

Epic Data Sources: Mount Sinai Data Warehouse

Intro to Query Tools for Research

Spring 2021



**Mount
Sinai**

Epic at a glance

- In place at MSH since 2011, MSQ since 2013
- In clinics since 2008 (now hundreds)
- Documentation from providers, nurses
- Visit details, timestamped events
- Orders, “flowsheets”
- Pulls in data from ADT, Bed Mgmt, Labs, Rads, Muse, Echo, Cath, Path, financials & more

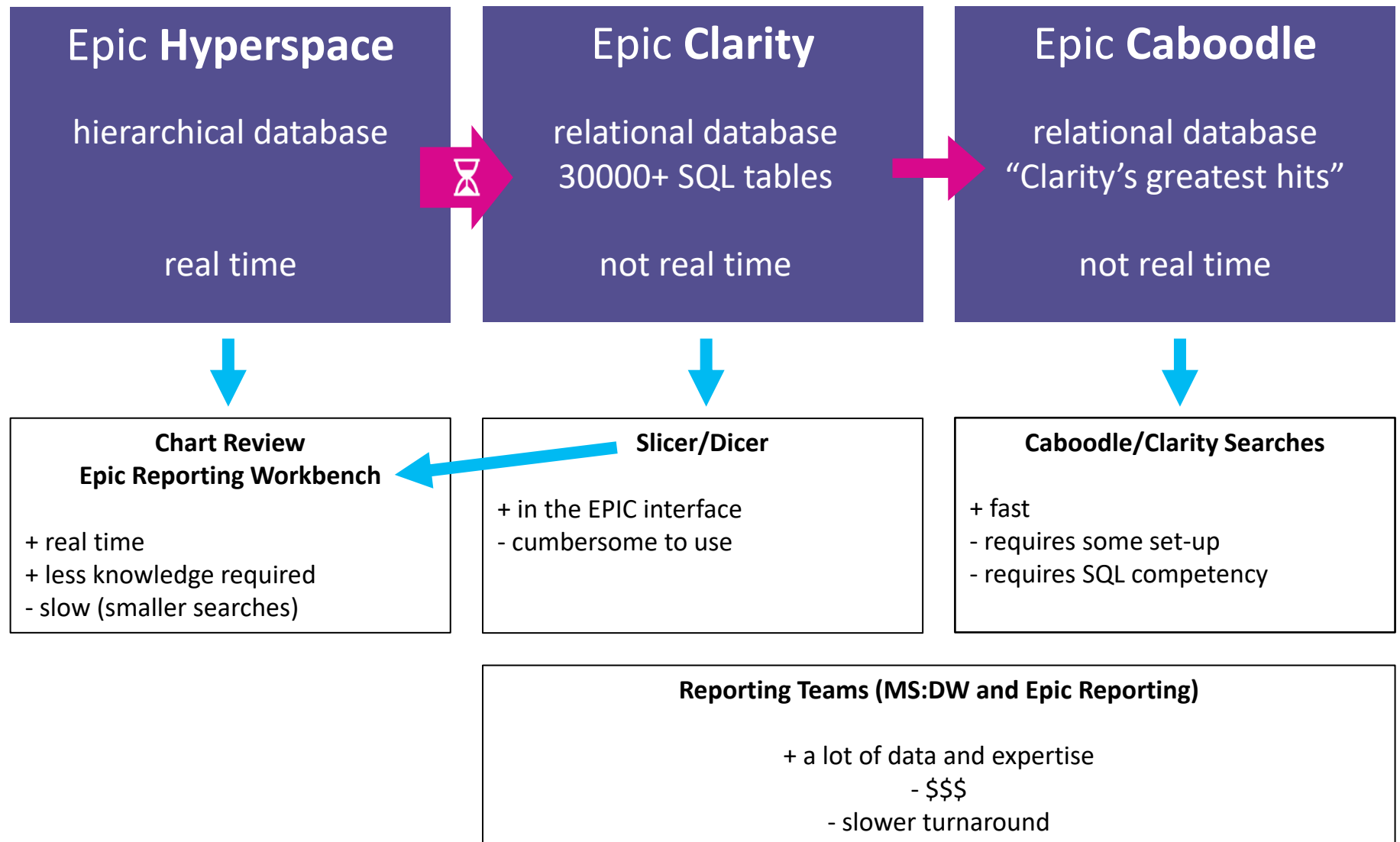


Self Service Tools

- Built in Epic Tools
 - Slicer Dicer
 - Reporting Workbench
- Need to Work with Data Team
 - Caboodle (not self-service)
- Other Sources
 - Mount Sinai Data Warehouse (MSDW)



Data in Epic



MSDW at a glance

- Data from MSHS going back to 2003
 - 4+ million patients, 25+ million visits,
 - Billions of “facts”
- Some data is validated and reorganized
 - Some data is obfuscated (HIV, VIPs)
- Includes data from:
 - ADT
 - Financial systems
 - Labs
 - Path
 - And other stand alone systems



Data Sources

Mount Sinai Data Warehouse collects data from various sources, including EPIC, CERNER and others. [More Info](#)

8 million
Patient Records

3.9 million

Patients w/ Clinical Data

98 million

Patient Visits

(inpatient, outpatient and ED)

115 million

Diagnoses

(ICD-9 and ICD-10 Coded)

355 million

Test Results

(lab, radiology and pathology)

166 million

Medication Records

(e.g. prescription and med-admins)

47 million

Clinical Documentations

(e.g. progress notes, discharge summaries, and operative reports)

Gender		Top 5 Races		Top 5 Ethnicities		Age Groups	
Female	2,139,091	White	998,230	Hispanic	70,774	0 - 2	52,470
Male	1,659,054	African American	270,141	Puerto Rican	39,990	3 - 12	221,781
Indeterminate	24,745	Hispanic Latino	135,338	Dominican	32,396	13 - 18	140,007
Total	3,822,890	Asian	39,529	Latin American	28,795	19 - 33	752,328
		Asian (Pac. Islander)	26,037	Mexican	8,418	34 - 48	869,510
		Other	2,353,614	Other	3,642,516	49 - 64	851,948
		Total	3,822,890	Total	3,822,890	65 - 78	556,970
						79 and up	357,406
						Other	20,471
						Total	3,822,890



Please submit support requests through the [MSDW ticketing system](#).

Data Services

The MSDW team offers a comprehensive list of [data services](#), including ad-hoc reports and these software tools...

Cohort Query Tool

The Cohort Query Tool (CQT) is an easy-to-use, self-service web app available to the research community at Mount Sinai to query the MSDW.

[Tutorial](#)[App \[↗\]\(#\)](#)

i2b2

The i2b2 web application provides self-service access to a de-identified data mart maintained by the MSDW team.

[Tutorial](#)[App \[↗\]\(#\)](#)

OMOP and OHDSI ATLAS

A web-based integrated platform for database exploration, standardized vocabulary browsing, cohort definition, and population-level analysis. Login using your @mssm.edu Active Directory credentials.

[Tutorial](#)[App \[↗\]\(#\)](#)

Data Dictionary

The Data Dictionary illustrates transactions and data points that are collected from each source. It is a good resource to gain an understanding of available data.

[App \[↗\]\(#\)](#)

TriNetX

TriNetX is now available. Please request an account [here](#) to get started.

[Tutorial](#)[App \[↗\]\(#\)](#)

Tools to Choose From



Please submit support requests through the [MSDW ticketing system](#).

Data Services

The MSDW team offers a comprehensive list of [data services](#), including ad-hoc reports and these software tools...

Cohort Query Tool

The Cohort Query Tool (CQT) is an easy-to-use, self-service web app available to the research community at Mount Sinai to query the MSDW.

[Tutorial](#)[App \[↗\]\(#\)](#)

i2b2

The i2b2 web application provides self-service access to a de-identified data mart maintained by the MSDW team.

[Tutorial](#)[App \[↗\]\(#\)](#)

OMOP and OHDSI ATLAS

A web-based integrated platform for database exploration, standardized vocabulary browsing, cohort definition, and population-level analysis. Login using your @mssm.edu Active Directory credentials.

[Tutorial](#)[App \[↗\]\(#\)](#)

Data Dictionary

The Data Dictionary illustrates transactions and data points that are collected from each source. It is a good resource to gain an understanding of available data.

[App \[↗\]\(#\)](#)

TriNetX

TriNetX is now available. Please request an account [here](#) to get started.

[Tutorial](#)[App \[↗\]\(#\)](#)

MSDW Cohort Query Tool

- Web-based way to browse
 - Aggregated, de-identified data
 - Data from Epic, Cerner, Eagle, TDS and other systems in the MSDW
- Can Union / Intersection sets
- If IRB-approved, can create sets of interest with identifiers



[Home](#)
[New Search](#)
[Metadata Browser](#)
[Data Dictionary](#)
[Datamart Admin ▾](#)

Available Searches

- [i Diagnosis](#)
- [i DRG Diagnosis](#)
- [i Surgical Procedure](#)
- [i Lab Result Numeric](#)
- [i Lab Report Text](#)
- [i Orders](#)
- [i Medication](#)
- [i EKG Report Text](#)
- [i Radiology Report](#)
- [i Pathology Report](#)
- [i Medication History](#)
- [i Past Surgical History](#)
- [i Past Medical History](#)
- [i Facility](#)**
- [i Clinical Documentation](#)

Build Code List

Search ▾ EPIC ▾ 



Sort By ▾

<input checked="" type="checkbox"/>	5119226: EMERGENCY DEPARTMENT	385855
<input type="checkbox"/>	2133001: EMERGENCY DEPT QUEENS	169602
<input type="checkbox"/>	5133002: EMERGENCY DEPARTMENT PSYCH	3619
<input type="checkbox"/>	5133006: EMERGENCY DEPT BILLING	36
<input type="checkbox"/>	5124174: EMERGENCY DEPT WALK IN	1
<input type="checkbox"/>	5124177: EMERGENCY DEPT PED	1

Query Filters

[Selected List](#)

[Save Library](#)

☒ Event Date From:  To: 

☐ Age

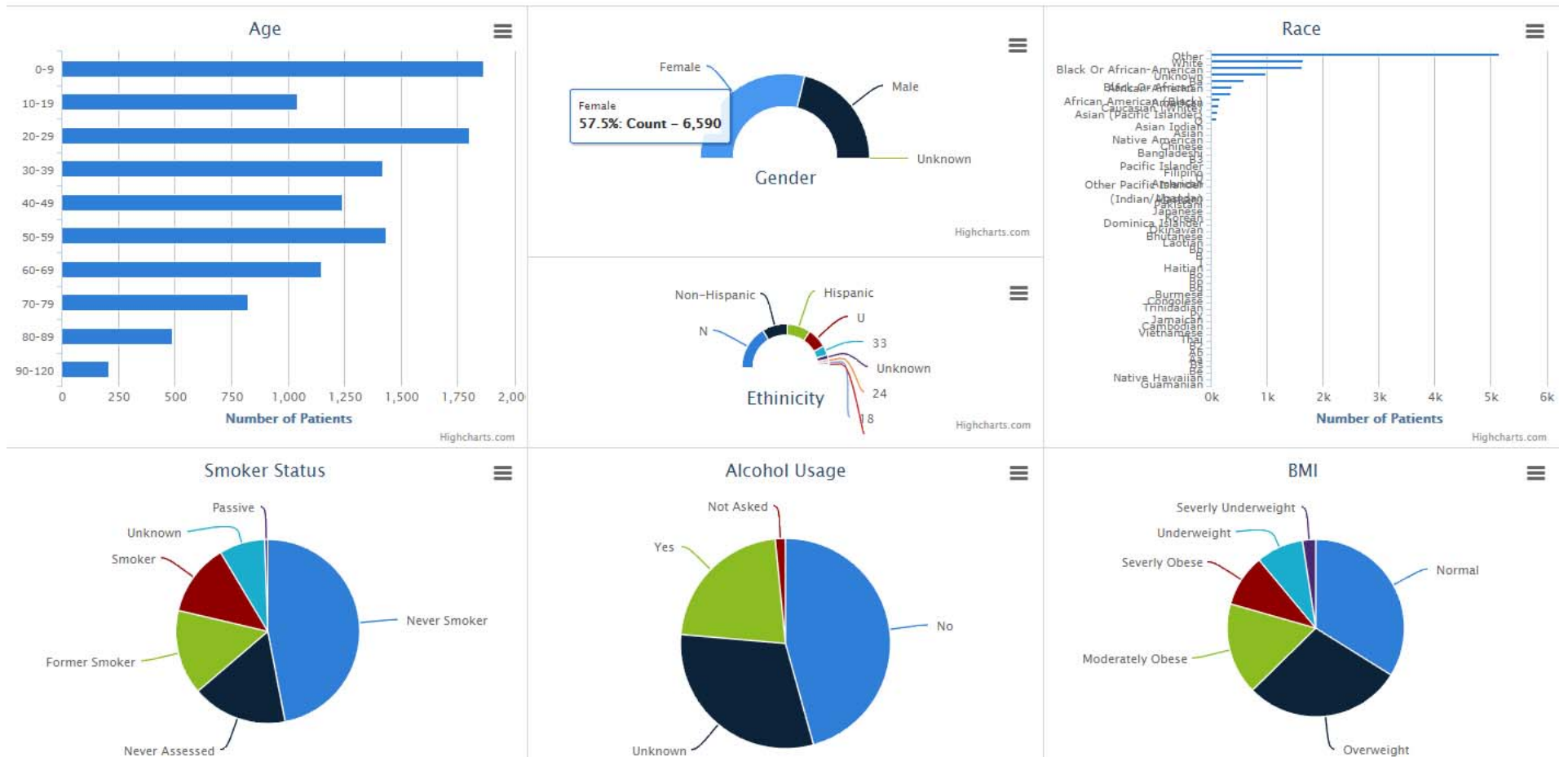
☐ Gender

☐ Race

☐ Death Status

Search parameters and selection tools

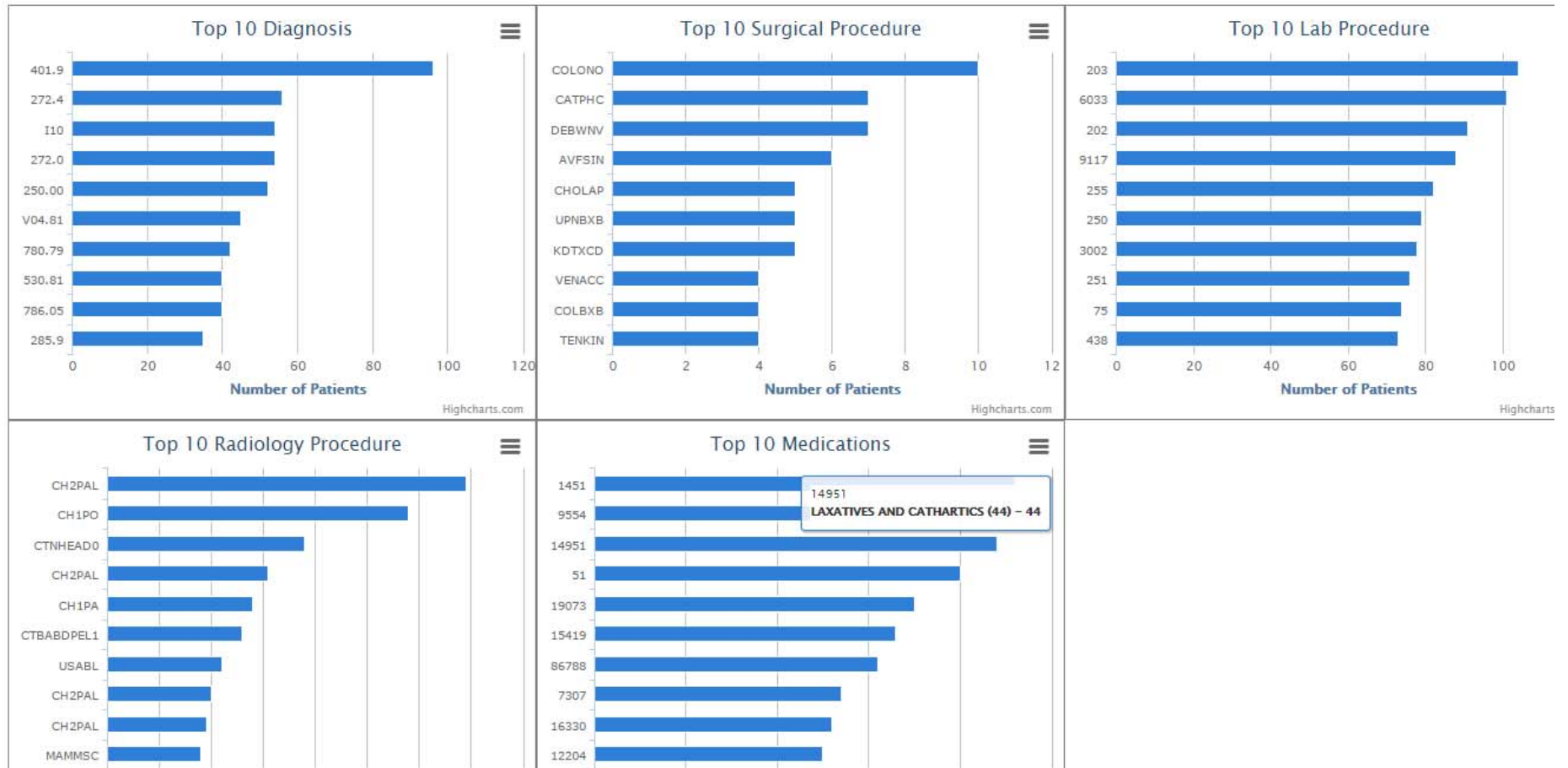
Demographics for Set: 95419 : ED1 ED1 : 11461



Example of Results from Query

Frequency chart for Set: 55945 : Diagnosis [55945] Diagnosis: 55945 : 108

Frequency Charts



Example of Results from Query

Panel for Set: 95419 - ED1 : ED1 (N = 11461)

[Return to Home Page](#)

Panel: All Demographics Detail - NoPHI Show Clear

Download

Displaying 200 of 11461 rows

Set Id	Msdw Person#	Gender	Race	Citizenship	Language	Marital Status Code	Religion	Address Zip	Address County	Address Country	Address Type	Age	Patient Ethnic Group	Interpreter Required	Interpreter Type	Dec Ind
95419	17326358	Female	White		English	Single	None	112**	Brooklyn (NY)	USA	home	51	24			No
95419	18154491	Female	Ba		English	Single	None	100**	New York	USA	home	<=18	N			No
95419	1836333	Female	Black Or African-American		English	Single	Christian	100**	New York	Usa	home	26	N	No		No
95419	18508183	Male	White		English	Married	None	100**	New York	Usa	home	64	Unknown			Yes
95419	629131	Female	Other		English	Significant Other/Life Partner	Catholic	104**	Bronx (NY)	Usa	home	49	U	No		No
95419	633220	Female	Black Or African-American		English	Single	Baptist	100**	New York	Usa	home	84	N			No
95419	633810	Male	Other		English	Separated	Other	100**	New York	Usa	home	62	N			No
95419	649001	Male	Other		English	Single	Other	100**	New York	USA	home	36	Hispanic			No
95419	651520	Female	Other		English	Single	Catholic	104**	New York	Usa	home	35	33	No		No
95419	654503	Male	Other		English	Single	Catholic	100**	New York	USA	home	36	Hispanic			No
95419	655948	Female	Black Or African-American		English	Married	None	100**	New York	Usa	home	37	N			No
95419	668249	Male	White		English	Married	Lutheran	105**	Westchester	USA	home	53	N			No
95419	671576	Male	White		English	Single	Jewish	100**	New York	Usa	home	84	U	No		No
95419	6789866	Male	Other		English	Single	None	100**	New York	USA	home	34	Hispanic	No		No
95419	685863	Female	Unknown		English	Unknown	Unknown	100**	New York	USA	home	68	Unknown			No
95419	18833014	Male	Other	USA	English	Single	Catholic	104**	Bronx (NY)	USA	home	<=18	Hispanic			No
95419	18842267	Male	Other		English	Single	None	100**	New York	USA	home	<=18	Hispanic			No
95419	1909401	Female	Other		English	Single	Hindu	114**	Queens (NY)	USA	home	28	N			No
95419	1942764	Male	Other		English	Single	Christian	100**	New York	Usa	home	78	Non-Hispanic			No
95419	1952440	Male	Other		Spanish	Single	Catholic	100**	New York	USA	home	62	33	No		No
95419	19840428	Male	African-American		English	Single	None	114**	Queens (NY)		home	46	N			No
95419	19895413	Female	White		English	Widow(er)	Jewish	117**	Suffolk	USA	home	82	N			No
95419	757265	Female	Other		English	Single	Pt Declined	100**	New York	Usa	home	37	U	No		No
95419	76229	Male	Other		English	Single	Catholic	100**	New York	USA	home	44	Hispanic	No		No
95419	776964	Female	Black Or African-American		English	Single	Pt Declined	100**	New York	Usa	home	29	Non-Hispanic	No		No

Download

Example of Deidentified Patient Level Results from Query

MSDW Cohort Query Tool

- Pros:
 - Self-service (No external costs)
 - Easy to use
 - Provides estimates of a wide range of patient and visit level factors
- Cons:
 - Only de-identified data available (need to work with analyst to pull data once project is IRB approved)
 - Data is aggregated or obfuscated



Additional MSDW Tools

The screenshot displays the MSDW website interface. At the top, a blue navigation bar contains the MSDW logo, links for Policies, Using MSDW, and About, along with links for Resident Tracking and Scientific Computing. Below the navigation bar, a green banner with a question mark icon prompts users to submit support requests through the MSDW ticketing system. The main content area is titled "Data Services" and includes a sub-header stating that the MSDW team offers a comprehensive list of data services, including ad-hoc reports and software tools. Five tool cards are displayed: Cohort Query Tool, i2b2, OMOP and OHDSI ATLAS, Data Dictionary, and TriNetX. Each card provides a brief description of the tool and includes links for a Tutorial and an App. The OMOP and OHDSI ATLAS card and the TriNetX card are highlighted with red borders. A red arrow points from the text "MSDW has trainings on their website for each tool." to the TriNetX card.

MSDW Policies Using MSDW About Resident Tracking Scientific Computing

? Please submit support requests through the [MSDW ticketing system](#).

Data Services

The MSDW team offers a comprehensive list of [data services](#), including ad-hoc reports and these software tools...

Cohort Query Tool

The Cohort Query Tool (CQT) is an easy-to-use, self-service web app available to the research community at Mount Sinai to query the MSDW.

[Tutorial](#) [App](#)

i2b2

The i2b2 web application provides self-service access to a de-identified data mart maintained by the MSDW team.

[Tutorial](#) [App](#)

OMOP and OHDSI ATLAS

A web-based integrated platform for database exploration, standardized vocabulary browsing, cohort definition, and population-level analysis. Login using your @mssm.edu Active Directory credentials.

[Tutorial](#) [App](#)

Data Dictionary

The Data Dictionary illustrates transactions and data points that are collected from each source. It is a good resource to gain an understanding of available data.

[App](#)

TriNetX

TriNetX is now available. Please request an account [here](#) to get started.

[Tutorial](#) [App](#)

MSDW has trainings on their website for each tool.

Comparison of MSDW Tools

	Custom Queries	Cohort Query Tool	OMOP Query Tool	TriNetX
Data type	All MSDW	Facilities, Diagnoses & DRGs, Procedures, Meds, Labs, Orders, EKG reports, Rads & Path reports, Clinical Documentation, Demographics	Facilities, Diagnoses & DRGs, Procedures, Meds, Labs, Orders, EKG reports, Rads & Path reports, Demographics	Providers, Diagnoses, Procedures, Meds, Labs, Demographics
Cost	\$180/hr	Free*	Free*	Free*
PHI	Yes	Yes (if IRB Approved)	No*	No*
Turnaround	Weeks	Seconds to Minutes	Minutes	Minutes
Capabilities/ Limitations	Custom search; Requires working with MSDW team	Contains all MSDW patients; Can search BioME patients; Can search Image Research Warehouse patients	Common data model and queries; Numerous query analysis tools available in github	Intuitive user interface; Connected to >300 million external patients (fee)

Important to Remember

- How are you using the data:
 - Research requires an IRB!!!
 - Speak with research faculty about requirements
 - IRB determines if study is exempt
- Start with least amount of PHI necessary
- You are responsible for your data
 - Keep data secure (computer, email, drives)
 - Reported data must be aggregated for research purposes

