

Brief Communication

Integrating real-world data to assess cardiac ablation device outcomes in a multicenter study using the OMOP common data model for regulatory decisions: implementation and evaluation

Yue Yu¹, Guoqian Jiang², Eric Brandt³, Tom Forsyth³, Sanket S. Dhruva ⁴, Shumin Zhang⁵, Jiajing Chen³, Peter A. Noseworthy⁶, Amit A. Doshi⁷, Kimberly Collison-Farr³, Dure Kim⁸, Joseph S. Ross⁹, Paul M. Coplan^{5,10}, and Joseph P. Drozda Jr³

¹Department of Quantitative Health Sciences, Mayo Clinic, Rochester, Minnesota, USA, ²Department of Artificial Intelligence and Informatics, Mayo Clinic, Rochester, Minnesota, USA, ³Mercy Research, Mercy, Chesterfield, Missouri, USA, ⁴School of Medicine, University of California San Francisco, and Section of Cardiology, Department of Medicine, San Francisco Veterans Affairs Medical Center, San Francisco, California, USA, ⁵MedTech Epidemiology and Real-World Data Sciences, Office of the Chief Medical Officer, Johnson & Johnson, New Brunswick, New Jersey, USA, ⁶Department of Cardiovascular Medicine, Mayo Clinic, Rochester, Minnesota, USA, ⁷Mercy Clinic, Mercy, St. Louis, Missouri, USA, ⁸National Evaluation System for Health Technology Coordinating Center (NESTcc), Medical Device Innovation Consortium, Arlington, Virginia, USA, ⁹Department of Internal Medicine, Yale School of Medicine, and the Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, New Haven, Connecticut, USA and ¹⁰Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Corresponding Author: Guoqian Jiang, MD, PhD, Department of Artificial Intelligence and Informatics, Mayo Clinic, 200 First Street, SW, Rochester, MN 55905, USA; jiang.guoqian@mayo.edu

Received 10 May 2022; Revised 10 August 2022; Editorial Decision 7 December 2022; Accepted 5 January 2023

ABSTRACT

The objective of this study is to describe application of the Observational Medical Outcomes Partnership (OMOP) common data model (CDM) to support medical device real-world evaluation in a National Evaluation System for health Technology Coordinating Center (NESTcc) Test-Case involving 2 healthcare systems, Mercy Health and Mayo Clinic. CDM implementation was coordinated across 2 healthcare systems with multiple hospitals to aggregate both medical device data from supply chain databases and patient outcomes and covariates from electronic health record data. Several data quality assurance (QA) analyses were implemented on the OMOP CDM to validate the data extraction, transformation, and load (ETL) process. OMOP CDM-based data of relevant patient encounters were successfully established to support studies for FDA regulatory submissions. QA analyses verified that the data transformation was robust between data sources and OMOP CDM. Our efforts provided useful insights in real-world data integration using OMOP CDM for medical device evaluation coordinated across multiple healthcare systems.

Key words: medical device, OMOP CDM, UDI, medical data standardization

Lay Summary

Data standardization is essential for integrating and linking medical device identification information to diverse data sources across different institutions for research use. A common data model (CDM) could facilitate medical device data standardization and integration. In this study, we described the application of the Observational Medical Outcomes Partnership (OMOP) CDM to support ThermoCool medical device real-world evaluation involving 2 healthcare systems, Mercy Health and Mayo Clinic. Specifically, we collected the device data from the supply chain database by unique device identifiers and related patients' clinical data from the electronic health record (EHR) database. Then, OMOP CDM-based database was designed and implemented to standardize and integrate device data and EHR data together in both Mercy Health and Mayo Clinic. Quality assurance verified that the data transformation was robust between data sources and OMOP CDM. Our efforts provided useful insights into real-world data integration using OMOP CDM for medical device evaluation coordinated across multiple healthcare systems. And our process could be generalized for use by other institutions.

INTRODUCTION

Data standardization is essential for integrating and linking of medical device identification information to diverse data sources, thereby enabling interoperability solutions that produce networks from which the data composite can support safety and effectiveness evaluations.¹ With the implementation of unique device identifiers (UDIs) in electronic health information sources as a key strategy from the US Food and Drug Administration (FDA),^{2,3} a UDI Research and Surveillance Database (UDIR)⁴ containing the linked clinical and device information were proposed in the research community to enable longitudinal assessment of device safety, performance and quality. Some studies have discussed how to use UDI to facilitate real-world medical device surveillance research.^{5,6} However, for these studies, a big challenge remains regarding how to standardize and integrate medical device and electronic health record (EHR)-related data together, especially for studies that include multiple institutions. Meanwhile, a variety of common data models (CDMs) have been developed to provide a standardized approach to store and organize clinical research data.⁷⁻¹¹ The assumption is that CDM-based solutions can facilitate meaningful collaborations by standardizing data collection and analysis processes across institutions, including distributed analytics. However, CDM-based solutions have not been widely used for medical device evaluation studies,^{12,13} and the applicability of CDMs to medical device evaluation studies,¹⁴ particularly whether they capture sufficient granularity of device identifiers and aggregate codes for procedures, remains an unanswered question.

In this context, we had an opportunity to implement and assess a CDM-based approach to integrate real-world data (RWD) to assess cardiac ablation catheter outcomes in a multicenter study. This was performed as part of a National Evaluation System for health Technology Coordinating Center (NESTcc) Test Case. The specific catheters of interest are the ThermoCool Smarttouch catheters, initially approved by the FDA in February 2014, and the ThermoCool Smarttouch Surround Flow catheters, initially approved by the FDA in August 2016. Both are products of Biosense Webster, Inc. (part of the Johnson & Johnson family companies) (Irvine, CA). After determining feasibility of leveraging RWD,^{14,15} the investigators, NESTcc and the medical device sponsor agreed to proceed with a follow-up study of the same medical devices, developing the data standards and analytic approach to turn EHR RWD into real-world evidence (RWE) in label expansion studies for submission to the FDA Center for Devices and Radiological Health (Electrophysiology). The focus of this report is to share how we, Mercy Health and

Mayo Clinic, standardized the ThermoCool device data and related EHR data using the Observational Medical Outcomes Partnership (OMOP) CDM to store the data at each institution. Using the UDI as an index key, we integrated the medical device data with the EHR data to demonstrate how the OMOP CDM can be used for medical device studies across institutions to support regulatory decision-making, evaluating the data quality of the distributed system in the process.

METHODS

Establishment of work group and preparation for the OMOP CDM implementation

The rationale for having a distributed analysis was that the required study sample size for adequate statistical power could not be obtained at one healthcare system alone. In addition, inclusion of more than one healthcare system was required to expand the generalizability of the study findings. Sharing patient EHR data, even if deidentified, between healthcare systems was complicated; this necessitated separate analyses of treatment effects in each healthcare system with as much analytical and variable consistency as possible. Use of a CDM to create a common data structure across the healthcare systems enabled this without sharing patient level data. Moreover, we chose OMOP CDM as a standard CDM to deploy the data repository across the 2 healthcare systems, Mercy and Mayo Clinic, because it is an open source CDM, and some OMOP CDM infrastructure had already been existed at Mayo Clinic.

To facilitate the OMOP CDM implementation and validation across the 2 healthcare systems, we established an OMOP Work Group of project investigators to coordinate defining and capturing the data elements and integrating them during the data transformation from Mercy and Mayo Clinic's EHR repositories to their OMOP CDM instances. As the initial step, the workgroup had several discussions to harmonize the fundamental definitions of data elements for the OMOP CDM implementation. The preliminary preparation process included the following tasks. First, we decided to use OMOP v5.3.1 for this project, since this version aligned with study needs and some resources already established within the Mayo Clinic environment. In addition, the latest OMOP v6.0 had not been synchronized with the open-source tools (eg, Atlas¹⁶) developed in the Observational Health Data Sciences and Informatics (OHDSI) community. Second, we chose a subset of Clinical and Health System tables present in OMOP v5.3.1 according to the data needs for the analysis. Third, to ensure the alignment across conditions, devices, drugs, observations, procedures, providers and visits at both

Mercy and Mayo Clinic, we determined the common set of standard vocabularies included in the OMOP CDM (SNOMED CT, RxNorm, LOINC, NUCC, CPT, and ICD-10 PCS) that we would use to standardize these clinical concepts. Finally, we collected data from all patients with cardiac ablation procedure codes (CPT codes), including those patients who had received ablation with one of the ThermoCool ST or STSF catheters as our primary patient population. The ThermoCool catheters were identified by the Global Trade Item Numbers (GTINs, the device identifier part of the UDI) from the supply chain system.

Device data sourcing and data integration with EHR

One of the goals of this project was to determine whether UDI could facilitate RWD-based observational studies to generate RWE. To this end, both Mercy and Mayo Clinic used a catalog of GTINs, which were captured via point-of-care scanning and recorded in the supply chain management system, to collect all cardiac ablation catheter and ablation sheath device data. Then, we integrated the device data with clinical data from EHR (the source data) as a single dataset for the OMOP CDM implementation. The supply chain system focuses on capturing device usage-related data, while the EHR is used for retrieving patients' clinical data. The patients' clinic identifiers helped us to link the medical device data to the patient data.

At Mercy, the Epic EHR (Verona, WI) and Omnicell OptiFlex (Mountain View, CA) supply management systems were utilized to isolate device records of interest, and the Mercy-developed Real-World Evidence Insights Data Platform (RWE IDP) was utilized as a transformation model to collect the EHR data extracted from Epic. Beginning in 2016, ablation catheter devices and other supplies needed for the procedures were scanned into inventory upon receipt by Mercy. At the point of use, they were scanned again to establish a relationship to the patients' clinical charts in Epic. Prior to 2016, the devices were recorded using custom billing procedures within the patient encounter data in the EHR, and these were mapped to GTINs. Clinical data elements and device-related elements were brought together in the RWE IDP using a combination of patient and encounter identifiers, including the device GTINs. All information (ie, encounters, diagnoses, procedures, devices, supplies, etc) for cohort patients was brought together in the RWE IDP, with filters for the devices of interest applied during population of the OMOP table structures. Refer to the top half of [Figure 1](#) for an illustration of this data flow.

Mayo Clinic device data were sourced from the system's Supply Information Management System (SIMS) that contains medical device data before 2018, and a Supply+ (Cardinal Health) in Epic systems that contain data since 2018. Both the SIMS and Supply+ databases record the use of the medical devices in patient care through scanning barcodes at Mayo Clinic, so these databases contain the UDI or corresponding catalog number information for a medical device. Moreover, the EHR data at Mayo Clinic was from the Unified Data Platform (UDP). UDP is an EHR data warehouse that integrates diverse data across multiple databases at Mayo Clinic and presents these data in a standard format. The EHR data from UDP were combined with the device data of SIMS and Supply+ using the Mayo Clinic patient number. The Mayo data integration model structure is shown in the bottom half of [Figure 1](#).

OMOP CDM-based repository implementation

In this step, we created data manipulation scripts to extract, transform and load (ETL) device data from the supply chain system and

clinical data from EHR into an OMOP CDM-based data repository at each site. Both healthcare systems used SQL to develop the scripts to perform the ETL task.

To ensure the mappings during the ETL process of the original EHR source values to the OMOP standard concept IDs were consistent across the 2 healthcare systems, we developed shared mappings of EHR data elements table by table. For the concept mapping of `visit_concept_id`, `observation_concept_id`, `place_of_service_concept_id`, and `specialty_concept_id`, we designed a manual mapping process to match the original concept to the OMOP standard concepts. Furthermore, for some other concepts (`condition_concept_id`, `measurement`, `drug_concept_id`, `procedure_concept_id`, `measurement_concept_id`), we developed an automatic mapping process through the OMOP `concept_relationship` table. We also randomly selected 100 standard concept IDs for each field and manually reviewed them, to make sure that the original codes (ICD, LOINC, CPT, etc) were correctly mapped to the standard OMOP concept IDs. For the `device_concept_id`, we used the GUDID device lookup API to map the device to the SNOMED CT codes and then to the OMOP standard concept IDs.¹⁷ Then, the ETL process was performed according to the data structure mappings and concept mappings. During the data ETL, the patient clinical data from EHR were transformed into 10 OMOP clinical/health system data tables, and the medical device data were converted into the `DEVICE_EXPOSURE` table.

To facilitate propensity score modeling and effectiveness and safety outcome analysis, we conducted a 2-step process to identify code lists for phenotypes, covariates, and outcome endpoints. First, we collected these code lists from literature or Sentinel Initiative reports. Next, we sent the code list to the clinical experts on our team for manual review. After manual evaluation, all confirmed code lists were loaded into 2 concept set tables in our OMOP database. The "concept_set" table records the concept set IDs and links those IDs to clinician-defined data variables such as congestive heart failure, intracardiac echocardiography (ICE), or chronic renal disease. Moreover, the specific concept code lists for each concept set were defined within the "concept_set_item" table. This table stores both the original concept codes (such as ICD, CPT) and its corresponding standard OMOP concept ID that maps from the OMOP concept relationship table.

Data quality assurance analyses on the OMOP CDM

To validate the data transformation quality of the ETL scripts, we first performed quality assurance (QA) analyses to compare records between the original database and the OMOP database. Specifically, we conducted a review of (1) record counts per variable; (2) record counts over time; and (3) null values across all tables. We then analyzed calculated outputs on OMOP data with originally recorded script outputs to (1) validate patient flow across platforms; (2) validate counts by phenotype code across platforms; (3) validate recorded encounters across care site locations; (4) validate visit record counts across platforms to ensure proper flow of patient encounters across platforms; and (5) validate device record counts across platforms. Regarding the phenotype count comparison, we used the original primary phenotype codes we had populated into the OMOP database to extract patients for ensuring consistent patient count totals with pre-extract counts on the original database.

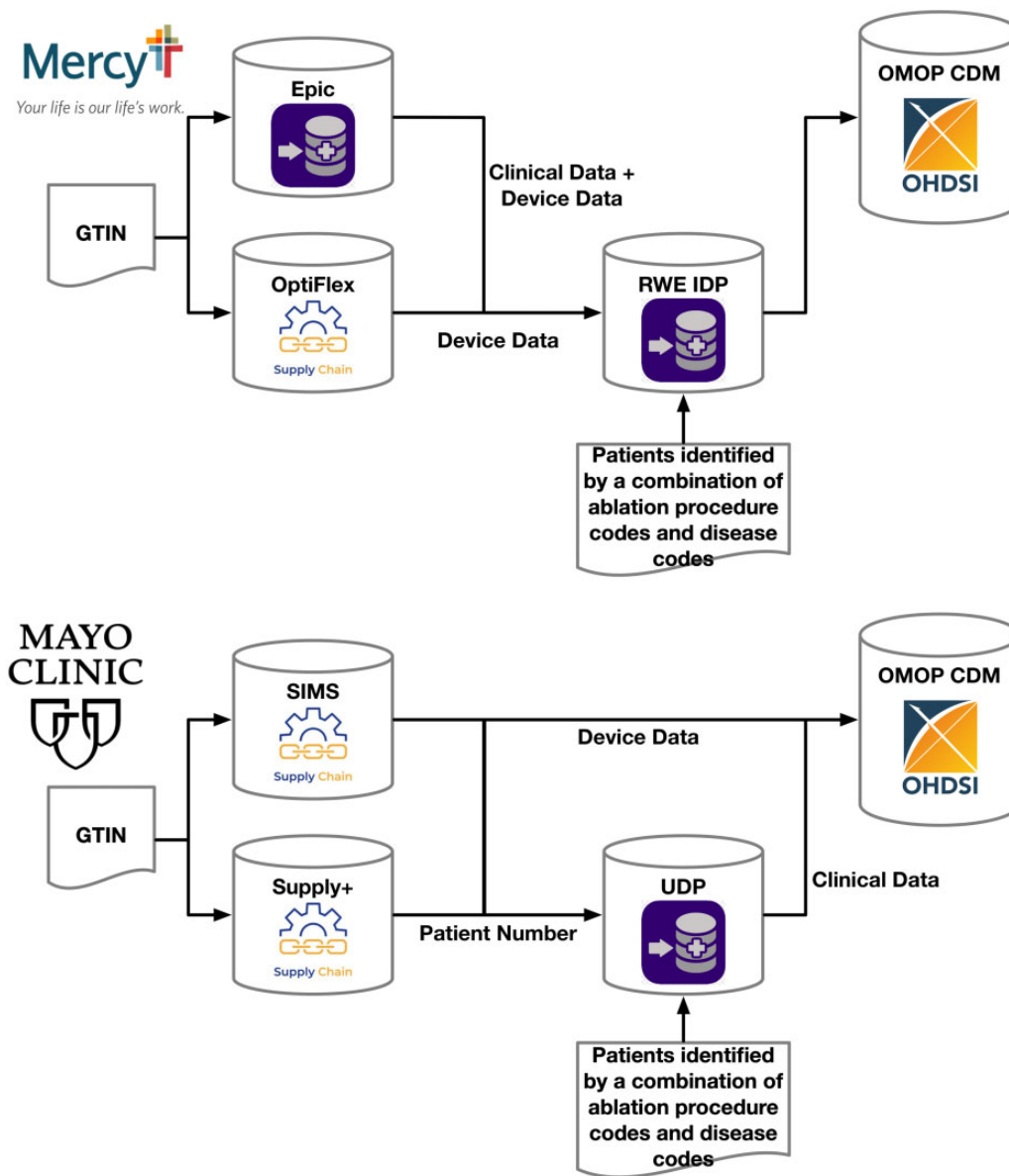


Figure 1. Data Integration model at Mercy and Mayo Clinic.

RESULTS

OMOP implementation results

In total, 20 757 patients at Mayo and 8 449 patients at Mercy between January 1, 2014 and April 30, 2021 were identified as a general population using cardiac ablation-related procedure codes and loaded into the OMOP CDM (shown in Figure 2). In addition, we counted and aggregated the device records by their GTIN device identifiers (shown in Table 1). The device record counts are distinct from the patient counts, since the catheters may be used for cardiac ablation procedures other than persistent AF and VT ablation (eg, paroxysmal AF or atrial flutter). Five of 16 GTINs associated with 5 125 ThermoCool device usage records (for 4 636 patients) were identified at Mayo Clinic, and most of the records were collected by 2 GTINs. Mercy collected 10 724 records (for 3 411 patients) with 4 GTINs in their supply chain system; records were almost evenly distributed across the 4 GTINs. In addition to those ThermoCool ST or STSF catheter patients, we also included

16 121 patients at Mayo and 5 038 patients at Mercy who had at least one cardiac ablation procedure code or were treated using a NaviStar ThermoCool (an older ThermoCool catheter) during our observational time window. These patients were identified as a potential control group.

Figure 2 shows the record counts for each OMOP CDM table implemented at Mayo Clinic and Mercy. The patient data collected from EHRs were transferred into 10 OMOP tables. The PERSON table is one of the main tables in our OMOP CDM-based database, and it recorded the patient demographic information. The patient clinical data were converted into the LOCATION, VISIT_OCCURRENCE, PROCEDURE_OCCURRENCE, CONDITION_OCCURRENCE, DEATH, DRUG_EXPOSURE, and OBSERVATION tables. The health system-related information was stored in the CARE_SITE and PROVIDER tables. The medical device usage records from the supply chain system were transformed into the DEVICE_EXPOSURE table. The medical device record counts are slightly higher than shown in

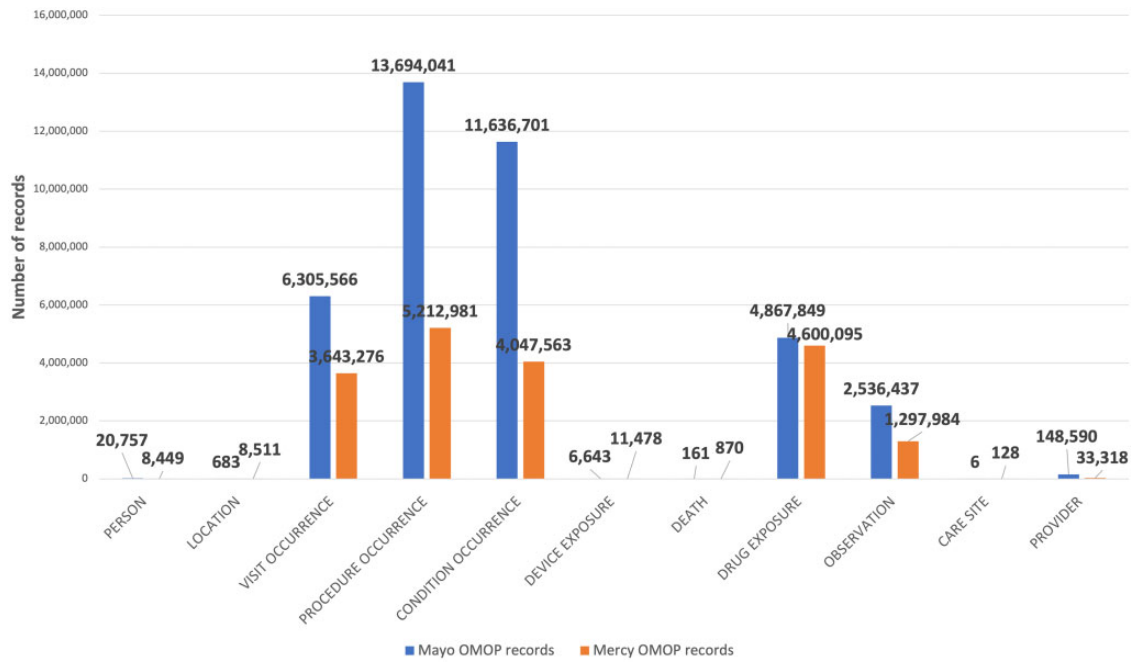


Figure 2. OMOP CDM statistics for tables used. CDM: common data model; OMOP: Observational Medical Outcomes Partnership.

Table 1. The record counts of ThermoCool devices in the 2 healthcare systems

Primary DI (GTIN)	Catalog #	Device	Mayo OMOP records	Mercy OMOP records
10846835008982	D133601	ThermoCool ST	0	0
10846835009002	D133602	ThermoCool ST	37	0
10846835009019	D133603	ThermoCool ST	0	0
10846835009163	D132701	ThermoCool ST	0	0
10846835009170	D132702	ThermoCool ST	0	0
10846835009187	D132703	ThermoCool ST	0	0
10846835009194	D132704	ThermoCool ST	7	2009
10846835009200	D132705	ThermoCool ST	3907	2600
10846835010145	D134801	ThermoCool ST SF	0	0
10846835010152	D134802	ThermoCool ST SF	0	0
10846835010169	D134803	ThermoCool ST SF	0	0
10846835010176	D134804	ThermoCool ST SF	16	2743
10846835010183	D134805	ThermoCool ST SF	1208	3372
10846835009774	D134701	ThermoCool ST SF	0	0
10846835009781	D134702	ThermoCool ST SF	0	0
10846835009798	D134703	ThermoCool ST SF	0	0
Total			5175	10 724

DI: device identifier; GTIN: Global Trade Item Number; Catalog #: Catalog Number of the device; OMOP: Observational Medical Outcomes Partnership; ST: ThermoCool SmartTouch; ST SF: ThermoCool SmartTouch SF.

Table 1 because some “NaviStar ThermoCool” device cases were included in this table as a potential control group.

We created a total of 111 concept sets to represent 20 796 standard codes used in identifying covariates and potential safety and effectiveness outcomes across the 2 healthcare systems. Figure 3 provides the overview statistics of concept sets and specific concept codes in different analysis usage categories.

OMOP QA analysis results

We compiled 26 total queries to demonstrate the faithfulness of data transformation between the OMOP CDM and source data platforms at both Mayo Clinic and Mercy. Table 2 shows the QA analy-

sis results of patient counts, indicating that all QA queries achieved perfect concordance results in patient, provider, visit, device identifier, and concept code counts between the source EHR dataset and the OMOP-transformed dataset. We considered the data successfully transferred if the counts are consistent between the source database and the target OMOP CDM. Moreover, we manually reviewed 100 concept mapping results for each of the automatic mapping fields. Supported by the comprehensive vocabulary and concept relationship designed by the OHDSI community, we found 100% mapping accuracy for the concept mapping of condition_concept_id, drug_concept_id, procedure_concept_id, and measurement_concept_id fields.

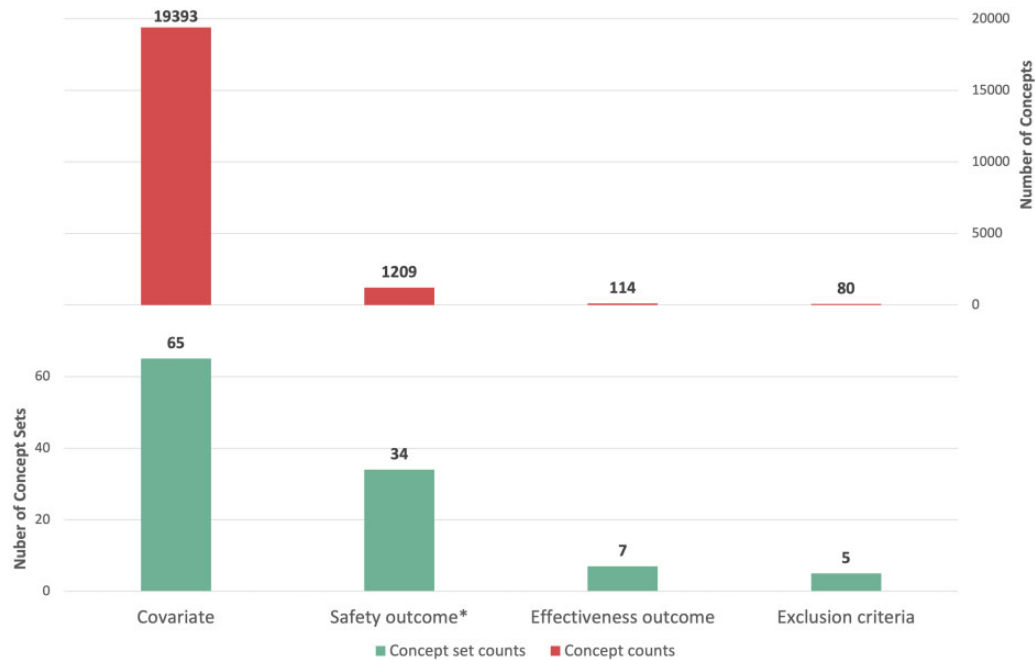


Figure 3. Concept set statistics by different analysis usage categories. *The codes in the “safety outcome” set are used to identify charts for review for potential safety events. All safety outcomes were identified by physician chart review.

DISCUSSION

In this study, we built OMOP CDM-based data repositories and evaluated the data quality in 2 healthcare systems to facilitate a real-world medical device label expansion study. To our knowledge, this study is the first retrospective comparative cohort study using data extracted from EHRs and hospital supply chain systems as the sole sources of clinical evidence for a premarket approval (PMA) device indication extension—a milestone of significant interest to manufacturers, FDA, and NESTcc.¹⁸ In the ThermoCool Phase I feasibility study, the project team investigated the CDM implementation status of 3 healthcare systems and discussed the potential use of the CDM-based approach for standardized device data capture and analytics.¹⁴ We found that the informatics approach could be used to capture study populations, device exposure, covariates, and safety and effectiveness outcomes in RWD for use in medical device studies supporting label extensions of cardiac ablation catheters. However, data quality issues may exist due to the variations of data sources in different institutions. To improve the consistency of data across institutions and support distributed data analytics, we used a CDM-based research repository to standardize data collection and analytic processes across institutions. So, in the label extension study phase of the project, 2 of 3 main participating sites of the NESTcc ThermoCool test case project that has sufficient samples for study use, Mercy Health and Mayo Clinic, implemented OMOP CDM to their data sources to improve the data consistency.

Our work made several contributions. First, by establishing the OMOP workgroup, we developed an effective process to coordinate the OMOP implementation and evaluation across 2 healthcare systems with different supply chain and EHR systems. Specifically, we aligned the data model version, database structure, data element definition, concept mapping, and ETL process between the 2 healthcare systems. Our experience illustrates how to deploy the same OMOP CDM successfully at different institutions, thereby providing a reference for the research community. Second, by integrating both device

data in the supply chain database and EHR data into the OMOP CDM, we demonstrated that the UDI could be utilized for medical device evaluation research to inform regulatory decision-making. Third, the QA analysis results demonstrated identical results between our source database and OMOP CDM-based database for patient, provider, visit, device identifier, and concept code counts, proving that the OMOP CDM-based database would satisfy the data quality requirement for the subsequent analysis tasks. Furthermore, a separate paper from our group reports the results of effectiveness and safety outcome analysis for the atrial fibrillation patients who were treated using the ThermoCool devices by using the data from our OMOP databases.¹⁹ Since the 2 healthcare systems deployed the same data model, we could design a unified data collection query and analysis script to perform distributed analytics.

We encountered some challenges during this study. Due to variation in institutional application infrastructure, it was sometimes difficult to collect the same data elements from the source databases. For example, when transforming the data into the drug exposure table of the OMOP CDM, Mercy collected all the outpatient prescribed and inpatient administered medications, while Mayo Clinic only captured inpatient administered drug data. In addition, mapping of some source EHR concepts to standard OMOP concepts was more challenging than expected. For example, to work around a lack of uniform discrete values for visit-defining variables, the label consisting of length of stay and emergency visit was utilized as a common tool for Mercy and Mayo to define stays as either inpatient (no emergency label and length of stay ≥ 24 h), outpatient (no emergency label and length of stay < 24 h), or emergency (emergency label). Finally, many encounter types could not be mapped to the OMOP standard concepts (eg, “Anticoagulation Visit” is too specific to find a mapping with a visit concept in the OMOP vocabulary); however, these were not integral to the study.

In summary, our study provided useful insights in medical device RWD integration using OMOP CDM for evaluation coord-

Table 2. QA analysis results of patient, provider, visit, device identifier, and concept codes at Mayo Clinic and Mercy

Use	Mayo source	Mayo OMOP	Mercy source	Mercy OMOP
Distinct patient count	20 757	20 757	8449	8449
Present qualifying ablation code count	14	14	13	13
Distinct patient count with qualifying ablation codes	20 669	20 669	8424	8424
Distinct patient count in procedure table	20 757	20 757	8446	8446
Distinct patient count in diagnosis table	20 757	20 757	8431	8431
Distinct VT patient count in diagnosis table	5198	5198	1616	1616
Distinct AF patient count in diagnosis table	15 774	15 774	6396	6396
Distinct patient count in drug exposure table	20 158	20 158	8414	8414
Distinct patient count in device exposure table	5973	5973	3857	3857
Distinct patient count in visit occurrence table	20 757	20 757	8433	8433
Distinct patient count in observation table	16 706	16 706	8420	8420
Distinct patient count in death table	161	161	870	870
Distinct provider count in provider table	46 971	46 971	33 318	33 318
Distinct visit count in device exposure table with study relevant device GTINs in Mayo/Mercy confirmed use list ^a	NA	NA	3761	3761
Distinct unique_device_id in device exposure table with study relevant device GTINs	13	13	7	7
Distinct patient count in device exposure table with study relevant device GTINs in Mayo/Mercy confirmed use list	4636	4636	3411	3411
Distinct visit count in device exposure table with GTIN matching device from J&J Reference List (includes Navistar) ^a	NA	NA	4070	4070
OMOP distinct concept_code count in concept_set_item table ^b	NA	19 656	NA	19 656
OMOP distinct concept_code vocabulary id count in concept_set_item table ^b	NA	19 740	NA	19 740
OMOP distinct concept_id count in concept_set_item table ^b	NA	15 305	NA	15 305
OMOP distinct concept_set_item_id count in concept_set_item table ^b	NA	13	NA	13
OMOP distinct procedure vocabulary-based concept_code count from concept_set_item table ^b	NA	7642	NA	7642
Distinct patient count in procedures table joined to concept_set_item table	20 728	20 728	8439	8439
Distinct patient count in procedures table joined to concept_set_item table with procedure-based vocabulary	20 728	20 728	8439	8439
Distinct patient count in diagnoses table joined to concept_set_item table	20 754	20 754	8430	8430
Distinct patient count in diagnoses table joined to concept_set_item table with diagnosis-based vocabulary	20 754	20 754	8430	8430

AF: atrial fibrillation; VT: ventricular tachycardia; GTIN: Global Trade Item Number; J&J: Johnson & Johnson.

^aThe device usage data from supply chain database could not be linked with the visit ID in EHR at Mayo Clinic. Thus, it shows NA here.

^bThe outcome concept list is not recorded in the source EHR database, so it is represented as not available (NA) here.

minated across multiple healthcare systems, and it demonstrates the opportunity to conduct medical device-related research across healthcare systems using distributed analytics. Adopting the OMOP CDM enabled distributed analytics without pooling the datasets from 2 healthcare systems into a centralized database, which also reduced efforts on data de-identification, data use agreements, and IRB approval for building a centralized database. Previous experience using the OMOP CDM by one informatics team members helped to reduce the learning curve on the OMOP CDM and its implementation. Establishment of an OMOP work group facilitated the effective coordination of informatics teams across the 2 sites. Our data integration process could be generalized for use by other institutions that record both device data from a supply chain system and clinical data from an EHR system. The institution could build their own OMOP database instance locally and join together in a standard analysis with other institutions having their own OMOP instances as we (Mercy and Mayo) did. A challenge for our approach is deciding how to map the local data to the standard concepts recommended by OMOP CDM and ensure they are aligned among the sites. Our solution was to develop shared mappings to align our concepts. In addition, although we only used ThermoCool catheters as a use case,

we believe that the data ETL process could be reused for other medical devices with UDI as the data collection criteria.

FUNDING

This work was supported by the Food and Drug Administration grant 1U01FD006292-01. Views expressed in written materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does any mention of trade names, commercial practices, or organization imply endorsement by the United States Government.

AUTHORS CONTRIBUTIONS

YY, GJ, EB, TF, JSR, PMC, and JPD contributed to the study design. YY, EB, and TF contributed to design the mapping between the original databases and the OMOP CDM. YY, EB, and TF implemented the ETL process at Mercy and Mayo. YY, EB and TF contributed to conduct the QA process. SS, SZ, JC, AAD, KC, and PAN designed the code lists for patient collection, phenotypes, covariates, and outcome endpoints. YY, GJ, EB, TF, SS, SZ, JSR, PMC, and JPD contributed to draft the manuscript. All the authors contributed to the manuscript review and editing.

ACKNOWLEDGMENTS

This project was supported by a research grant from the Medical Device Innovation Consortium (MDIC) as part of the National Evaluation System for Health Technology (NEST), an initiative funded by the U.S. Food and Drug Administration (FDA) Center for Device and Radiologic Health. Its contents are solely the responsibility of the authors and do not necessarily represent the official views nor the endorsements of the Department of Health and Human Services or the FDA. While MDIC provided feedback on project conception and design, the organization played no role in collection, management, analysis, and interpretation of the data. The research team, not the funder, made the decision to submit the manuscript for publication.

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY

The data underlying this article cannot be shared publicly due to the IRB requirement of the study.

REFERENCES

- Krucoff MW, Sedrakyan A, Normand SL. Bridging nmet medical device ecosystem needs with strategically coordinated registries networks. *JAMA* 2015; 314 (16): 1691–2.
- Center for Devices and Radiological Health, US Food and Drug Administration. Strengthening our national system for medical device postmarket surveillance. 2013. <http://www.fda.gov/downloads/AboutFDA/Center-Offices/OfficeofMedicalProductsandTobacco/CDRH/CDRHReports/UCM301924.pdf>. Accessed July 29, 2022.
- US Food and Drug Administration. 2018-2020 Strategic Priorities. Center for Devices and Radiological Health. 2018. <https://www.fda.gov/media/110478/download>. Accessed July 29, 2022.
- Drozda JP Jr, Roach J, Forsyth T, et al. Constructing the informatics and information technology foundations of a medical device evaluation system: a report from the FDA unique device identifier demonstration. *J Am Med Inform Assoc* 2018; 25 (2): 111–20.
- Dhalluin T, Fakhiri S, Bouzillé G, et al. Role of real-world digital data for orthopedic implant automated surveillance: a systematic review. *Expert Rev Med Devices* 2021; 18 (8): 799–810.
- Tcheng JE, Crowley J, Tomes M, et al. MDEpiNet UDI Demonstration Expert Workgroup. Unique device identifiers for coronary stent postmarket surveillance and research: a report from the Food and Drug Administration Medical Device Epidemiology Network Unique Device Identifier demonstration. *Am Heart J* 2014; 168 (4): 405–413.e2.
- FDA Sentinel Common Data Model. 2021. <https://www.sentinelinitiative.org/methods-data-tools/sentinel-common-data-model>. Accessed July 29, 2022.
- OMOP Common Data Model. 2021. <https://ohdsi.github.io/CommonDataModel/>. Accessed July 29, 2022.
- PCORnet Common Data Model Specification. 2021. https://pcorner.org/wp-content/uploads/2020/12/PCORnet-Common-Data-Model-v60-2020_10_221.pdf. Accessed July 29, 2022.
- I2B2 Common Data Model Documentation. 2021. <https://community.i2b2.org/wiki/display/BUN/i2b2+Common+Data+Model+Documentation>. Accessed July 29, 2022.
- Visweswaran S, Becich MJ, D'Itri VS, et al. Accrual to clinical trials (ACT): A Clinical and Translational Science Award Consortium Network. *JAMIA Open* 2018; 1 (2): 147–52.
- Jiang G, Yu Y, Kingsbury PR, Shah N. Augmenting medical device evaluation using a reusable unique device identifier interoperability solution based on the OHDSI Common Data Model. *Stud Health Technol Inform* 2019; 264: 1502–3.
- Choi S, Choi SJ, Kim JK, et al. Preliminary feasibility assessment of CDM-based active surveillance using current status of medical device data in medical records and OMOP-CDM. *Sci Rep* 2021; 11 (1): 24070.
- Jiang G, Dhruva SS, Chen J, et al. Feasibility of capturing real-world data from health information technology systems at multiple centers to assess cardiac ablation device outcomes: a fit-for-purpose informatics analysis report. *J Am Med Inform Assoc* 2021; 28 (10): 2241–50.
- Dhruva SS, Jiang G, Doshi AA, et al. Feasibility of using real-world data in the evaluation of cardiac ablation catheters: a test-case of the National Evaluation System for Health Technology Coordinating Center. *BMJ Surg Interv Health Technol* 2021; 3 (1): e000089.
- ATLAS. 2021. <https://atlas.ohdsi.org/>. Accessed July 29, 2022.
- https://accessgudid.nlm.nih.gov/resources/developers/device_snomed_api. Accessed July 29, 2022.
- GUDID Device Lookup API. <https://nestcc.org/portfolio-item/the-feasibility-of-using-real-world-data-in-the-evaluation-of-cardiac-ablation-catheters/>. Accessed July 29, 2022.
- Dhruva SS, Zhang S, Chen J, et al. Safety and effectiveness of a catheter with contact force and 6-hole irrigation for ablation of persistent atrial fibrillation in routine clinical practice. *JAMA Netw Open* 2022; 5 (8): e2227134.