
Patient last name	Required	Required, Must Support*
Social Security Number	Required if known	Required†
Medical Record Number	Not supported	Required†
Beneficiary Identifier	Required	Must Support
Date of Birth	Expected	Optional, Must Support
Sex (Administrative Gender)	Expected	Required, Must Support‡
Phone	Required	Optional, Must Support
Address	Required	Optional, Must Support
City	Required	Optional, Must Support
State	Required	Optional, Must Support
Zip Code	Required	Optional, Must Support

*First name and/or last name shall be present or a data absent reason extension shall be present
†Identifier is a required field in the US Core Patient profile, but it does not specify that you must use SSN or MRN
‡This is called "gender" in the US Core Patient profile. It is a required field. Birth sex is Optional, Must Support (extension)

Data governance: In describing the use of the FHIR specification to support patient matching to another data source like T-MSIS, it is important to note that the current use of FHIR at health systems for external data exchange is primarily in support of clinical care or for other operational purposes (e.g., prior authorization). For the purpose of research or public health surveillance, the transmission of identifiers to external organizations using FHIR APIs in order to link to another data source may require changes to health system information technology policies and/or specific language in contracts or data use agreements to cover this activity.

Other Considerations

While not directly related to the data transfer and patient-matching process, the following issues should be taken into consideration when linking administrative claims and EHR data.

Encounter-level linkage: When merging claims and EHR data, or EHR data from one source with another, decisions must be made on how to de-duplicate information within a patient record. This can occur if EHR data for the "same" encounter or service are received from multiple sources. For example, a physician orders lab work for their patient, which is collected and analyzed at an independent facility; both the physician and the lab facility may send the same results, with varying degrees of completeness. Similarly, a health system may have diagnosis records that were generated by the clinician or health system billing process that are to be merged with diagnosis records that are available from the patient's health plan(s). Given that multiple EHR encounters can be associated with a single claims encounter, or vice versa, it may be difficult to create encounter-based identifiers that span multiple sources. Initially, it may be simplest for researchers to rely on date of service when running analyses that combine records from both EHR and claims sources.

Quality control: A key aspect of any data asset is to ensure a minimum level of quality control. T-MSIS, for instance, has a full list of data quality measures that are checked against each reporting entity's

monthly submission. These programs are made publicly available by CMS on GitHub (https://github.com/CMSgov/T-MSIS-Data-Quality-Measures-Generation-Code/blob/main/Thresholds.xlsx). In addition, the CMS DQ Atlas (https://www.medicaid.gov/dq-atlas/welcome) provides a wealth of information regarding the quality of the T-MSIS Analytic Files (TAF) data on a yearly basis. Reporting entities' data quality is rated per topic per year (e.g., race and ethnicity 2020, income 2020) on a scale of low, medium, and high concern, unusable, and unclassified, with extensive details provided on the methodology used for each measure. Often the rating is based on the amount of missing values, rather than the values being suspect (e.g., race and ethnicity). Investigators can use this information to determine which reporting entities and years have adequate data to analyze, specific to their research question. Other payers may have similar quality measures for their data that can potentially be shared with interested investigators.

***Data latency**:* Due to a number of factors, including delays in submitting claims, revisions, and the curation of the analytic files, complete administrative claims data are not available to investigators for several months or years after an encounter. As an example, within T-MSIS, as of early 2024, complete TAF data is available through 2021, with preliminary data available for 2022. This latency will limit the potential for real-time linkage and will best support retrospective analyses.

Cost: The cost for payers to set up and maintain FHIR APIs to support patient linkage may be non-trivial. This could require mapping all relevant patient identifiers to a server that is connected to a FHIR API, implementing the required workflows, and making the relevant connections to the final storage location of the underlying claims data (e.g., the Virtual Research Data Center in the context of T-MSIS). For CMS and T-MSIS, this would require an increase to their existing budget as this is not within their current scope of activities. To offset cost burdens, incentives may be needed to motivate other payers to adopt a similar approach over whatever existing processes they have in place.

Conclusion

The goal of the " Making Medicaid Data More Accessible Through Common Data Models and FHIR APIs" project is to increase researcher access to and uptake of Medicaid data. The multi-part project is a collaboration between two federal agencies, FDA and NIH/NLM. One part involved the transformation of T-MSIS TAF-RIF data into two CDM formats, which provides a more familiar data structure for improved access to a broader suite of analytic tools and methods. Another part was about how FHIR APIs can be used to create linked resources that combine EHR data with administrative claims. The use case presented in this white paper focused on a theoretical linkage between CMS/T-MSIS TAF-RIF and EHR data, but these considerations also apply to linkages with data from other health insurance payers. APIs provide an alternative strategy for data exchange over the traditional "finder file" approach. While there may currently be governance challenges and cost barriers in creating the necessary infrastructure to support these workflows, the widespread adoption of FHIR APIs by health systems and payers means that an API-based approach could become more tractable in the near future.

References

1. Making Medicaid Data More Accessible Through Common Data Models and FHIR APIs. Office of the Assistant Secretary for Planning and Evaluation. Accessed 6/4/2023, https://aspe.hhs.gov/making-medicaid-data-more-accessible-through-common-data-models-fhir-apis
2. Health Level 7. FHIR Specification Home Page. Accessed August 24, 2015. http://hl7.org/fhir/
3. Dhruva SS, Ross JS, Akar JG, et al. Aggregating multiple real-world data sources using a patient-centered health-data-sharing platform. *npj Digital Medicine*. 2020;3(1):60. doi:10.1038/s41746-020-0265-z
4. El-Sappagh S, Ali F, Hendawi A, Jang J-H, Kwak K-S. A mobile health monitoring-and-treatment system based on integration of the SSN sensor ontology and the HL7 FHIR standard. *BMC Medical Informatics and Decision Making*. 2019;19(1):97. doi:10.1186/s12911-019-0806-z
5. Assistant Secretary for Planning and Evaluation (ASPE). Data Capacity for Patient-Centered Outcomes Research through Creation of an Electronic Care Plan for People with Multiple Chronic Conditions 2.0: Development of the Patient-Facing Application. April 15, 2024. https://aspe.hhs.gov/data-capacity-patient-centered-outcomes-research-through-creation-electronic-care-plan-people
6. Centers for Medicare and Medicaid Services (CMS). Policies and Regulations. Accessed March 25, 2024. https://www.cms.gov/priorities/key-initiatives/burden-reduction/interoperability/policies-and-regulations
7. Office of Health Policy, Evaluation ASoPa. HHS Projects to Build Data Capacity for Patient-Centered Outcomes Research: Completed Projects FY2010 through FY2015. 2017. https://aspe.hhs.gov/sites/default/files/migrated_legacy_files/183196/OSPCORTFCompletedProjectsReport.pdf
8. Centers for Medicare and Medicaid Services (CMS). Beneficiary Claims Data API. Accessed March 25, 2024. https://bcda.cms.gov/
9. Advancing Interoperability and Improving Prior Authorization Processes Proposed Rule CMS-0057-P: Fact Sheet. Centers for Medicare & Medicaid Services; 2022. https://www.cms.gov/newsroom/fact-sheets/advancing-interoperability-and-improving-prior-authorization-processes-proposed-rule-cms-0057-p-fact
10. Office of the National Coordinator for Health Information Technology (ONC). Certification of Health IT. Accessed November 11, 2023. https://www.healthit.gov/topic/certification-ehrs/certification-health-it
11. Centers for Medicare and Medicaid Services (CMS). Important Research Data Request & Access Policy Changes. Accessed March 25, 2024. https://www.cms.gov/data-research/files-order/data-disclosures-and-data-use-agreements-duas/important-research-data-request-access-policy-changes-0
12. Health Level 7 (HL7). US Core Implementation Guide. Accessed March 25, 2024. https://www.hl7.org/fhir/us/core/
13. Culbertson A, Goel S, Madden MB, et al. The Building Blocks of Interoperability. A Multisite Analysis of Patient Demographic Attributes Available for Matching. *Appl Clin Inform*. 04 05 2017;8(2):322-336. doi:10.4338/ACI-2016-11-RA-0196
14. Marsolo K, Kiernan D, Toh S, et al. Assessing the impact of privacy-preserving record linkage on record overlap and patient demographic and clinical characteristics in PCORnet®, the National Patient-Centered Clinical Research Network. *J Am Med Inform Assoc*. Feb 16 2023;30(3):447-455. doi:10.1093/jamia/ocac229

15. Grannis SJ, Williams JL, Kasthuri S, Murray M, Xu H. Evaluation of real-world referential and probabilistic patient matching to advance patient identification strategy. *J Am Med Inform Assoc*. 2022;29(8):1409-1415. doi:10.1093/jamia/ocac068
16. Patient Identity and Patient Record Matching. HealthIT.gov. Accessed 9/6/2023, https://www.healthit.gov/topic/interoperability/standards-and-technology/patient-identity-and-patient-record-matching