HeartShare EHR Integration Manual

HeartShare Data Portal and EHR Integration Working Group

2024-04-18

Contents

About this Document

This document is intended to be the current documentation of specifications pertinent to clinical data integration across the HeartShare study. The most recent version will be available on github at (github.com/HeartShareStudy/PHENOTYPES). The best way to pose questions and suggestions about the format of this document, its content, or the process surrounding its maintenance, please start an issue on: (github.com/HeartShareStudy/PHENOTYPES/issues). This will start a threaded public conversation in which others can participate and it will also note when the issue is resolved in future updates.

CURRENT VERSION: NEW RELEASE v1.0.20240418.

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)

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

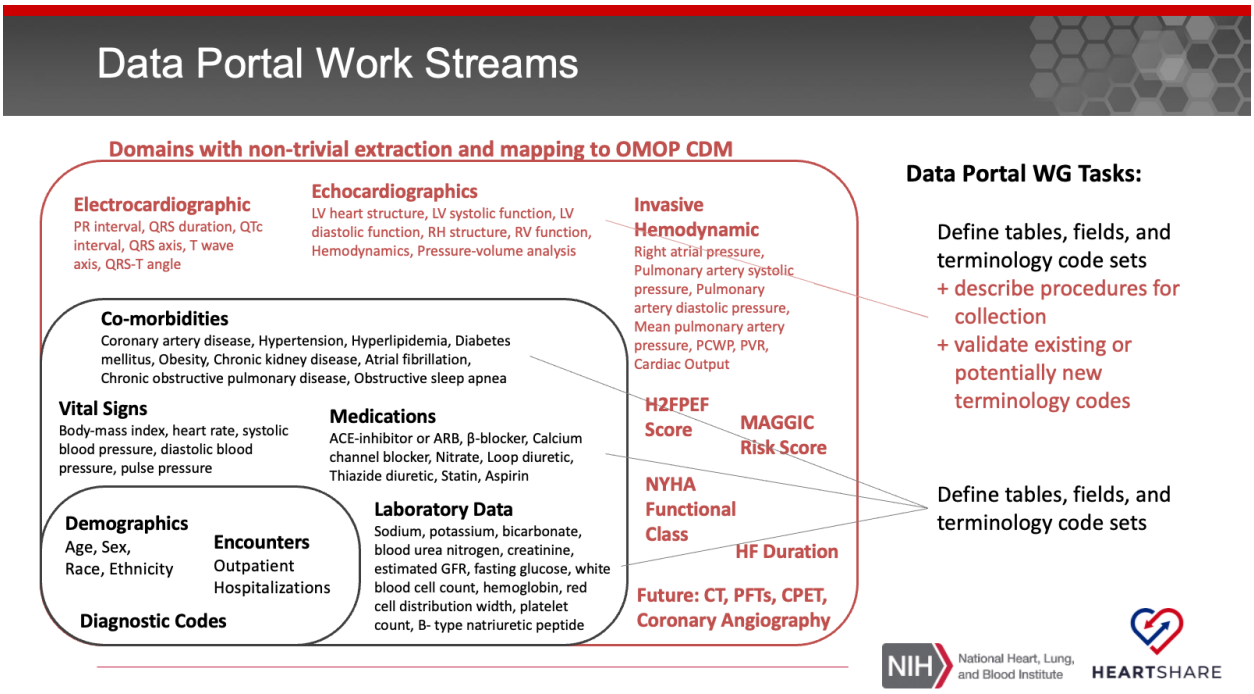
This is a broad category and intended to include *all* HF patients and their appropriate comparator groups with EHR records at HeartShare CCs. Since this category is broad and includes current and past patients who are not HeartShare participants, the regulatory framework for collecting and analyzing 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**.

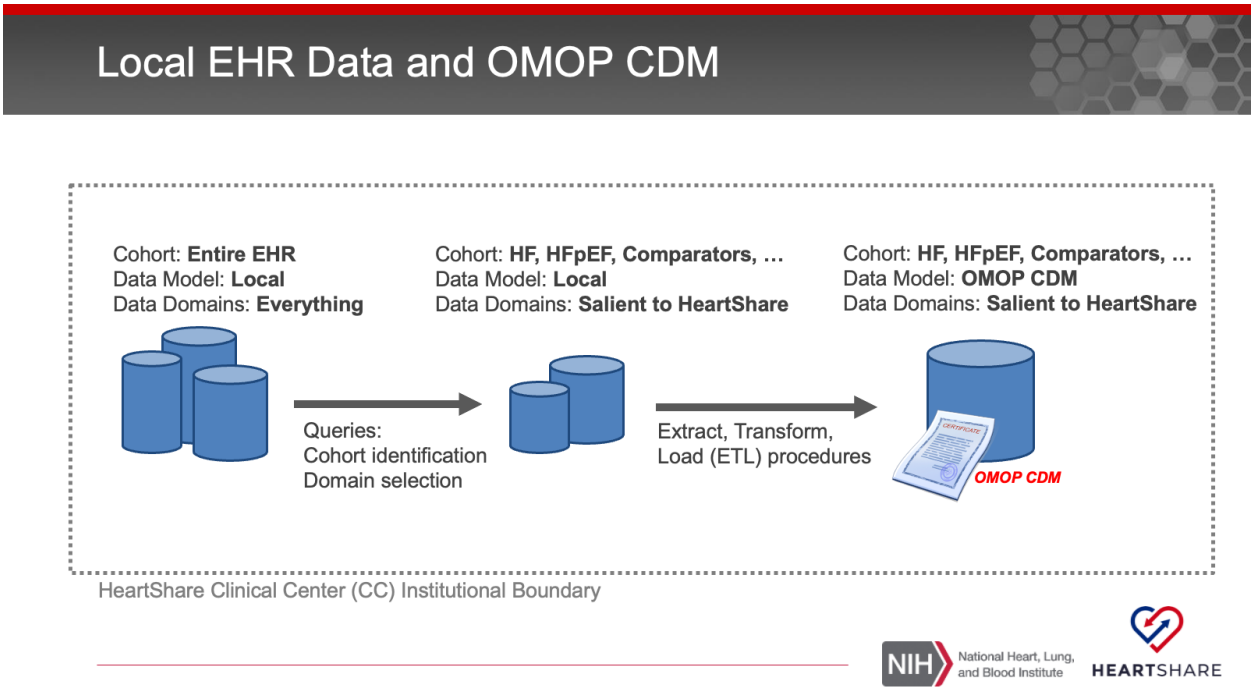
1.4 What EHR Data Are in Scope?

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. The following diagram shows a high-level roadmap of the domains we will be seeking. Note that there will be domains (red) for which there may not be a clear pathway for mapping to the OMOP CDM. In addition to specifying the requirements, the DP and EHRI WG will set guidelines on the procedures for obtaining data from ancillary data sources where they reside. Furthermore, there may be a need to examine the pertinent vocabularies and propose additions where needed. (e.g., by working through the OHDSI community's CDM Refresh Process.) This will offer an opportunity for the work in HeartShare to inform EHR-based phenotyping in the field in general.



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.

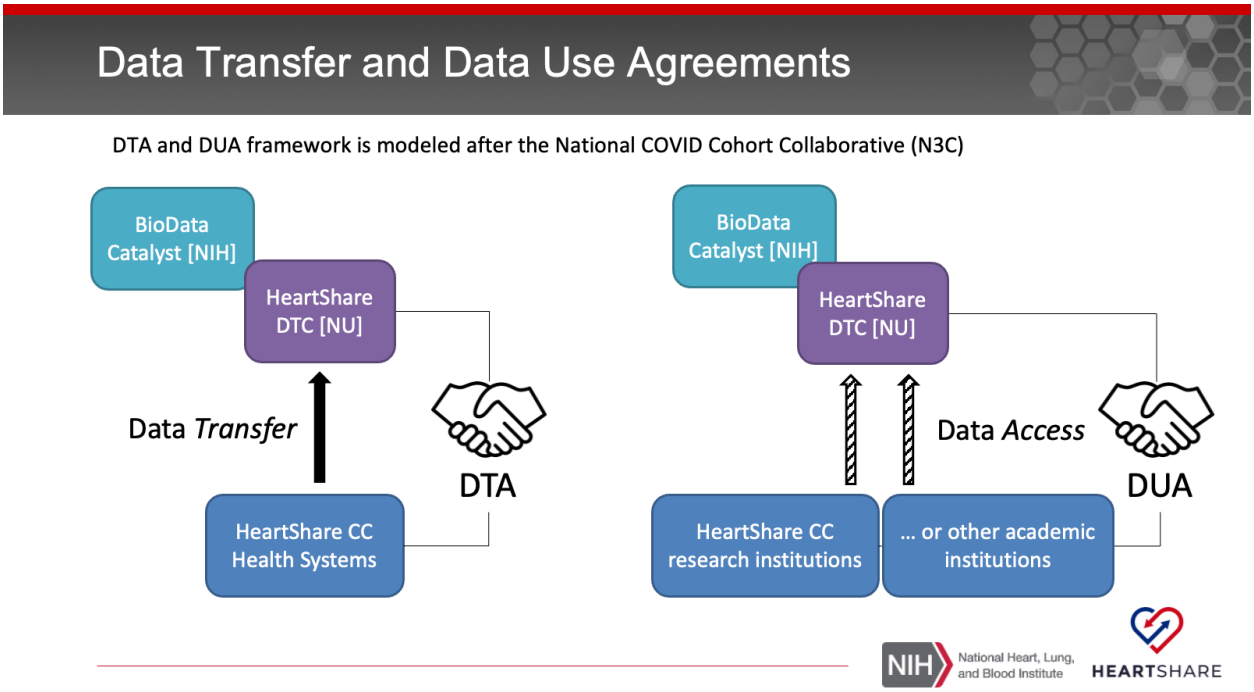
This step is highly variable and depends on the local EHR data stores within each CC. This step will rely on cohort definitions to identify the appropriate patients.

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

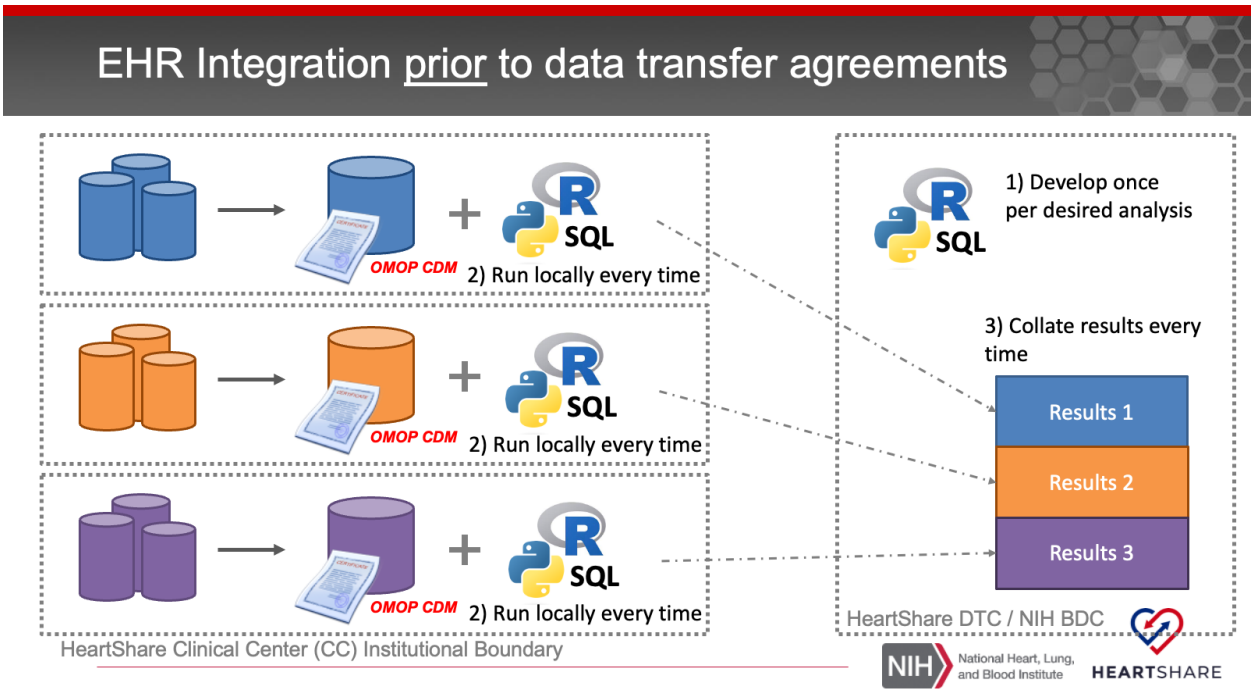
In this step, the exported records with the associated domain field values will be transformed into relational database tables that are concordant with the OHDSI OMOP Common Data Model (CDM).

1.6 Where Will the EHR Data Be Analyzed?

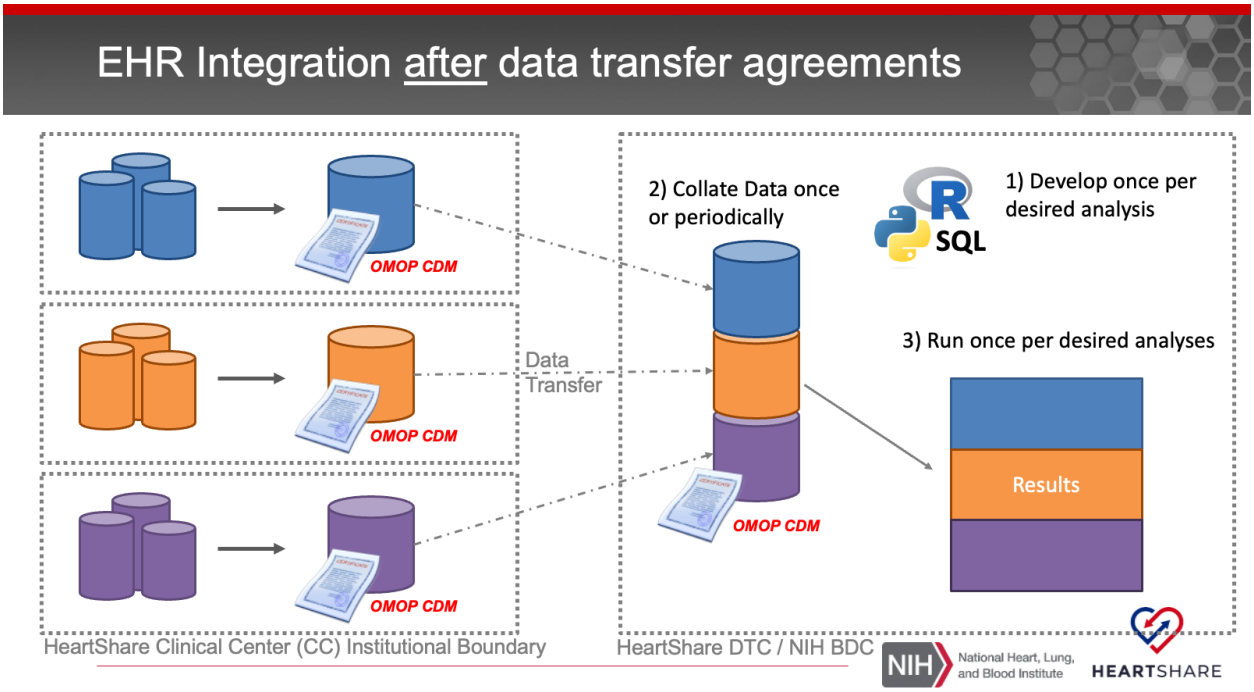
The HeartShare DTC will work with the HeartShare CCs to develop appropriate data transfer agreements (DTA) that cover the different subsets of patient populations in the shared EHR data. Furthermore, all institutions whose investigators will require access to the data will need to sign a data use agreement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



There are two workflows for data analysis.



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 a central location managed by the HeartShare DTC followed by one central analysis

Chapter 2

HeartShare Cohort Details

2.1 Cohort

As currently proposed, there will be **three cohorts** of HF patients whose records are to be included in the OMOP tables. This does not include their comparator groups yet as the cohort definition criteria for them is still being developed.

2.1.1 Historical HF Patients at CC

2.1.1.1 Selection criteria

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

- Age: $\geq$ 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 do not 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.

2.1.1.2 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
```

2.1.2 HF Patients Prospectively Enrolled in the HeartShare Study

Patients in the HF arms of the HeartShare Study, both the HeartShare HF Registry and the HFpEF Deep Phenotyping Cohorts, are also to be included as they enroll in the study and provide informed consent and EHR authorization via the Eureka app.

2.2 OMOP Tables Usage

All the tables are congruent with **OMOP CDM v.5.3.1**. Required tables are expected to be populated with the data and are listed in **bold letters**. Optional tables can be left empty for now are listed in gray. A link to the corresponding table definitions in the OMOP CDM is provided for each table.

If possible, we ask that patient data for the tables are populated starting 10 years prior to the date patient qualified for the cohort (selection criteria looks in the EHR records starting with year 2016; If patient qualified in 2020, cohort should include patients' data from 2010 onwards).

2.2.1 Clinical Data Tables

https://ohdsi.github.io/CommonDataModel/cdm53.html#Clinical_Data_Tables

- **PERSON**
  - We'll be collecting patient's 5-digit address zip codes only. Patient's most recent address should be stored in PERSON table, in *location_id* field. More information is provided later, under LOCATION table
- **OBSERVATION_PERIOD**
- **VISIT_OCCURENCE**
- VISIT_DETAIL
- **CONDITION_OCCURRENCE**
- **DRUG_EXPOSURE**
  - If possible, we recommend loading all of the medication data. For sites that are building de novo OMOP tables, selective loading should focus on cardiovascular drugs that fall under the following ATC classes: alimentary tract and metabolism, blood and blood forming organs, and cardiovascular systems. Since the list of rxnorm codes is a long one, provided below is sql code, to run against OMOP vocabulary tables. The code returns rxnorm codes that belong to ATC (sub)class specified in the variable `@ATC_code` (in the example, all rxnorm codes for ATC class **A10** - Drugs used in Diabetes will be returned)

```
-- Returns RxNorm codes for selected ATC class
DECLARE @ATC_code AS varchar(10);
SET  @ATC_code = 'A10'

;with t as (
        select concept_code, concept_name
            from [cdm].[concept]
            where vocabulary_id in ('ATC') and concept_code like @ATC_code+'%'
       )

, atc as   (
        select c.concept_code,c.concept_id,
            left(c.concept_code,case when len(c.concept_code)>=1 then 1 end)
            [ATC 1st_code]
            ,t1.concept_name [ATC 1st_name],
            left(c.concept_code,case when len(c.concept_code)>=3 then 3 end)
            [ATC 2st_code]
```

```
            ,t2.concept_name [ATC 2st_name],
            left(c.concept_code,case when len(c.concept_code)>=4 then 4 end)
            [ATC 3st_code]
            ,t3.concept_name [ATC 3st_name],
            left(c.concept_code,case when len(c.concept_code)>=5 then 5 end)
            [ATC 4st_code]
            ,t4.concept_name [ATC 4st_name],
            left(c.concept_code,case when len(c.concept_code)>=7 then 7 end)
            [ATC 5st_code]
            ,t5.concept_name [ATC 5st_name]
            from [cdm].[concept] c
            left join t as t1 on t1.concept_code= left(c.concept_code
            ,case when len(c.concept_code)>=1 then 1 end)
            left join t as t2 on t2.concept_code=left(c.concept_code
            ,case when len(c.concept_code)>=3 then 3 end)
            left join t as t3 on t3.concept_code=left(c.concept_code
            ,case when len(c.concept_code)>=4 then 4 end)
            left join t as t4 on t4.concept_code=left(c.concept_code
            ,case when len(c.concept_code)>=5 then 5 end)
            left join t as t5 on t5.concept_code=left(c.concept_code
            ,case when len(c.concept_code)>=7 then 7 end)
            where
                 c.vocabulary_id in ('ATC')
                 and t5.concept_name is not null
                 and c.concept_code like @ATC_code+'%'
          )
select [ATC 1st_code]+isnull('-'+[ATC 1st_name],'') ATC_1,
            [ATC 2st_code]+isnull('-'+[ATC 2st_name],'') ATC_2,
            [ATC 3st_code]+isnull('-'+[ATC 3st_name],'') ATC_3,
            [ATC 4st_code]+isnull('-'+[ATC 4st_name],'') ATC_4,
            [ATC 5st_code]+isnull('-'+[ATC 5st_name],'') ATC_5,
       c.concept_code RxNorm_code,c.concept_class_id
            ,c.standard_concept, c.concept_name RxNorm_desc
from atc d
join [heartshare_stg].[cdm].[concept_ancestor] a
 on  a.ancestor_concept_id=d.concept_id
join [cdm].[concept] c on c.concept_id=a.[descendant_concept_id]
where c.vocabulary_id='rxnorm'
order by d.concept_code, a.min_levels_of_separation
```

- **PROCEDURE_OCCURRENCE**
  - If possible, we recommend loading all procedures. In particular, we would like sites to confirm the loading of cardiovascular procedures and diagnostic radiology procedures of the heart.

Procedure Group	CPT4 Codes
Cardiac Catheterization Procedures	93451, 93452, 93453, 93454, 93455, 93456, 93457, 93458, 93459, 93460, 93461, 93462, 93463, 93464, 93503, 93505, 93563, 93564, 93565, 93566, 93567, 93568, 93571, 93572, 93580, 93581, 93582, 93583, 93590, 93591, 93592, 93593, 93594, 93595, 93596, 93597, 93598
Cardiography Procedures	93000, 93005, 93010, 93015, 93016, 93017, 93018, 93024, 93025, 93040, 93041, 93042, 93050

Procedure Group	CPT4 Codes
Cardiovascular Monitoring Services	93224, 93225, 93226, 93227, 93229, 93241, 93242, 93243, 93244, 93245, 93246, 93247, 93248, 93268, 93270, 93271, 93272, 93278
Echocardiography Procedures	93303, 93304, 93306, 93307, 93308, 93312, 93313, 93314, 93315, 93316, 93317, 93318, 93319, 93320, 93321, 93325, 93350, 93351, 93352, 93355, 93356
Home and Outpatient International Normalized Ratio	93792, 93793
Implantable, Insertable, and Wearable Cardiac Dev	93260, 93261, 93279, 93280, 93281, 93282, 93283, 93284, 93285, 93286, 93287, 93288, 93289, 93290, 93291, 93292, 93293, 93294, 93295, 93296, 93297, 93298, 93299
Intracardiac Electrophysiological Procedures	93600, 93602, 93603, 93609, 93610, 93612, 93613, 93615, 93616, 93618, 93619, 93620, 93621, 93622, 93623, 93624, 93631, 93640, 93641, 93642, 93644, 93650, 93653, 93654, 93655, 93656, 93657, 93660, 93662
Noninvasive Physiologic Studies and Procedures	93701, 93702, 93724, 93740, 93745, 93750, 93770, 93784, 93784, 93786, 93786, 93788, 93788, 93790, 93790
Other Cardiovascular Procedures	93797, 93798, 93799
Peripheral Arterial Disease Rehabilitation	93668
Therapeutic Cardiovascular Services and Procedures	92920, 92921, 92924, 92925, 92928, 92929, 92933, 92934, 92937, 92938, 92941, 92943, 92944, 92950, 92953, 92960, 92961, 92970, 92971, 92973, 92974, 92975, 92977, 92978, 92978, 92979, 92979, 92986, 92987, 92990, 92997, 92998, 1012989
Diagnostic Radiology Procedures of the Heart	75557, 75559, 75561, 75563, 75565, 75571, 75572, 75573, 75574

- DEVICE_EXPOSURE
- **MEASUREMENT**
  - If possible, we recommend loading all of measurement data, including vitals. High value labs and vitals are listed in the table.

Lab/Vital	Description	LOINC code
Sodium	Sodium [Moles/volume] in Serum or Plasma	2951-2
	Sodium [Moles/volume] in Blood	2947-0
	Sodium [Moles/volume] in Arterial blood	32717-1
	Sodium [Moles/volume] in Capillary blood	39792-7
	Sodium [Moles/volume] in Serum, Plasma or Blood	77139-4
	Sodium [Moles/volume] in Mixed venous blood	41657-8
	Sodium [Moles/volume] in Venous blood	39791-9
Potassium	Potassium [Moles/volume] in Serum or Plasma	2823-3
	Potassium [Moles/volume] in Blood	6298-4
	Potassium [Moles/volume] in Arterial blood	32713-0
	Potassium [Moles/volume] in Capillary blood	39790-1
	Potassium [Moles/volume] in Mixed venous blood	41656-0
	Potassium [Moles/volume] in Serum, Plasma or Blood	77142-8
	Potassium [Moles/volume] in Venous blood	39789-3
	Potassium [Mass/volume] in Blood	75940-7
Bicarbonate	Bicarbonate [Moles/volume] in Blood	1959-6
	Bicarbonate [Moles/volume] in Serum or Plasma	1963-8

Lab/Vital	Description	LOINC code
	Bicarbonate [Moles/volume] in Capillary blood	1961-2
	Bicarbonate [Moles/volume] in Central venous blood	97543-3
	Bicarbonate [Moles/volume] in Mixed venous blood	19229-4
	Bicarbonate [Moles/volume] in Venous blood	14627-4
	Bicarbonate [Moles/volume] standard in Arterial blood	19230-2
	Bicarbonate [Moles/volume] in Arterial blood	1960-4
	Bicarbonate [Moles/volume] standard in Capillary blood	19231-0
	Bicarbonate [Moles/volume] standard in Central venous blood	97544-1
	Bicarbonate [Moles/volume] standard in Mixed venous blood	19233-6
	Bicarbonate [Moles/volume] standard in Plasma	69964-5
	Bicarbonate [Moles/volume] standard in Venous blood	19232-8
Blood Urea Nitrogen	Urea nitrogen [Mass/volume] in Arterial blood	12961-9
	Urea nitrogen [Mass/volume] in Peripheral blood	12963-5
	Urea nitrogen [Mass/volume] in Venous blood	12962-7
	Urea nitrogen [Mass/volume] in Serum or Plasma	3094-0
	Urea nitrogen [Mass/volume] in Blood	6299-2
	Urea nitrogen [Moles/volume] in Serum or Plasma	14937-7
	Urea nitrogen [Moles/volume] in Blood	59570-2
Creatinine	Creatinine [Mass/volume] in Serum or Plasma	2160-0
	Creatinine [Mass/volume] in Blood	38483-4
	Creatinine [Moles/volume] in Serum, Plasma or Blood	77140-2
	Creatinine [Moles/volume] in Serum or Plasma	14682-9
	Creatinine [Moles/volume] in Blood	59826-8
	Creatinine [Mass/volume] in Arterial blood	21232-4
estimated GFR	Glomerular filtration rate/1.73 sq M.predicted among blacks [Volume Rate/Area] in Serum, Plasma or Blood by Creatinine-based formula (CKD-EPI)	88293-6
	Glomerular filtration rate/1.73 sq M.predicted among non-blacks [Volume Rate/Area] in Serum, Plasma or Blood by Creatinine-based formula (CKD-EPI)	88294-4
	Glomerular filtration rate/1.73 sq M.predicted among females [Volume Rate/Area] in Serum, Plasma or Blood by Creatinine-based formula (MDRD)	50044-7
	Glomerular filtration rate/1.73 sq M.predicted among non-blacks [Volume Rate/Area] in Serum, Plasma or Blood by Creatinine-based formula (MDRD)	48642-3
	Glomerular filtration rate/1.73 sq M.predicted [Volume Rate/Area] in Serum, Plasma or Blood by Creatinine-based formula (CKD-EPI)	62238-1
	Glomerular filtration rate/1.73 sq M.predicted among blacks [Volume Rate/Area] in Serum, Plasma or Blood by Creatinine-based formula (MDRD)	48643-1
	Glomerular filtration rate/1.73 sq M.predicted [Volume Rate/Area] in Serum, Plasma or Blood by Creatinine-based formula (MDRD)	77147-7
	Glomerular filtration rate/1.73 sq M.predicted among males [Volume Rate/Area] in Serum, Plasma or Blood by Creatinine-based formula (MDRD)	70969-1
Fasting Glucose	Fasting glucose [Mass/volume] in Capillary blood	1556-0
	Fasting glucose [Mass/volume] in Capillary blood by Glucometer	41604-0
	Fasting glucose [Mass/volume] in Serum or Plasma	1558-6

Lab/Vital	Description	LOINC code
	Fasting glucose [Mass/volume] in Venous blood	1557-8
	Fasting glucose [Moles/volume] in Blood	76629-5
	Fasting glucose [Moles/volume] in Capillary blood by Glucometer	14770-2
	Fasting glucose [Moles/volume] in Serum or Plasma	14771-0
	Fasting glucose [Moles/volume] in Serum, Plasma or Blood	77145-1
	Glucose [Mass/volume] in Serum or Plasma –10 hours fasting	10450-5
Glucose	Glucose [Mass/volume] in Serum, Plasma or Blood	74774-1
	Glucose [Moles/volume] in Blood	15074-8
	Glucose [Moles/volume] in Capillary blood	51596-5
	Glucose [Moles/volume] in Serum or Plasma	14749-6
	Glucose [Moles/volume] in Serum, Plasma or Blood	77135-2
	Glucose [Moles/volume] in Venous blood	39480-9
	Glucose [Mass/volume] in Venous blood	41652-9
	Glucose [Mass/volume] in Blood	2339-0
	Glucose [Mass/volume] in Serum or Plasma	2345-7
White Blood Cell count	Leukocytes [#/volume] in Blood by Automated count	6690-2
	Leukocytes [#/volume] in Blood by Manual count	804-5
	Leukocytes [#/volume] in Blood	26464-8
	Leukocytes [#/volume] in Blood by Estimate	49498-9
Hemoglobin	Hemoglobin [Mass/volume] in Blood	718-7
	Hemoglobin [Mass/volume] in Venous blood	30350-3
	Hemoglobin [Mass/volume] in Blood by Oximetry	55782-7
	Hemoglobin [Mass/volume] in Capillary blood	30352-9
	Hemoglobin [Mass/volume] in Blood by calculation	20509-6
	Hemoglobin [Mass/volume] in Capillary blood by Oximetry	97556-5
	Hemoglobin [Mass/volume] in Central venous blood by calculation	97550-8
	Hemoglobin [Mass/volume] in Central venous blood by Oximetry	97555-7
	Hemoglobin [Mass/volume] in Mixed venous blood	30351-1
	Hemoglobin [Mass/volume] in Mixed venous blood by Oximetry	76768-1
	Hemoglobin [Mass/volume] in Venous blood by Oximetry	76769-9
	Hemoglobin [Moles/volume] in Arterial blood	75928-2
	Hemoglobin [Moles/volume] in Blood	59260-0
	Hemoglobin [Moles/volume] in Venous blood	93846-4
	Hemoglobin [Mass/volume] in Arterial blood	30313-1
Hemoglobin A1c	Hemoglobin A1c/Hemoglobin.total in Blood	4548-4
	Hemoglobin A1c/Hemoglobin.total in Blood by JDS/JSCC protocol	62388-4
	Hemoglobin A1c/Hemoglobin.total in Blood by HPLC	17856-6
	Hemoglobin A1c/Hemoglobin.total in Blood by Electrophoresis	4549-2
Platelets count	Platelets [#/volume] in Blood by Automated count	777-3
	Platelets [#/volume] in Blood by Manual count	778-1
	Platelets [#/volume] in Plasma by Automated count	13056-7
	Platelets [#/volume] in Plasma	26516-5
	Platelets [#/volume] in Blood	26515-7
	Platelets [#/volume] in Blood by Automated count.optical	97995-5
	Platelets [#/volume] in Blood by Estimate	49497-1
	Platelets reticulated [#/volume] in Blood by Automated count	71692-8
Red Cell Distrib. Width	Erythrocyte distribution width [Ratio] by Automated count	788-0
	Erythrocyte distribution width [Entitic volume]	30384-2

Lab/Vital	Description	LOINC code
	Erythrocyte distribution width [Entitic volume] by Automated count	21000-5
	Erythrocyte distribution width [Ratio]	30385-9
BNP	Natriuretic peptide B [Mass/volume] in Blood	42637-9
	Natriuretic peptide B [Mass/volume] in Serum or Plasma	30934-4
	Natriuretic peptide B [Moles/volume] in Serum or Plasma	47092-2
	Natriuretic peptide B [Mass or Moles/volume] in Serum or Plasma	35257-5
	Natriuretic peptide.B prohormone N-Terminal [Mass/volume] in Serum or Plasma	33762-6
	Natriuretic peptide.B prohormone N-Terminal [Moles/volume] in Serum or Plasma	33763-4
	Natriuretic peptide.B prohormone N-Terminal [Mass/volume] in Serum or Plasma by Immunoassay	83107-3
	Natriuretic peptide.B prohormone N-Terminal [Moles/volume] in Serum or Plasma by Immunoassay	83108-1
	Natriuretic peptide.B prohormone N-Terminal [Mass/volume] in Blood by Immunoassay	71425-3
Cholesterol	Cholesterol [Mass/volume] in Serum or Plasma	2093-3
	Cholesterol [Moles/volume] in Serum or Plasma	14647-2
HDL	Cholesterol in HDL [Mass/volume] in Serum or Plasma	2085-9
	Cholesterol in HDL [Mass/volume] in Serum or Plasma by Electrophoresis	49130-8
	Cholesterol in HDL [Moles/volume] in Serum or Plasma	14646-4
	Cholesterol in HDL [Presence] in Serum or Plasma	27340-9
	Cholesterol in HDL [Presence] in Serum or Plasma by Electrophoresis	12771-2
LDL	Cholesterol in LDL [Mass/volume] in Serum or Plasma	2089-1
	Cholesterol in LDL [Mass/volume] in Serum or Plasma by Calculated by Martin-Hopkins	96259-7
	Cholesterol in LDL [Mass/volume] in Serum or Plasma by calculation	13457-7
	Cholesterol in LDL [Mass/volume] in Serum or Plasma by Direct assay	18262-6
	Cholesterol in LDL [Mass/volume] in Serum or Plasma by Electrophoresis	49132-4
	Cholesterol in LDL [Moles/volume] in Serum or Plasma	22748-8
	Cholesterol in LDL [Moles/volume] in Serum or Plasma by Calculated by Martin-Hopkins	96258-9
	Cholesterol in LDL [Moles/volume] in Serum or Plasma by calculation	39469-2
	Cholesterol in LDL [Moles/volume] in Serum or Plasma by Direct assay	69419-0
	Cholesterol.in LDL (real) [Mass/volume] in Serum or Plasma by VAP	55440-2
	Cholesterol.in LDL.small dense [Mass/volume] in Serum or Plasma	90364-1
	Cholesterol.in LDL.small dense [Moles/volume] in Serum or Plasma	96959-2
Triglyceride	Triglyceride [Mass/volume] in Blood	3043-7
	Triglyceride [Mass/volume] in Serum or Plasma	2571-8

Lab/Vital	Description	LOINC code
	Triglyceride [Mass/volume] in Serum or Plasma –12 hours fasting	1644-4
	Triglyceride [Mass/volume] in Serum or Plasma –fasting	3048-6
	Triglyceride [Moles/volume] in Blood	70218-3
	Triglyceride [Moles/volume] in Serum or Plasma	14927-8
	Triglyceride [Moles/volume] in Serum or Plasma –12 hours fasting	30524-3
	Triglyceride [Moles/volume] in Serum or Plasma –fasting	47210-0
Magnesium	Magnesium [Mass/volume] in Blood	21377-7
	Magnesium [Mass/volume] in Serum or Plasma	19123-9
	Magnesium [Moles/volume] in Blood	2593-2
	Magnesium [Moles/volume] in Serum or Plasma	2601-3
	Magnesium Ionized [Mass/volume] in Serum or Plasma	32698-3
	Magnesium Ionized [Moles/volume] in Blood by Ion-selective membrane electrode (ISE)	73572-0
	Magnesium Ionized [Moles/volume] in Serum or Plasma	2600-5
Albumin	Albumin [Mass/volume] in Serum or Plasma	1751-7
	Albumin [Mass/volume] in Serum or Plasma by Electrophoresis	2862-1
	Albumin [Mass/volume] in Blood by Bromocresol purple (BCP) dye binding method	76631-1
	Albumin [Mass/volume] in Serum or Plasma by Bromocresol green (BCG) dye binding method	61151-7
	Albumin [Mass/volume] in Serum or Plasma by Bromocresol purple (BCP) dye binding method	61152-5
	Albumin [Mass/volume] in Serum, Plasma or Blood by Bromocresol purple (BCP) dye binding method	77148-5
	Albumin [Moles/volume] in Serum or Plasma	54347-0
	Albumin [Moles/volume] in Serum or Plasma by Bromocresol green (BCG) dye binding method	62235-7
	Albumin [Moles/volume] in Serum or Plasma by Bromocresol purple (BCP) dye binding method	62234-0
Lymphocytes	Lymphocytes [#/volume] in Blood by Automated count	731-0
	Lymphocytes [#/volume] in Blood	26474-7
	Lymphocytes [#/volume] in Blood by Manual count	732-8
Uric Acid	Urate [Mass/volume] in Serum or Plasma	3084-1
	Urate [Moles/volume] in Serum or Plasma	14933-6
	Urate [Mass/volume] in Blood	98981-4
	Urate [Moles/volume] in Urine	14934-4
	Urate [Mass/volume] in Urine	3086-6
	Urate [Mass/volume] in 2 hour Urine	57386-5
	Urate [Mass/volume] in 12 hour Urine	57332-9
	Urate [Moles/volume] in 24 hour Urine	25997-8
	Urate [Mass/volume] in 24 hour Urine	21587-1
Body-mass index	Body mass index (BMI) [Percentile]	59574-4
	Body mass index (BMI) [Percentile] Per age	59575-1
	Body mass index (BMI) [Percentile] Per age and sex	59576-9
	Body mass index (BMI) [Ratio]	39156-5
	Body mass index (BMI) [Ratio] Estimated	89270-3
Height	Body height	8302-2
Weight	Body weight	29463-7
Heart rate	Heart rate	8867-4
Systolic blood pressure	Systolic blood pressure	8480-6

Lab/Vital	Description	LOINC code
Diastolic blood pressure	Diastolic blood pressure	8462-4

  - See details for loading echocardiogram data under Echo measurement.

- **OBSERVATION**
- **DEATH**
- **NOTE**
  - We recommend loading notes for echocardiogram, cardiac cath, and ECG if possible. Our initial priority will be loading echo notes locally. At Northwestern, we intend to use Philter 1.0 text de-identification tool. It has some nice features and it is certified. More information can be found in the JAMIA Open, Volume 6, Issue 3, October 2023, ooad045 article. Another possibility is to use an open source tool from Stanford called TiDE.
- NOTE_NLP
- SPECIMEN
- FACT_RELATIONSHIP

2.2.2 Health System Data Tables

https://ohdsi.github.io/CommonDataModel/cdm53.html#Health_System_Data_Tables

- **LOCATION**
  - We'll be collecting patient's 5-digit address zip codes only. For this purpose, LOCATION table needs to be populated with all USA 5-digit zip codes. A file in CSV format will be distributed to sites for loading into the table. This will also allow us to explore relevance of socioeconomic status, since various SES indicators exist at zip code level.
- **LOCATION_HISTORY**
  - OMOP CDM v.5.3.1 does not have a way to record patient's historical addresses. That is why LOCATION_HISTORY table has been introduced in OMOP CDM v.6.0. If address history is available for the patients, please create and populate this table. OMOP CDM guideline is that current (most recent) patient address should be stored in PERSON table, in *location_id* field.
- CARE_SITE
- PROVIDER

2.2.3 Health Economics Data Tables

https://ohdsi.github.io/CommonDataModel/cdm53.html#Health_Economics_Data_Tables

- PAYER_PLAN_PERIOD
- COST

2.2.4 Standardized Derived Elements

https://ohdsi.github.io/CommonDataModel/cdm53.html#Standardized_Derived_Elements

- **DRUG_ERA**
- DOSE_ERA
- **CONDITION_ERA**

2.2.5 Vocabulary Tables

https://ohdsi.github.io/CommonDataModel/cdm53.html#Vocabulary_Tables

- **CONCEPT**
- **VOCABULARY**
- **DOMAIN**
- **CONCEPT_CLASS**
- **CONCEPT_RELATIONSHIP**
- **RELATIONSHIP**
- **CONCEPT_SYNONYM**
- **CONCEPT_ANCESTOR**
- SOURCE_TO_CONCEPT_MAP
- DRUG_STRENGTH
- **COHORT_DEFINITION**
  - As HeartShare study involves multiple cohorts, we will need to be able to keep track of cohorts and its subjects. COHORT_DEFINITION table will be used to define cohorts of interest. The table should be populated with records as shown in the table below.

cohort_definition_id	cohort_definition_name	definition_type_concept_id	subject_concept_id
1	HeartShare EHR	0	1147026
2	HeartShare Registry	0	1147026
3	HeartShare Deep Phenotyping	0	1147026

- **COHORT**
  - Participating sites will be responsible for assignment of patients into cohorts. For *subject_id* in COHORT table use *person_id* from PERSON table and for *cohort_definition_id* use appropriate *cohort_definition_id* from COHORT_DEFINITION table. For HeartShare EHR cohort, use *observation_period_start_date* and *observation_period_end_date* from OBSERVATION_PERIOD table as *cohort_start_date* and *cohort_end_date*. For other cohorts, use patient's enrollment date as *cohort_start_date* and use *observation_period_end_date* for *cohort_end_date*. If patient withdraws from the study or completes the participation update *cohort_end_date* with the date of the event.
- ATTRIBUTE_DEFINITION

2.2.6 Vocabularies to Load

Vocabularies can be downloaded from https://athena.ohdsi.org/vocabulary/list. Size of OMOP vocabularies is typically in the range of ~20GB. At a minimum, it is recommended to use the vocabularies below:

ID (CDM V4.5)	CODE (CDM V5)	NAME	LATEST UPDATE
1	SNOMED	Systematic Nomenclature of Medicine - Clinical Terms (IHTSDO)	27-Jan-22
2	ICD9CM	International Classification of Diseases, Ninth Revision, Clinical Modification, Volume 1 and 2 (NCHS)	30-Sep-14
3	ICD9Proc	International Classification of Diseases, Ninth Revision, Clinical Modification, Volume 3 (NCHS)	30-Sep-14
4	CPT4	Current Procedural Terminology version 4 (AMA)	30-Apr-23
5	HCPCS	Healthcare Common Procedure Coding System (CMS)	30-Jun-23
6	LOINC	Logical Observation Identifiers Names and Codes (Regenstrief Institute)	14-Aug-23
8	RxNorm	RxNorm (NLM)	2-Jul-23
9	NDC	National Drug Code (FDA and manufacturers)	26-Aug-23
12	Gender	OMOP Gender	
13	Race	Race and Ethnicity Code Set (USBC)	
14	CMS Place of Service	Place of Service Codes for Professional Claims (CMS)	
21	ATC	WHO Anatomic Therapeutic Chemical Classification	6-Sep-21
35	ICD10PCS	ICD-10 Procedure Coding System (CMS)	30-Sep-20
40	DRG	Diagnosis-related group (CMS)	
41	MDC	Major Diagnostic Categories (CMS)	
44	Ethnicity	OMOP Ethnicity	
70	ICD10CM	International Classification of Diseases, Tenth Revision, Clinical Modification (NCHS)	30-Sep-22
82	RxNorm Extension	OMOP RxNorm Extension	23-Aug-23
88	CVX	CDC Vaccine Administered CVX (NCIRD)	17-Apr-23
115	Provider	OMOP Provider	
128	OMOP Extension	OMOP Extension (OHDSI)	30-May-23

2.2.7 Echo Measurement

Echo measurements should be loaded into the MEASUREMENT table. Most of the echo variables could be matched to appropriate LOINC code(s). For variables where suitable standard codes were not found, custom vocabulary and custom concept mappings were created for capturing those values.The echo measurements should be able to be linked to echos in the PROCEDURE_OCCURRENCE table. For left ventricular ejection fraction, we recommend mapping the value in the echo reported by the echocardiographer to the most appropriate LOINC code. If the method of measurement is unknown, then use LOINC Code 8806-2, Left ventricular Ejection fraction by 2D echo.

Echo Variables Mappings

Echo Variables Custom Vocabulary

Previous Data Calls

EHR Data Call for demographic and LVEF

Approval date: Monday 2022-12-19

Due date: Wednesday 2023-03-01

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: $\geq$ 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 $\geq$ 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*

NIH National Heart, Lung, and Blood Institute HEARTSHARE