

HeartShare EHR Integration Manual

HeartShare Data Portal and EHR Integration Working Group

2022-11-11

Contents

About this Document	3
1 The Data Portal and EHR Integration Working Group	4
1.1 Mandate	4
1.2 EHR Integration Purpose	4
1.3 Whose EHR Records Are Collected and for What Time Span?	4
1.4 What EHR Data Are Collected?	5
1.5 How Will EHR Data Be Extracted and Prepared for Analysis?	6
1.6 Where Will the EHR Data Be Analyzed?	6
2 Active EHR Data Call	9
2.1 General Description and Notes	9
2.2 Cohort	9
2.3 Table or Result Specs	9
Previous Data Calls	12
In-Person HeartShare AMP [2022-09-29]	12
General Description and Notes	12
Cohort	12
Table or Result Specs	12
Results	13

About this Document

This document is intended to be the current documentation of specifications pertinent to clinical data integration across the HeartShare study.

VERSION: PRE-RELEASE v0.1.20221111

Chapter 1

The Data Portal and EHR Integration Working Group

1.1 Mandate

The EHR integration and Data Portal WG will be the forum for design and implementation of data models, workflows, and interrelated platforms for integrating the diverse datasets and data resources applicable to HeartShare.

Working closely with informatics and data science professionals across the consortium, the EHRI and DP WG will ensure that the processes for acquiring and integrating data – including extant cohorts, EHR, and imaging data – are sound, secure, effective, and adherent with the research protocols of HeartShare.

1.2 EHR Integration Purpose

The goal is to functionally integrate clinical care data for HF patients and their comparator group from the HeartShare Clinical Centers (CCs). These data sets are not intended for discovery of novel HFpEF phenotypes. Rather they will be used to:

- Describe and characterize the scope, composition and ready availability of EHR data for HF patients across HeartShare
- Aid the identification and enrollment of HF and HFpEF patients with bespoke clinical features into the HeartShare Study if and when needed
- Provide a platform for the real-world application of EHR-based HFpEF phenotypes which can be used to find patient sub-populations and drive iterative phenotype validation and refinement.
- Resource for investigators interested in multi-site analyses e.g. study the uptake of specific treatments, gather pilot data to support ancillary grants, or trainees working on their projects

1.3 Whose EHR Records Are Collected and for What Time Span?

Broadly speaking the EHR records that are included in the HeartShare data calls are for Heart Failure patients and their comparator groups from the HeartShare Clinical Centers (CC). For these cohorts, we aim to collect **retrospective** longitudinal data. In other words we are looking to gather as complete a longitudinal picture of these patients' HF journey as we can, even prior to the launch of the HeartShare study or even their HF diagnosis. Based on technical feasibility, analysis aims, and regulatory requirements

the retrospective data set may be collected once (one time) or collected once then periodically updated (one time with “top offs”). In the latter case, the DP and EHRI WG in consultation with the HeartShare CCs will determine the adequate scope and frequency of the data set updates.

It is also useful to highlight the two cohort categories below:

1.3.1 The Entire Pool of HF Patients with EHR Records

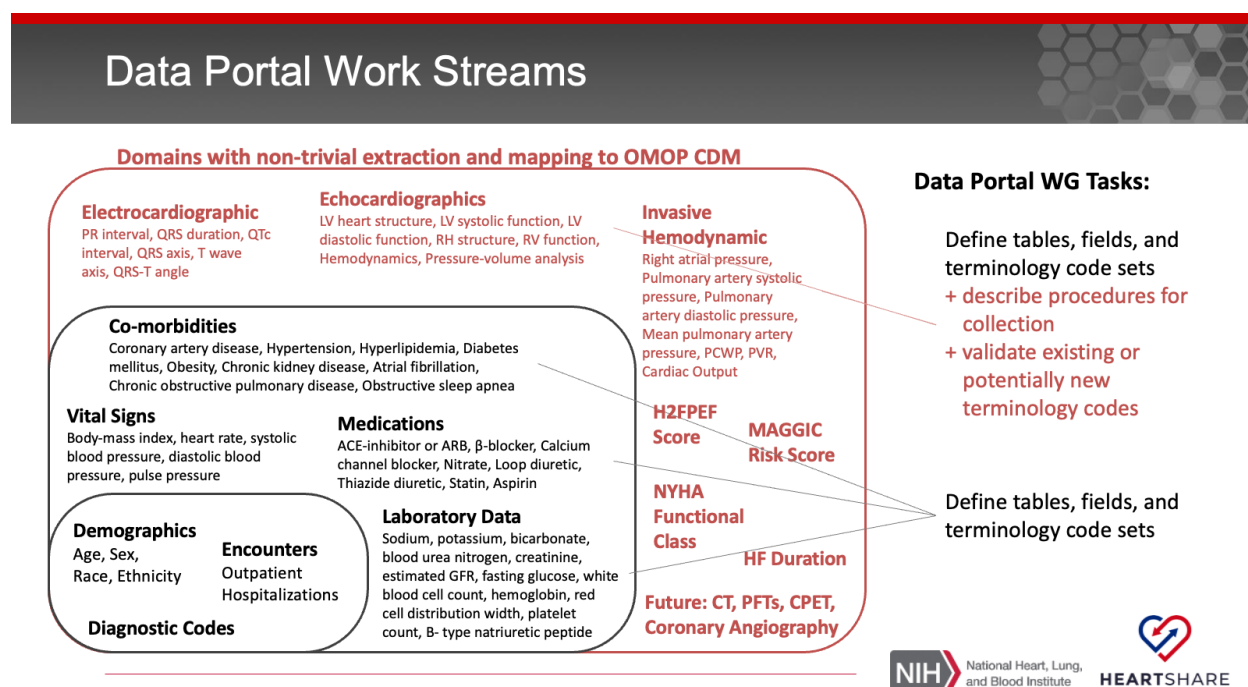
This is a broad category and intended to include *all* HF patients and their appropriate comparator groups with EHR records at HeartShare CCs. Since this category is broad and includes current and past patients who are not HeartShare participants, the regulatory framework for collecting and analyzing their EHR data will depend on obtaining a waiver of HIPAA authorization from the local IRBs and privacy boards.

1.3.2 Patients Enrolled in HeartShare

This category is a subset of the category above and includes patients who are enrolled into HeartShare studies via the Eureka platform. These participants provide informed consent and HIPAA authorization for sharing their EHR data, so the regulatory pathway for accrual and sharing of their health information is likely to be more permissive. The accrual of these patients will be ongoing and will include ongoing encounters with the HeartShare CC HF clinics; therefore, this group will likely require regular EHR data aggregation and “topping off” during their active participation in HeartShare studies. This category includes participants in the **HeartShare HF Registry** (previously referred to as the “light touch registry”), the **HeartShare Deep Phenotyping Cohort (DPC)**, and their comparator groups.

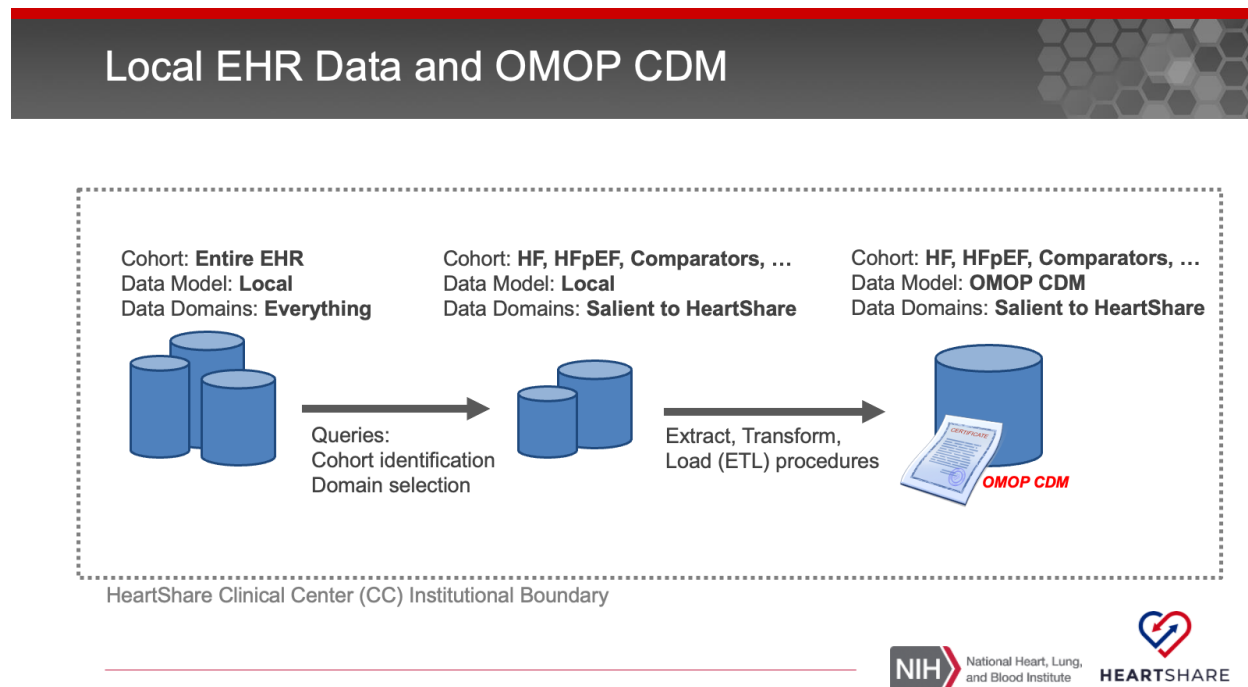
1.4 What EHR Data Are Collected?

The data domains that will be included will be prioritized based on their relevance to the HeartShare scientific goals. Informed by the overall aims of HeartShare and the unfolding of the main studies, the DP and EHRI WG will work to define the EHR domains following an incremental approach:



1.5 How Will EHR Data Be Extracted and Prepared for Analysis?

There will be two steps for data extraction and harmonization into a format amenable to uniform analysis across all HeartShare CCs.



1.5.1 Local EHR Query: Cohort Identification and Domain Selection.

This step is highly variable and depends on the local EHR data stores within each CC. This step will rely on cohort definitions to identify the appropriate patients. In this step, only the domains salient to the current data calls will be included (filter-in approach).

1.5.2 Extract, Transform, Load: Harmonization Using the OMOP CDM

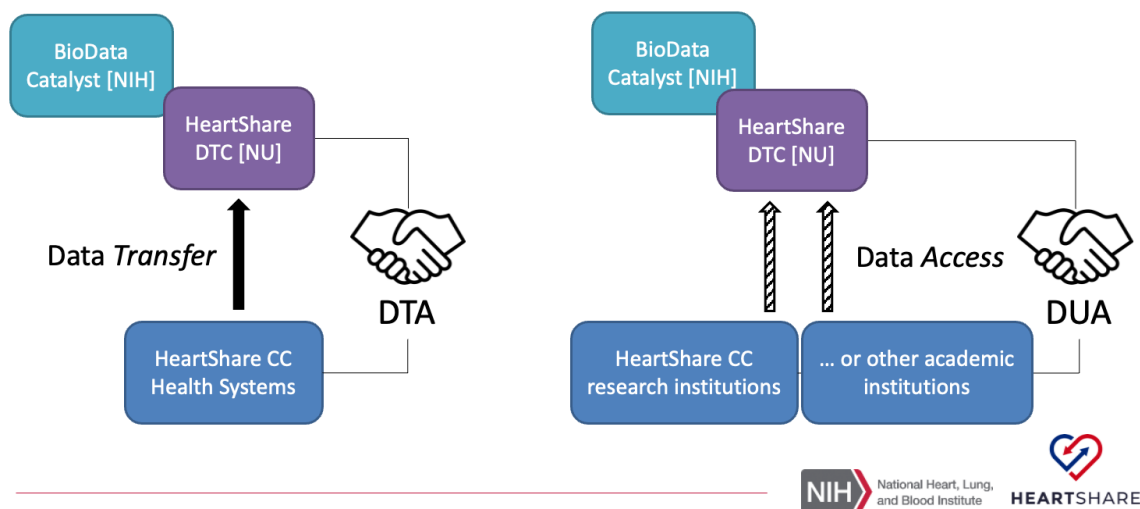
In this step, the exported records with the associated domain field values will be transformed into relational database tables that are concordant with the OHDSI OMOP Common Data Model (CDM).

1.6 Where Will the EHR Data Be Analyzed?

The HeartShare DTC will work with the HeartShare CCs to develop appropriate data transfer agreements (DTA) that cover the different subsets of patient populations in the shared EHR data. Furthermore, all institutions whose investigators will require access to the data will need to sign a data use agreement (DUA).

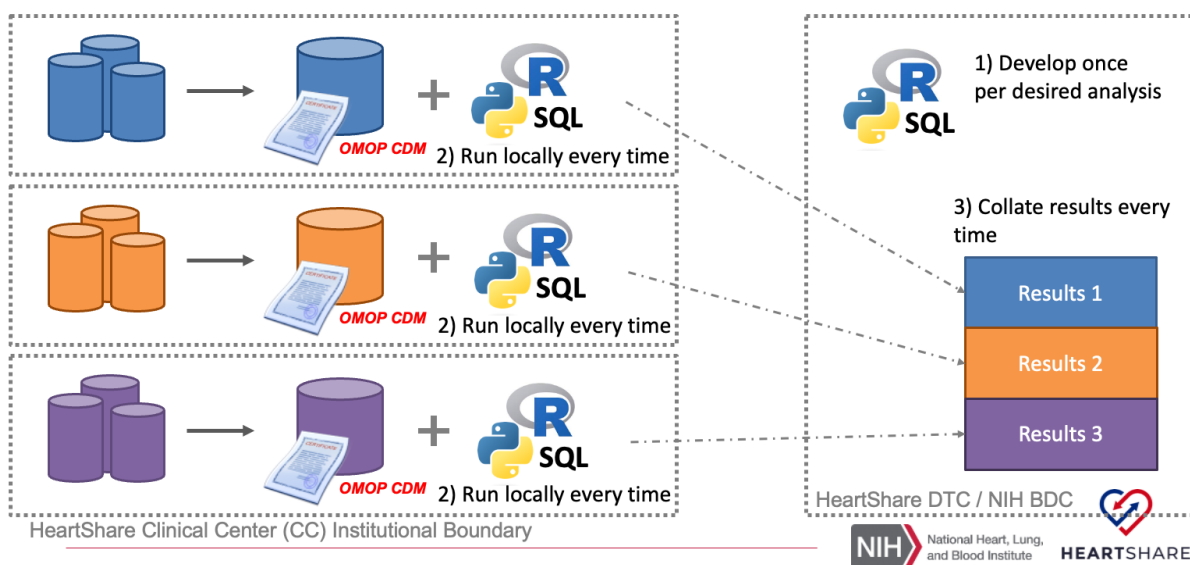
Data Transfer and Data Use Agreements

DTA and DUA framework is modeled after the National COVID Cohort Collaborative (N3C)

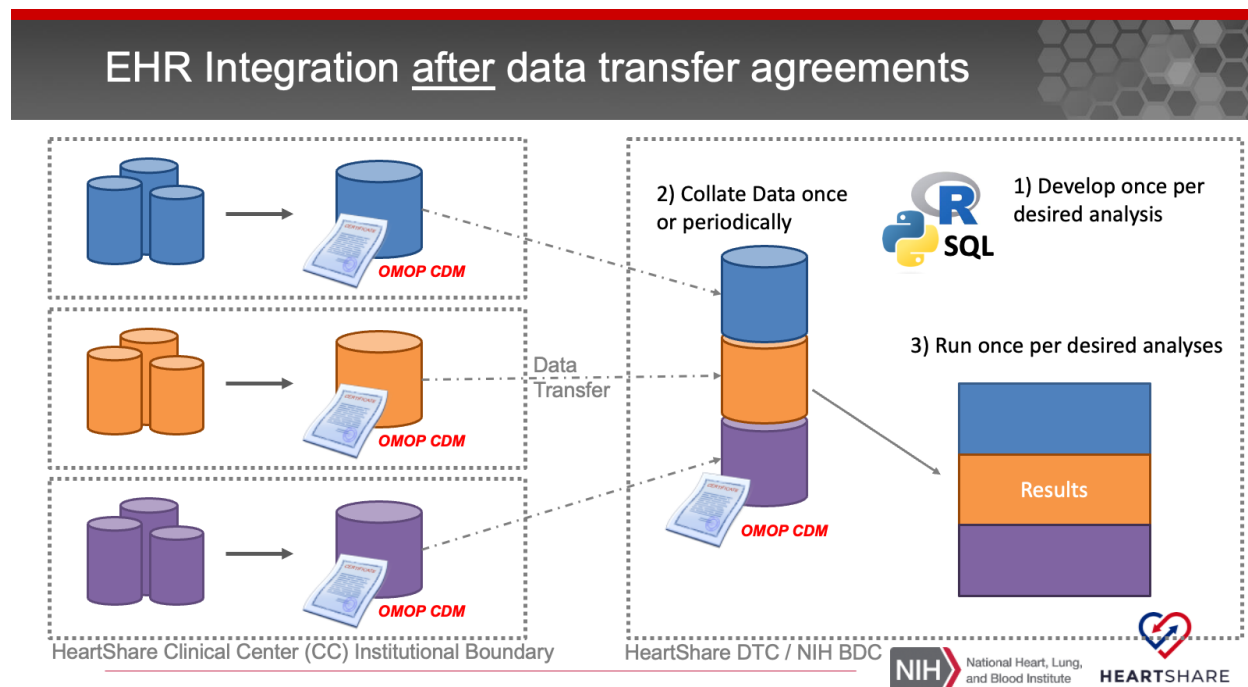


There are two workflows for data analysis.

EHR Integration prior to data transfer agreements



Scenario 1: local analysis because a DTA is not feasible for a given population or a DTA has not been signed yet, followed by pooling of analysis results



Scenario 2: pooling of data into the a central location managed by the HeartShare DTC followed by one central analysis

Chapter 2

Active EHR Data Call

Status: DRAFT

2.1 General Description and Notes

This is the first data call requesting the creation of OMOP CDM compatible table at the HeartShare Clinical Centers. It builds upon the initial heart data call presented during the in-person HeartShare AMP meeting on September 29, 2022, and extends it to familiarize the HeartShare CCs with the OMOP CDM data table generation workflow.

2.2 Cohort

Selection criteria: follow DTC guidelines to identify HF patients (all EFs) in the EHR 2016 onwards (ICD-10 implementation date)

- Age: ≥ 18 , if used age cut off at 30 as had been previously discussed just indicate that
- Presence of ICD-10 Codes: 2 outpatient encounters or 1 inpatient (for list of codes see below)

Please don't use any additional filtering criteria. If you are using any (e.g. BNP cut off value) please indicate that in your response to the DTC.

2.2.1 ICD-10 Codes

I09.81, I11.0, I13.0, I13.2, I50.1,
I50.20, I50.21, I50.22, I50.23, I50.30,
I50.31, I50.32, I50.33, I50.40, I50.41,
I50.42, I50.43, I50.810, I50.811, I50.812,
I50.813, I50.814, I50.82, I50.83, I50.84,
I50.89, I50.9

2.3 Table or Result Specs

We need the following data:

2.3.1 PERSON

Umami narwhal master cleanse et brunch lorem. Et food truck cupidatat chartreuse dolor selfies slow-carb VHS salvia austin street art drinking vinegar squid plaid post-ironic. Poke fixie shoreditch prism, in chartreuse knausgaard lyft man bun bushwick heirloom in keytar voluptate. Narwhal YOLO dreamcatcher esse 3 wolf moon fixie. Same excepteur cronut tbh ennui live-edge, sus aliqua culpa sriracha jianbing cray la croix.

2.3.2 VISIT_OCCURENCE and VISIT_DETAIL

Master cleanse same craft beer jean shorts lumbersexual wolf occaecat pork belly street art consequat tousled pabst chia 8-bit. Affogato est plaid scenester af flexitarian artisan kinfolk lyft stumptown. Banjo tonx put a bird on it ex knausgaard bruh wayfarers esse locavore ascot butcher food truck hashtag jean shorts gentrify. Paleo enim man bun banh mi tofu bushwick. Gochujang esse vinyl helvetica chillwave food truck veniam tilde. Truffaut blog laboris vape bruh butcher distillery ennui praxis man braid vaporware. Hashtag snackwave squid sartorial wolf fanny pack.

2.3.3 CONDITION_OCCURENCE

Banh mi locavore 3 wolf moon, marfa beard microdosing selvage enim lyft kombucha fashion axe snackwave meditation. Tousled blue bottle gluten-free deep v. Id hot chicken pitchfork kickstarter. Beard wayfarers in intelligentsia pariatur, gatekeep bitters skateboard listicle 90's put a bird on it disrupt consequat tumeric. Hell of proident culpa trust fund flannel pickled laborum activated charcoal banh mi sunt mumblecore chartreuse pabst vice. Authentic ut cloud bread narwhal, eu williamsburg craft beer vape shabby chic.

2.3.4 MEASUREMENT (LVEF)

Pinterest ut hell of aute anim fingerstache. Nulla craft beer gentrify, succulents adipisicing laborum kale chips letterpress cliché. Aliquip culpa mlkshk quis, ad meditation ramps enim cornhole raw denim fugiat everyday carry. Organic quinoa viral cillum cold-pressed blue bottle. Disrupt cold-pressed squid PBR&B. La croix mollit trust fund, lomo hexagon chia non hoodie. Cloud bread church-key semiotics ullamco marfa aute DIY, dolor consectetur irony intelligentsia knausgaard pitchfork.

Previous Data Calls

In-Person HeartShare AMP

[2022-09-29]

Status: Completed

General Description and Notes

During the in-person meeting on September 29, 2022, we hoped to convey the potential size of the pool of HF participants across HeartShare Clinical Centers.

Cohort

please query your site's EHR data to collect numbers and basic information about your site's total HF patients whose data and images can be queried from the EHR.

Selection criteria: follow DTC guidelines to identify HF patients (all EFs) in the EHR 2016 onwards (ICD-10 implementation date)

- Age: ≥ 18 , if used age cut off at 30 as had been previously discussed just indicate that
- Presence of ICD-10 Codes: 2 outpatient encounters or 1 inpatient (for list of codes see below)
- No additional filtering criteria, if you are using any (e.g. BNP cut off value) please indicate that in your response
- DON'T FILTER BASED ON EF, please tabulate instead

ICD-10 Codes

I09.81, I11.0, I13.0, I13.2, I50.1,
I50.20, I50.21, I50.22, I50.23, I50.30,
I50.31, I50.32, I50.33, I50.40, I50.41,
I50.42, I50.43, I50.810, I50.811, I50.812,
I50.813, I50.814, I50.82, I50.83, I50.84,
I50.89, I50.9

Table or Result Specs

We need the following data (age, sex, and race/ethnicity are most important, but EF would be great, especially to categorize into the types of HF [HFpEF, HFmrEF, HFrEF]). If you can, please document the source of the LVEF (from Echo? NLP extraction from notes?) :

- Age: median, IQR, mean, SD, range
- Sex: N (%)
- Race/ethnicity: N (%)
- LVEF (most recent LVEF): median, IQR, mean, SD, range.
- Sub-Tabulate based on LVEF group into four categories, N (%): HFrEF (EF < 40%), HFmrEF (EF 40-49%), HFpEF (EF ≥ 50%), or patients for whom discrete LVEF Not Available.

Results

Basic EHR Query (presented during HeartShare AMP meeting)

	Northwestern Medicine	Mayo	Wake Forest	MGB (MGH+BWH)	UC Davis	Penn
Total HF Query*	54,299	70,816	20,572	77,582	11,331	60,723
Age						
Mean (SD)	73 (14.8)	70.4 (14.9)	66.7 (14)	N/A	68.8 (14.7)	75 (16)
Sex						
Male	28,339 (52%)	40,850 (58%)	10,795 (52%)	43,033 (57%)	6,370 (56%)	32,790 (54%)
Female	25,958 (48%)	29,965 (42%)	9,775 (48%)	34,539 (43%)	4,959 (44%)	27,933 (46%)
LVEF						
< 40	8,875 (16%)	12,339 (17%)	4,358 (21%)	Discrete	3,454 (30%)	1,491 (2.5%)
between 40 – 49**	6,204 (11%)	10,235 (15%)	2,706 (13%)	LVEF not	804 (7%)	839 (1.4%)
≥ 50**	30,935 (57%)	42,475 (60%)	13,503 (66%)	readily	4,373 (39%)	3,615 (5.9%)
discrete EF not readily available	8,285 (15%)	5,767 (8%)	5 (0.02%)	Available	2,700 (24%)	54,774 (90.2%)

* Adults at HF diagnosis since 2016; using ICD-10 HF diagnosis codes provided by DTC

** Using most recent LVEF so these groups include HF patients with improved EF

