

PHYSICIAN SCIENTIST COLLECTIVE

UCI'S HEALTH DATA RESEARCH RESOURCES

ICTS • IPH

OFFICE OF RESEARCH INFORMATION

01/2026





Kai Zheng, PhD, FACMI

**Professor, Informatics & Emergency Medicine
Director, ICTS • IPH Office of Research Information**

kai.zheng@uci.edu



Funding

- **Institute for Clinical and Translational Science (ICTS)**
- **Institute for Precision Health (IPH)**
- **Various extramural federal research grants and contracts**

Mission

To provide and innovate informatics solutions to facilitate clinical research, accelerate results translation, and engender the transformation of UCI Health hospitals and clinics into a learning health system.

**Research Cyberinfrastructure
Center (RCIC)**

High-performance and high-
throughput computing; Research data
storage

**Center for Artificial
Intelligence in Diagnostic
Medicine (CAIDM)**

Access to and analysis of medical
imaging data; AI/ML; Deep learning

**Genomics Research and
Technology Hub (GRT Hub)**

Sequencing facilities; computational
-omics

**Cancer Center Experimental
Tissue Resource (ETR)**

Tissue and body fluids procurement,
management, and analysis

ICTS • IPH Office of Research Information

Enterprise Data & Analytics (EDA)

Operational reporting; Business intelligence; Management of enterprise data warehouses; Research support, particularly for quality improvement projects

UCI Health Information Technology Services (HS ITS)

Management of Epic EHR and other clinical information systems; HIPAA-compliant data centers/cloud computing capability; information security; integration of research in clinical care



Service Areas

- 1. Honest broker for retrieving UCI Health patient data according to IRB-approved research protocols**
- 2. Partially de-identified, self-service research data warehouses**
- 3. Self-service cohort discovery tools**
- 4. Advanced data analytics: artificial intelligence / machine learning (AI/ML), natural language processing (NLP)**
- 5. Research data sharing networks**
- 6. Research software development**
- 7. Researcher engagement and enrichment**



Service Areas

1. **Honest broker for retrieving UCI Health patient data according to IRB-approved research protocols**
2. **Partially de-identified, self-service research data warehouses**
3. **Self-service cohort discovery tools**
4. **Advanced data analytics: artificial intelligence / machine learning (AI/ML), natural language processing (NLP)**
5. **Research data sharing networks**
6. **Research software development**
7. **Researcher engagement and enrichment**

Center for Artificial Intelligence
in Diagnostic Medicine

Peter Chang
changp6@hs.uci.edu



Medicine
imaging center

Financial
cost data

DISCHARGE SUMMARY

SUMMARY OF TREATMENT PLANNING:

Two major problems were identified at the admission of this adolescent:

1. Mood swings.
2. Oppositional and defiant behavior.

A developmentally appropriate group oriented therapy program was the primary treatment modality for this adolescent. He participated in at least eight psychoeducational and activity groups. The attending psychiatrist provided evaluation for and management of psychotropic medications and collaborated with the treatment team. The clinical therapist facilitated individual, group, and family therapy at least twice per week.

COURSE IN HOSPITAL: During his hospitalization, the patient was seen initially as very depressed, withdrawn, some impulsive behavior observed, also oppositional behavior was displayed on the unit. The patient also talked with a therapist about his family conflicts. He was initiated on an antidepressant medication, Zoloft, and he continued with Adderall. He responded well to Zoloft, was less depressed. He continued with behavior problems and mood swings. A mood stabilizer was added to his treatment and with a positive response to it.

DIAGNOSTIC AND THERAPEUTIC TEST/EVALUATIONS: Sleep-deprived EEG was done, which was reported as normal. His last Depakote level was 57 as per 06/04/04. His laboratory basic metabolic panel, CBC, TSH were reported within normal limits.

CONSULTATIONS: He was seen by our medical consultant for a complete history and medical examination. No major acute problems were reported, only the acne. Treatment was initiated with face wash medication.

FINAL DIAGNOSIS:

AXIS I: ADHD (Attention Deficit Hyperactivity Disorder), rule out Bipolar Disorder and ODD (Oppositional Defiant Disorder).

AXIS II: Deferred.

AXIS III: Acne.

AXIS IV: Psychosocial stressors: Severe, family conflicts and educational problems.

AXIS V: AF: 45 to 50.

CONDITIONS ON DISCHARGE: The patient had appropriate mood and was not engaging in self-injurious behavior. He denied suicidal or homicidal ideation.

Height: 5 foot 8 inches. Weight: 134. Blood pressure: 120/54. Pulse: 104. Respirations: 18. Temperature: 99.

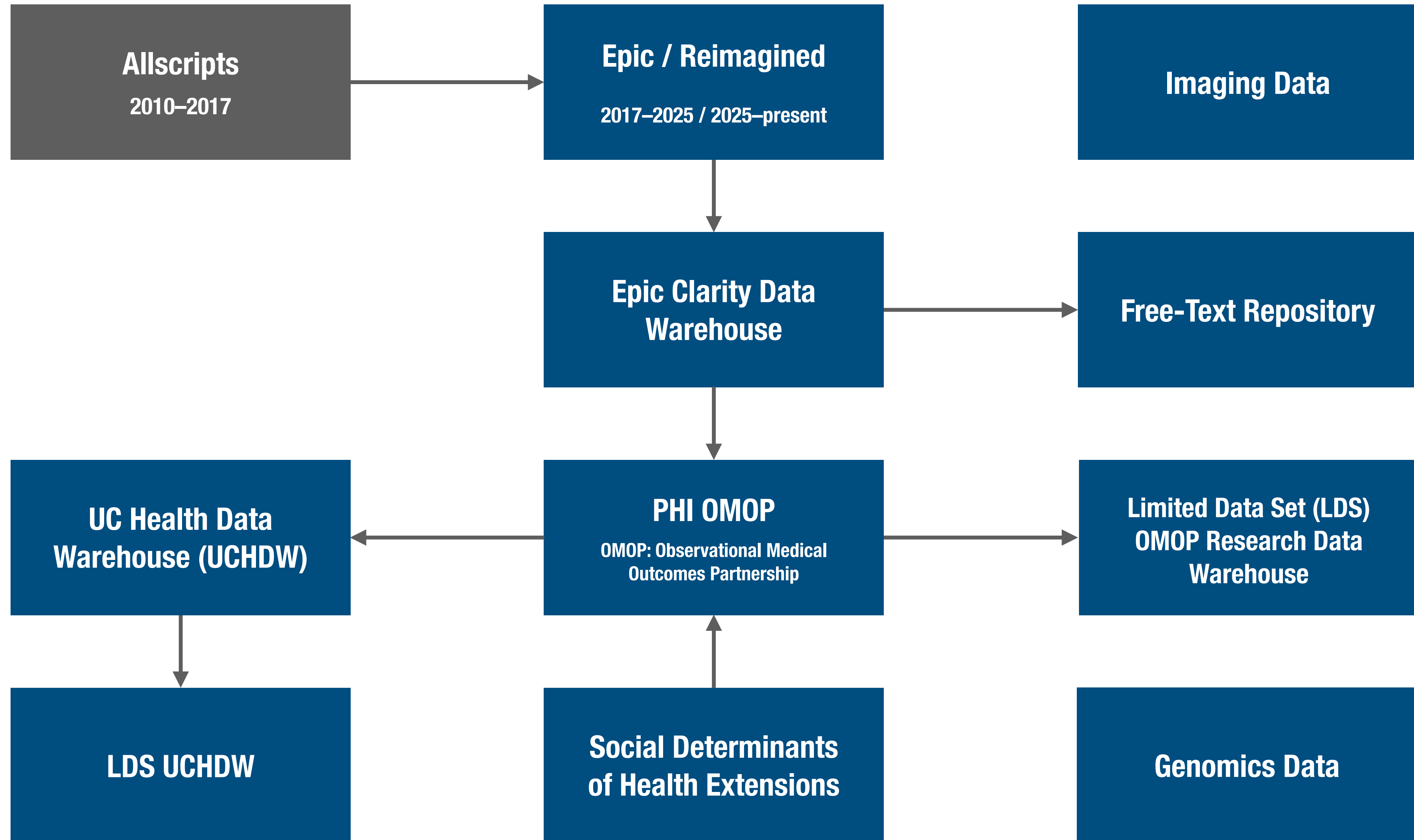
PROGNOSIS: Guarded.

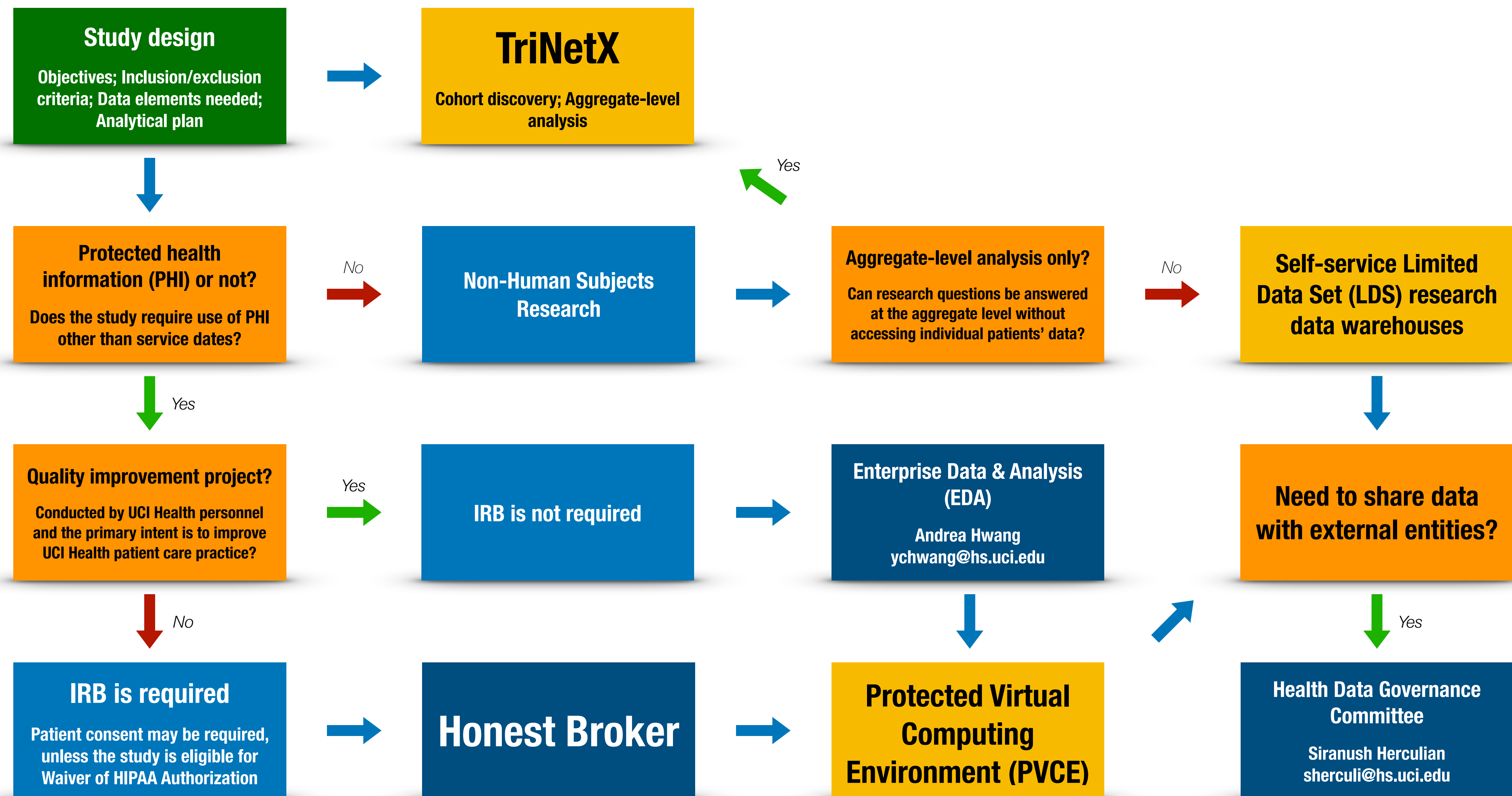
IPH Office of
Information (ORI)

ori@uci.edu

Data & Analysis
(EDA)

Andrea Hwang
ang@hs.uci.edu





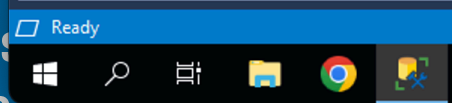
The screenshot shows a SQL query editor window titled 'SQLQuery1.sql - omop.uci.edu.LDS_OMOP_RDW (HS)'. The menu bar includes File, Edit, View, Query, Project, Tools, and Window. The toolbar contains icons for opening, saving, and executing queries. The 'Query' dropdown menu is open, showing 'LDS_OMOP_RDW' as the selected database. The 'Execute' button is visible. The 'Object Explorer' pane at the bottom shows a connection to 'omop.uci.edu (SQL Server 15.0.4249.2 - HS)zhengkai' with a 'Databases' folder expanded.

**Prot
informat
Does the stu
other th**

Quality improvement
Conducted by
and the primary
UCI Health

IRB

Patient consent
unless the
Waiver of IRB



Protected Health Information

x

+

<

>

C

https://research.uci.edu/human-research-protections/assessing-risks-and-benefits/privacy-and-confidentiality/protected-health-information-hipaa/

☆

📁

|

📄

👤

:

The Privacy Rule and Research

As noted above, HIPAA affects only that research which uses, creates, or discloses PHI. Research involves the collection, analysis, and use of information about an individual derived from his or her medical history, physical examinations, laboratory tests, interviews, questionnaires, or other sources. The Privacy Rule protects PHI while it is being used or disclosed for research purposes, regardless of whether the information originates from a health care provider or another source. In general, there are two types of human research that would involve the use or disclosure of PHI:

- Studies involving review of existing medical records as a source of research information. Retrospective studies often rely on such information. Sometimes prospective studies do it also, for example, when they contact a participant's physician to obtain information about the participant's health history.
- Studies that create new medical information because a health care service is being performed as part of the research, such as testing of a new way of diagnosing a health condition or a new drug or device for treating a health condition. Virtually all sponsored clinical trials that submit data to the U.S. Food and Drug Administration (FDA) will involve PHI.

The IRB's Role

The IRB acts as a Privacy Board (required by HIPAA) to review the use/disclosure of PHI and to determine whether the subjects should sign an authorization (an addendum to the consent to participate in research) or if a waiver of consent (roughly analogous to a Waiver of Consent under the [Common Rule](#)) may be granted.

When the IRB determines that subjects should sign a HIPAA research authorization in order to use or disclose PHI for research, subjects are to sign the UC HIPAA research authorization as a part of the informed consent process for participation in the study.

Requesting a Waiver of HIPAA Authorization

Although it is always preferred to get permission / authorization to use an individual's PHI, HIPAA permits research using PHI without obtaining authorization. This is referred to as a waiver of HIPAA research authorization, which is granted by the UCI IRB.

In order to waive HIPAA Authorization, the IRB must determine that the study meets all of the following criteria:

- The use or disclosure of PHI involves no more than minimal risk
- Granting of the waiver will not adversely affect privacy rights and welfare of the individuals whose records will be used
- The project could not practicably be conducted without a waiver
- The project could not practicably be conducted without use of PHI
- The privacy risks are reasonable relative to the anticipated benefits of research
- An adequate plan to protect identifiers from improper use and disclosure is included in the research proposal
- An adequate plan to destroy the identifiers at the earliest opportunity, or justification for retaining identifiers, is included in the research proposal
- The project plan includes written assurances that PHI will not be re-used or disclosed for other purposes
- Whenever appropriate, the subjects will be provided with additional pertinent information after participation

Clinical Activities as Research: When IRB Review, Consent, Research HIPAA and California Bill of Rights Apply

The University of California, Office of the President has advised when IRB Review, Consent, Research HIPAA and the California Bill of Rights apply to clinical activities that are treated as research. The following table illustrates this advisory and practice at UCI.

	Expanded Access	Humanitarian Use	Right to Try
IRB Review	Yes	Yes	Yes
Consent Required	Yes	Yes	Yes
California Bill of Rights[1]	Yes	No	Yes
HIPAA Research Authorization	Yes	No[2]	No[3]

Please contact [HRP Staff](#) for any questions.

[1] REQUIRED FOR “medical experiments.” [HSC 24174](#)

[2] REQUIRED IF IRB has access to PHI (unusual at UCI)

[3] REQUIRED IF non-UCI covered components will access protected health information (PHI)

Human Research Protections

Do You Need IRB

IRB Forms

IRB Forms

The IRB's Role

When the IRB determines that subjects should sign a HIPAA research authorization in order to use or disclose PHI for research, subjects are to sign the UC HIPAA research authorization as a part of the informed consent process for participation in the study.

Although it is always preferred to get permission / authorization to use an individual's PHI, HIPAA permits research using PHI without obtaining authorization. This is referred to as a waiver of HIPAA research authorization, which is granted by the UCI IRB.

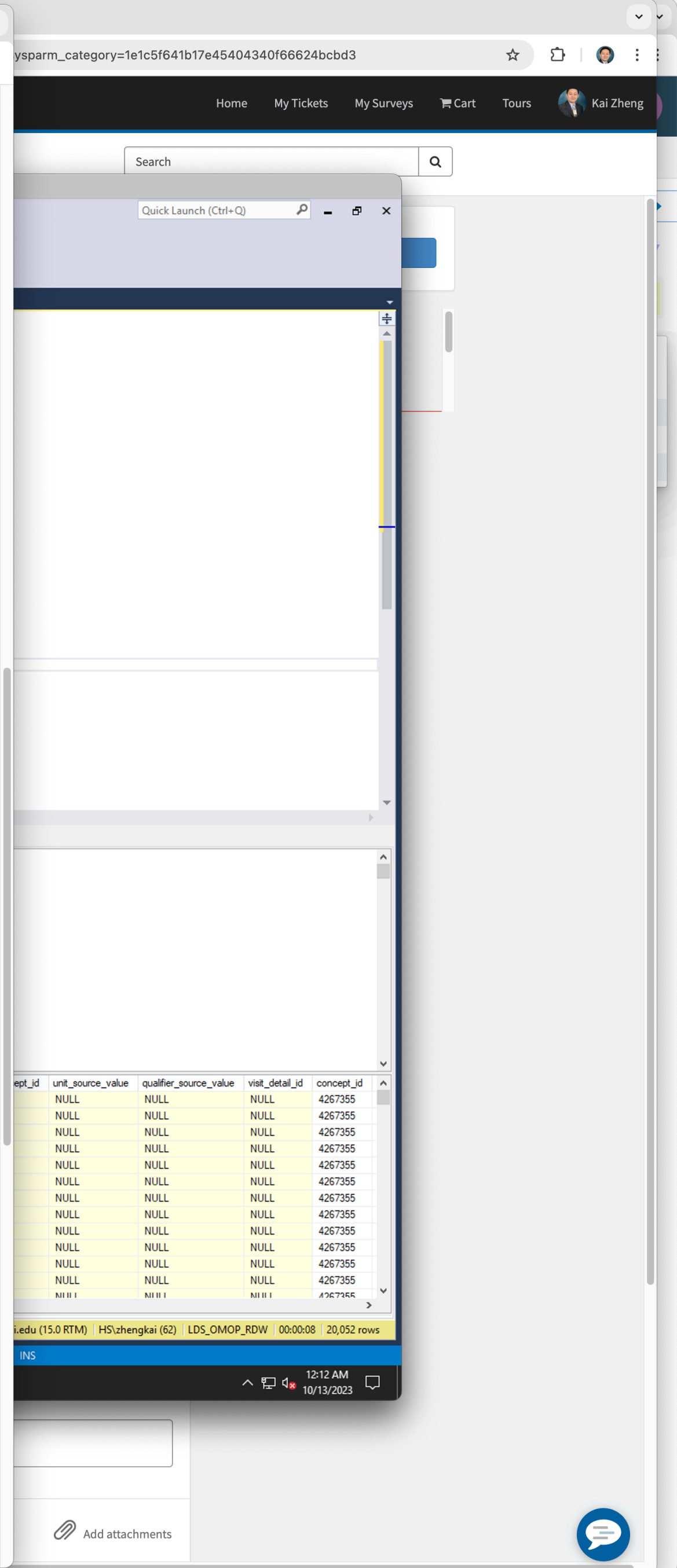
- The use or disclosure of PHI involves no more than minimal risk
- Granting of the waiver will not adversely affect privacy rights and welfare of the individuals whose records will be used
- The project could not practicably be conducted without a waiver
- The project could not practicably be conducted without use of PHI
- The privacy risks are reasonable relative to the anticipated benefits of research
- An adequate plan to protect identifiers from improper use and disclosure is included in the research proposal
- An adequate plan to destroy the identifiers at the earliest opportunity, or justification for retaining identifiers, is included in the research proposal
- The project plan includes written assurances that PHI will not be re-used or disclosed for other purposes
- Whenever appropriate, the subjects will be provided with additional pertinent information after participation

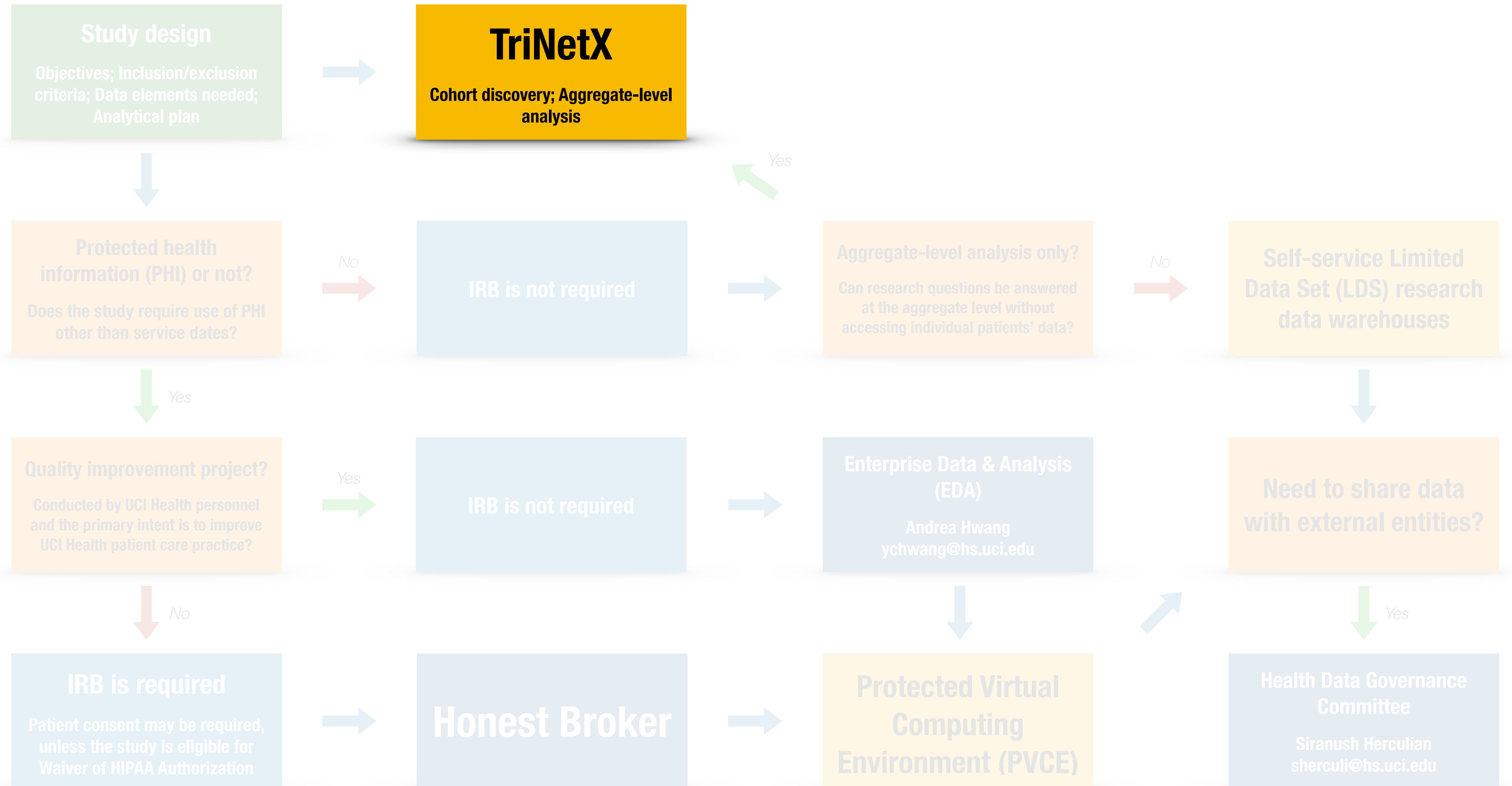
The University of California, Office of the President has advised when IRB Review, Consent, Research HIPAA and the California Bill of Rights apply to clinical activities that are treated as research. The following table illustrates this advisory and practice at UCI.

Please contact [HRP Staff](#) for any questions.

[2] REQUIRED IF IRB has access to PHI (unusual at UCI)

[3] REQUIRED IF non-UCI covered components will access protected health information (PHI)







TRIAL DESIGN

DATA SETS & ANALYTICS

NON-US ONCOLOGY

RWE GENERATION

DRUG SAFETY

Explore. Discover. Connect.

TriNetX

A cohort discovery tool for constructing patient cohorts at UCI Health or globally, conducting aggregate-level analyses, or requesting de-identified patient-level data for research purposes.

Discover the power of TriNetX, the largest real-world data, as we revolutionize drug development and medical research. Beyond merely providing data, we offer end-to-end support throughout the drug development lifecycle. Our extensive research capabilities empower researchers to publish groundbreaking findings in prestigious medical journals while leveraging our platform technologies, integrated analytics, and downloadable data sets. From optimizing clinical trial designs to refining post-market safety measures, we offer comprehensive solutions that drive innovation and elevate healthcare standards worldwide. Join us in shaping a safer, healthier future, where every data point contributes to better patient outcomes and medical advancements.

Empowering Global Research: Leading the Way in Real-World Data for Comprehensive Drug Development



TRIAL DESIGN

DATA SETS & ANALYTICS

NON-US ONCOLOGY

RWE GENERATION

DRUG SAFETY

Explore. Discover. Connect.

TriNetX

Access is free to all UCI employees and students. Email ori@uci.edu for an account.

Welcome to the world's largest, living ecosystem of global real-world data and evidence for the life sciences and healthcare communities.

Discover the power of TriNetX, the largest real-world data, as we revolutionize drug development and medical research. Beyond merely providing data, we offer end-to-end support throughout the drug development lifecycle. Our extensive research capabilities empower researchers to publish groundbreaking findings in prestigious medical journals while leveraging our platform technologies, integrated analytics, and downloadable data sets. From optimizing clinical trial designs to refining post-market safety measures, we offer comprehensive solutions that drive innovation and elevate healthcare standards worldwide. Join us in shaping a safer, healthier future, where every data point contributes to better patient outcomes and medical advancements.

Empowering Global Research: Leading the Way in Real-World Data for Comprehensive Drug Development

My Studies / IBD / Analyze Criteria / Final

- Query Builder
- Healthcare Organizations (HCOs)
- Explore Cohort
- Analyze Criteria
- Rate of Arrival
- Summary Statistics
- Analytics
- Pending Datasets
- Available Datasets
- Connect
- Study Management
- Design Assistance

Final

Oct 18, 2024 at 9:40 am by Kai Zheng

Patients

2,820

HCOs

1

New Analysis

View History

Analyze Criteria

1=

View

New Query

Hide

Filter By: Filter

Sort By: Most Recent

Request Dataset

Final

by Kai Zheng

Oct 18, 2024 9:40 AM

Patients

2,820

HCOs

1

Network

UC Irvine Health

My Studies / IBD / Explore Cohort / Demographics

- 🔍 Query Builder
- 🛠 Healthcare Organizations (HCOs)
- 👤 Explore Cohort
- Demographics
- Diagnoses
- Oncology
- Procedures
- Medications
- Labs
- Genomics

- 📊 Analyze Criteria
- 📈 Rate of Arrival
- 📊 Summary Statistics
- 📊 Analytics
- 📄 Pending Datasets
- 📄 Available Datasets
- 🔄 Connect
- ⚙ Study Management
- 🔗 Design Assistance

★ **Final** ✎

Oct 18, 2024 at 9:38 am by Kai Zheng

Patients
2,820

HCOs
1

Run

View History

Demographics

Grouped

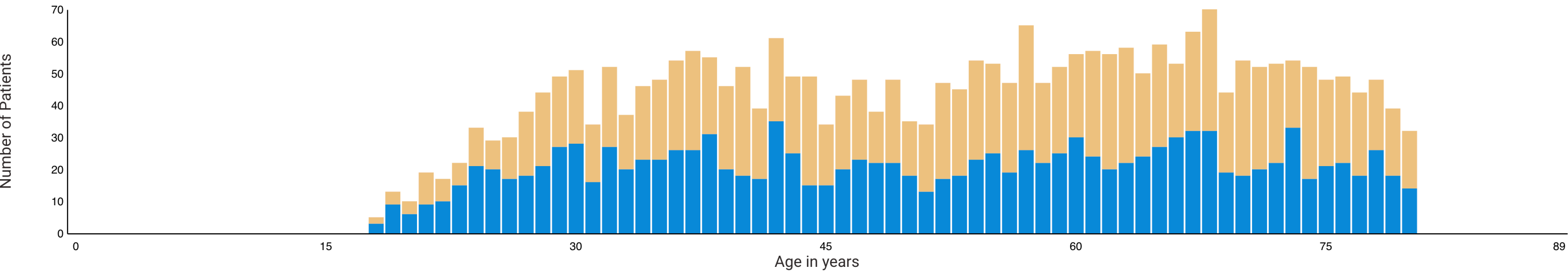
Stacked

☑ F

☑ M

☑ U

📶



Patients 90 and Older: 0

Total Patients	Minimum Age	Maximum Age	Mean Age	Standard Deviation
2,820	18	80	52	17

Sex

Female	53.19%	
Male	46.81%	
Unknown	0.00%	

Ethnicity

Not Hispanic or Latino	80.85%	
Hispanic or Latino	14.89%	
Unknown Ethnicity	4.60%	

Race

White	70.21%	
Other Race	13.12%	
Asian	7.80%	
Unknown Race	6.02%	
Black or African American	2.83%	
American Indian or Alaska Nati...	0.35%	
Native Hawaiian or Other Paci...	0.35%	

New Query

Hide ▶

Filter By: Filter ▾

Sort By: Most Recent ▾

Request Dataset

★ **Final** ✎

by Kai Zheng

Oct 18, 2024 9:51 AM

🗑

👤 Patients2,820

🏢 HCOs1

🌐 NetworkUC Irvine Health

My Studies / IBD / Explore Cohort / Diagnoses

- 🔍

Query Builder
- +

Healthcare Organizations (HCOs)
- 👤

Explore Cohort
- Demographics
- Diagnoses
- Oncology
- Procedures
- Medications
- Labs
- Genomics
- 🔍

Analyze Criteria
- 📈

Rate of Arrival
- 📊

Summary Statistics
- 📊

Analytics
- 📄

Pending Datasets
- 📄

Available Datasets
- 🔄

Connect
- ⚙️

Study Management
- 🔗

Design Assistance

★ Final ✎

Oct 18, 2024 at 9:38 am by Kai Zheng

Patients

2,820

HCOs

1

Run

View History

Diagnoses

Diagnoses within

3M6M12M24MAnytime?

AllAcuteChronic?

⬆️

Search...

⬆️

⬆️

✕

Diagnoses				Patients	% of Cohort
>	ICD-10-CM	K00-K95	Diseases of the digestive system	2,820	100%
>	ICD-10-CM	Z00-Z99	Factors influencing health status and contact with health services	2,430	86%
>	ICD-10-CM	R00-R99	Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	2,190	78%
>	ICD-10-CM	E00-E89	Endocrine, nutritional and metabolic diseases	1,770	63%
>	ICD-10-CM	M00-M99	Diseases of the musculoskeletal system and connective tissue	1,450	51%
>	ICD-10-CM	G00-G99	Diseases of the nervous system	1,370	49%
>	ICD-10-CM	D50-D89	Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	1,200	43%
>	ICD-10-CM	N00-N99	Diseases of the genitourinary system	1,200	43%
>	ICD-10-CM	I00-I99	Diseases of the circulatory system	1,190	42%
>	ICD-10-CM	F01-F99	Mental, Behavioral and Neurodevelopmental disorders	1,140	40%
>	ICD-10-CM	J00-J99	Diseases of the respiratory system	1,040	37%
>	ICD-10-CM	C00-D49	Neoplasms	1,020	36%
>	ICD-10-CM	L00-L99	Diseases of the skin and subcutaneous tissue	970	34%
>	ICD-10-CM	S00-T88	Injury, poisoning and certain other consequences of external causes	900	32%
>	ICD-10-CM	A00-B99	Certain infectious and parasitic diseases	840	30%
>	ICD-10-CM	H00-H59	Diseases of the eye and adnexa	600	21%
>	ICD-10-CM	V00-Y99	External causes of morbidity	530	19%
>	ICD-10-CM	H60-H95	Diseases of the ear and mastoid process	340	12%
>	ICD-10-CM	Q00-Q99	Congenital malformations, deformations and chromosomal abnormalities	270	10%
>	ICD-10-CM	U00-U85	Codes for special purposes	180	6%
>	ICD-10-CM	O00-O9A	Pregnancy, childbirth and the puerperium	130	5%
>	ICD-10-CM	P00-P96	Certain conditions originating in the perinatal period	20	1%

New Query

Hide ▶

Filter By: Filter ▾

Sort By: Most Recent ▾

Request Dataset

★ Final ✎

by Kai Zheng

Oct 18, 2024 9:40 AM

🗑️

👤 Patients

2,820

🏠 HCOs

1

🌐 Network

UC Irvine Health

My Studies / IBD / Explore Cohort / Medications

- 🔍

Query Builder
- +

Healthcare Organizations (HCOs)
- 👤

Explore Cohort
- Demographics
- Diagnoses
- Oncology
- Procedures
- Medications
- Labs
- Genomics
- 🔍

Analyze Criteria
- 📈

Rate of Arrival
- 📊

Summary Statistics
- 📊

Analytics
- 📄

Pending Datasets
- 📄

Available Datasets
- 🔄

Connect
- ⚙️

Study Management
- 🔗

Design Assistance

★

Final

✎

Oct 18, 2024 at 9:38 am by Kai Zheng

Patients

2,820

HCOs

1

Run

View History

Medications

Medications within

3M6M12M24MAnytime?

VAATC

⬆

Search...

⬆

⬇

✕

Medication				Patients	% of Cohort
>	VA	HS000	Hormones/synthetics/modifiers	2,820	100%
>	VA	GA000	Gastrointestinal medications	2,720	96%
>	VA	DE000	Dermatological agents	2,710	96%
>	VA	OP000	Ophthalmic agents	2,670	95%
>	VA	NT000	Nasal and throat agents,topical	2,640	94%
>	VA	CN000	Central nervous system medications	2,620	93%
>	VA	RE000	Respiratory tract medications	2,610	93%
>	VA	AM000	Antimicrobials	2,580	91%
>	VA	TN000	Therapeutic nutrients/minerals/electrolytes	2,530	90%
>	VA	IR000	Irrigation/dialysis solutions	2,300	82%
>	VA	PH000	Pharmaceutical aids/reagents	2,260	80%
>	VA	RS000	Rectal,local	2,260	80%
>	VA	IM000	Immunological agents	2,250	80%
>	VA	CV000	Cardiovascular medications	2,200	78%
>	VA	GU000	Genitourinary medications	2,200	78%
>	VA	MS000	Musculoskeletal medications	2,020	72%
>	VA	AH000	Antihistamines	1,740	62%
>	VA	AU000	Autonomic medications	1,710	61%
>	VA	VT000	Vitamins	1,580	56%
>	VA	BL000	Blood products/modifiers/volume expanders	1,290	46%
>	VA	OR000	Dental and oral agents,topical	1,280	45%
>	VA	OT000	Otic agents	1,170	41%
>	VA	Other	Other medications (va)	990	35%

New Query

Hide ▶

Filter By: Filter ▾

Sort By: Most Recent ▾

Request Dataset

★

Final

✎

by Kai Zheng

Oct 18, 2024 9:40 AM

🗑

👤

Patients

2,820

🏠

HCOs

1

🌐

Network

UC Irvine Health

TriNetX

Studies

Connect

Browse Network

What's New

1

My Studies / IBD / Rate of Arrival

Query Builder

Healthcare Organizations (HCOs)

Explore Cohort

Analyze Criteria

Rate of Arrival

Summary Statistics

Analytics

Pending Datasets

Available Datasets

Connect

Study Management

Design Assistance

★ Final

Oct 30, 2024 at 2:06 pm by Kai Zheng

Patients

2,820

HCOs

1

Run

View History

Patient Arrival Rate

Show graph

Healthcare Organization (HCO)	Historic Arrivals (Monthly Avg Over Past 3 Yrs.)	Predicted Arrivals (Monthly Avg Over Next 1 Yr.)	Trend
UC Irvine Health	34.7	30.3	

New Query

Hide

Filter By: Filter

Sort By: Most Recent

Request Dataset

★ Final

by Kai Zheng

Oct 30, 2024 2:06 PM

Patients

2,820

HCOs

1

Network

UC Irvine Health

- 🔍 Query Builder
- ✚ Healthcare Organizations (HCOs)
- 👤 Explore Cohort
- 🔑 Analyze Criteria
- 📈 Rate of Arrival
- 📊 Summary Statistics
- 📊 Analytics
- 📄 Pending Datasets
- 📄 Available Datasets
- 🌐 Connect
- ⚙️ Study Management
- 🔗 Design Assistance

★ Final ✎
Oct 18, 2024 at 9:40 am by Kai Zheng

🌐 UC Irvine Health
1 of 1 HCOs online

🔿 Filters
Any medical center type | Any data type | 18 - 80 years, Any sex

Patients
2,820

HCOs
1

🔄

MUST HAVE

CANNOT HAVE

🔍 Search Term...

🔍 Search Term...

✕ Collapse All Groups

Ungrouped Terms			
MUST HAVE			
ICD-10-CM	K50	Crohn's disease [regional enteritis]	3,540
OR			
ICD-10-CM	K51	Ulcerative colitis	3,920
AND			
VA	HS050	ADRENAL CORTICOSTEROIDS	401,880
CANNOT HAVE			
RxNorm	327361	adalimumab	4,060
OR			
TNX Curated	9063	C reactive protein [Mass/volume] in Serum, Plasma or Blood	84,650
> 🔿 ≥ 5 mg/L, most recent value			

AND

Group 1

Group Rules ▾ ⬆ ⬇ 📄 🗑

Unnamed Group ✎ This group occurred since 1 year ago

+ Terms

MUST HAVE			
CANNOT HAVE			
CPT	1007622	Endoscopy Procedures on the Rectum	32,580

✕ Collapse All Groups

+ Create a New Group

New Query

Hide ▶

Filter By: Filter ▾Sort By: Most Recent ▾

Request Dataset

★ Final ✎
by Kai Zheng
Oct 18, 2024 9:40 AM
🗑

👤 Patients	2,820
🏢 HCOs	1
🌐 Network	UC Irvine Health

My Studies / IBD / Query Builder / Networks

- 🔍 Query Builder
- ✚ Healthcare Organizations (HCOs)
- 👤 Explore Cohort
- 🔹 Analyze Criteria
- 📈 Rate of Arrival
- 📊 Summary Statistics
- 📊 Analytics
- 📄 Pending Datasets
- 📄 Available Datasets
- 🌐 Connect
- ⚙️ Study Management
- 🔗 Design Assistance

★ Final ✎

Oct 18, 2024 at 9:40 am by Kai Zheng

Patients

HCOs

2,820

1



🌐 UC Irvine Health
1 of 1 HCOs online

🔹 Filters
Any medical center type | Any data type | 18 - 80 years, Any sex



Select a network from the list below

APAC Collaborative Network | 5,285,593 Patients

● 15 of 15 Online | Last Update: 4 days ago



Connect



Analytics



Data sets

EMEA Collaborative Network | 19,891,151 Patients

● 27 of 27 Online | Last Update: an hour ago



Connect



Analytics



Data sets

Global Collaborative Network | 155,718,354 Patients

● 131 of 131 Online | Last Update: 35 minutes ago



Connect



Analytics



Data sets

LATAM Collaborative Network | 31,930,289 Patients

● 32 of 32 Online | Last Update: 35 minutes ago



Connect



Analytics



Data sets

Research | 133,769,938 Patients

● 95 of 95 Online | Last Update: 35 minutes ago



Connect



Analytics



Data sets

UC Collaborative Network | 2,691,240 Patients

● 3 of 3 Online | Last Update: 9 days ago



Connect



Analytics



Data sets

UC Irvine Health | 1,161,440 Patients

● 1 of 1 Online | Last Update: 2 months ago



Connect



Analytics



Data sets

US Collaborative Network | 118,293,935 Patients

● 67 of 67 Online | Last Update: 12 hours ago



Connect



Analytics



Data sets

Cancel

+ New Query

Hide ▶

Filter By: Filter ▾

Sort By: Most Recent ▾

Request Dataset

★ Final ✎

by Kai Zheng

Oct 18, 2024 9:40 AM



👤 Patients

2,820

🏠 HCOs

1

🌐 Network

UC Irvine Health

TriNetX

Studies ▾

Connect

Browse Network

What's New ▾

?

KZ

My Studies / IBD / Query Builder

Query Builder

Healthcare Organizations (HCOs)

Explore Cohort

Analyze Criteria

Rate of Arrival

Summary Statistics

Analytics

Pending Datasets

Available Datasets

Connect

Study Management

Design Assistance

★ Unnamed

Oct 18, 2024 at 9:47 am by Kai Zheng

Research 95 of 95 HCOs online

Filters Any medical center type | Any data type | 18 - 80 years, Any sex

MUST HAVE

Search Term...

Collapse All Groups

Ungrouped Terms

MUST HAVE

ICD-10-CM K50 Crohn's disease [regional enteritis] 232,510

OR

ICD-10-CM K51 Ulcerative colitis

AND

VA HS050 ADRENAL CORTICOSTEROIDS

AND

Group 1

Unnamed Group This group occurred since 1 year ago

MUST HAVE

CPT 1007622 Endoscopy Procedures on the Rectum 3,965,290

Collapse All Groups

+ Create a New Group

Patients 219,875

HCOs 91

Filter By: Filter ▾

Sort By: Most Recent ▾

Request Dataset

★ Unnamed

by Kai Zheng

Oct 18, 2024 9:47 AM

Patients 219,875

HCOs 91

Network Research

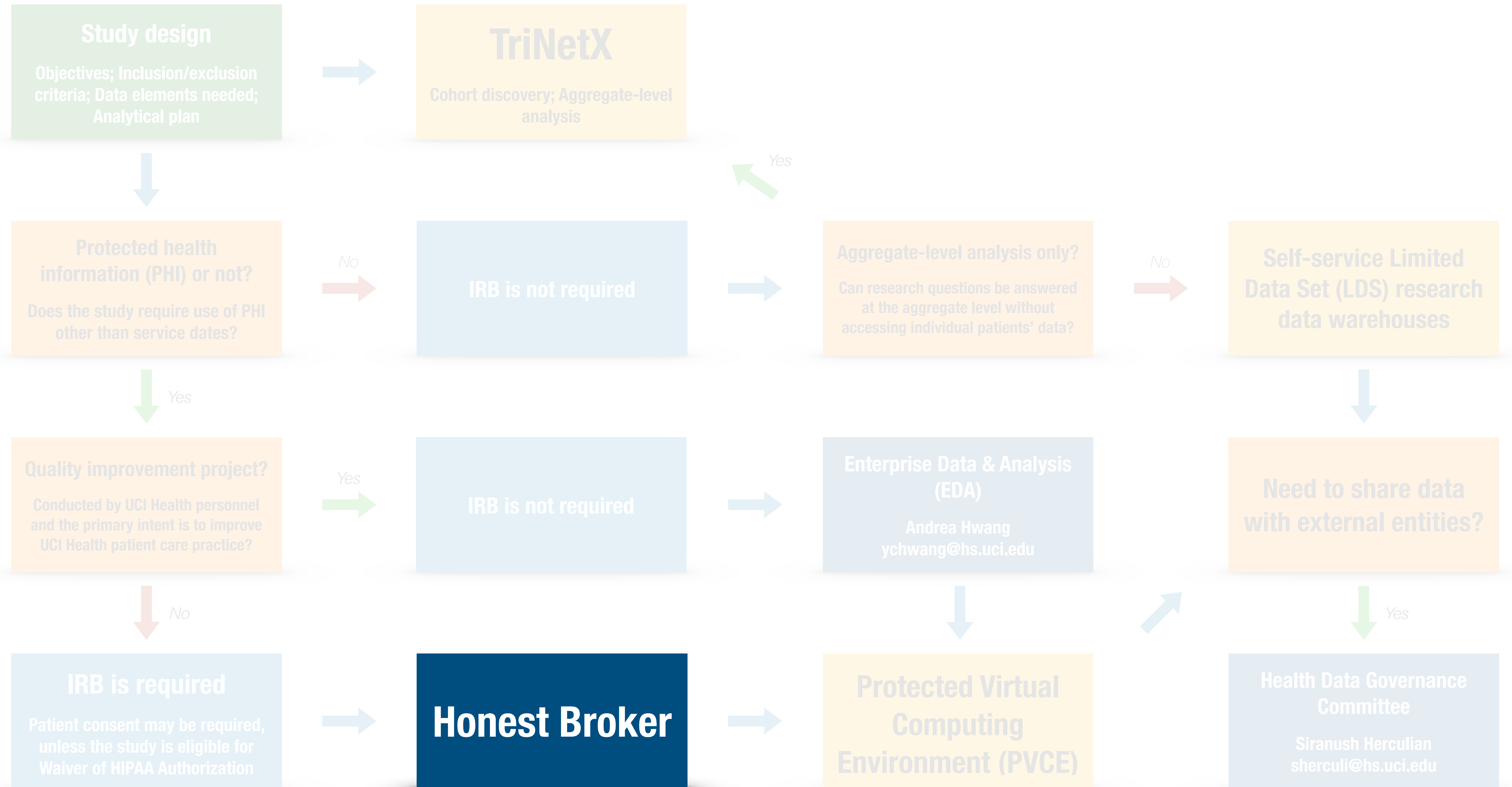
★ Final

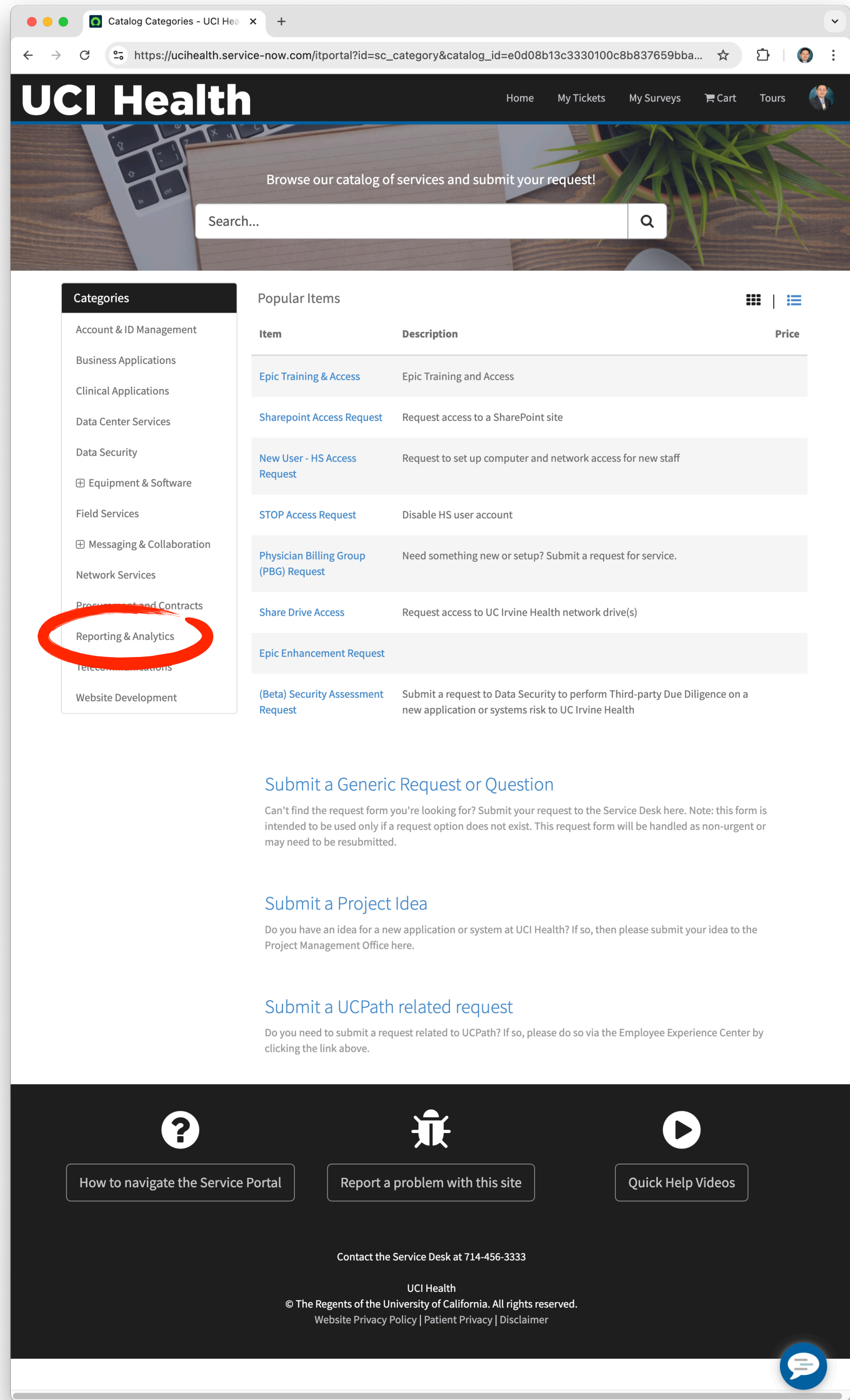
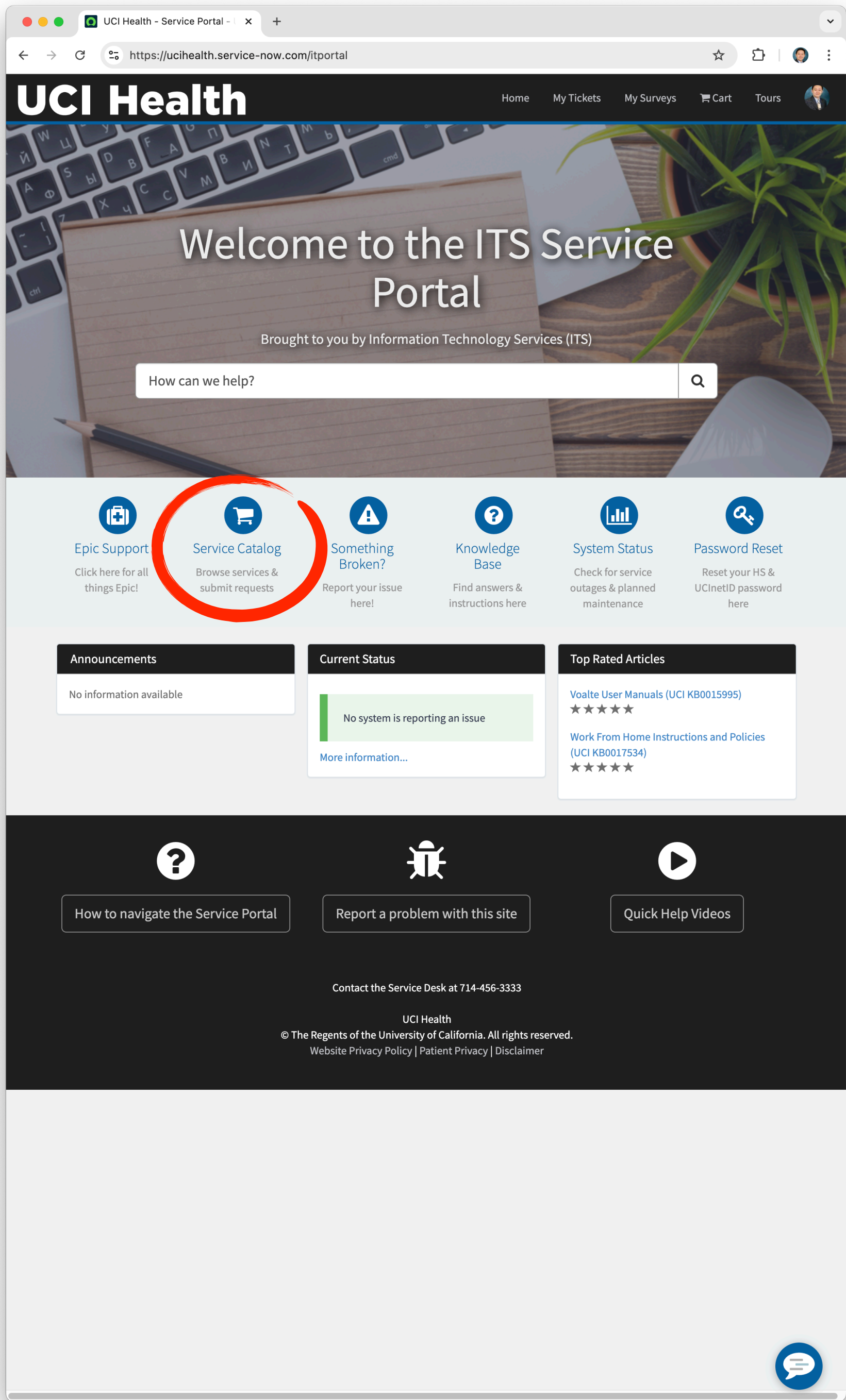
Patients 2,820

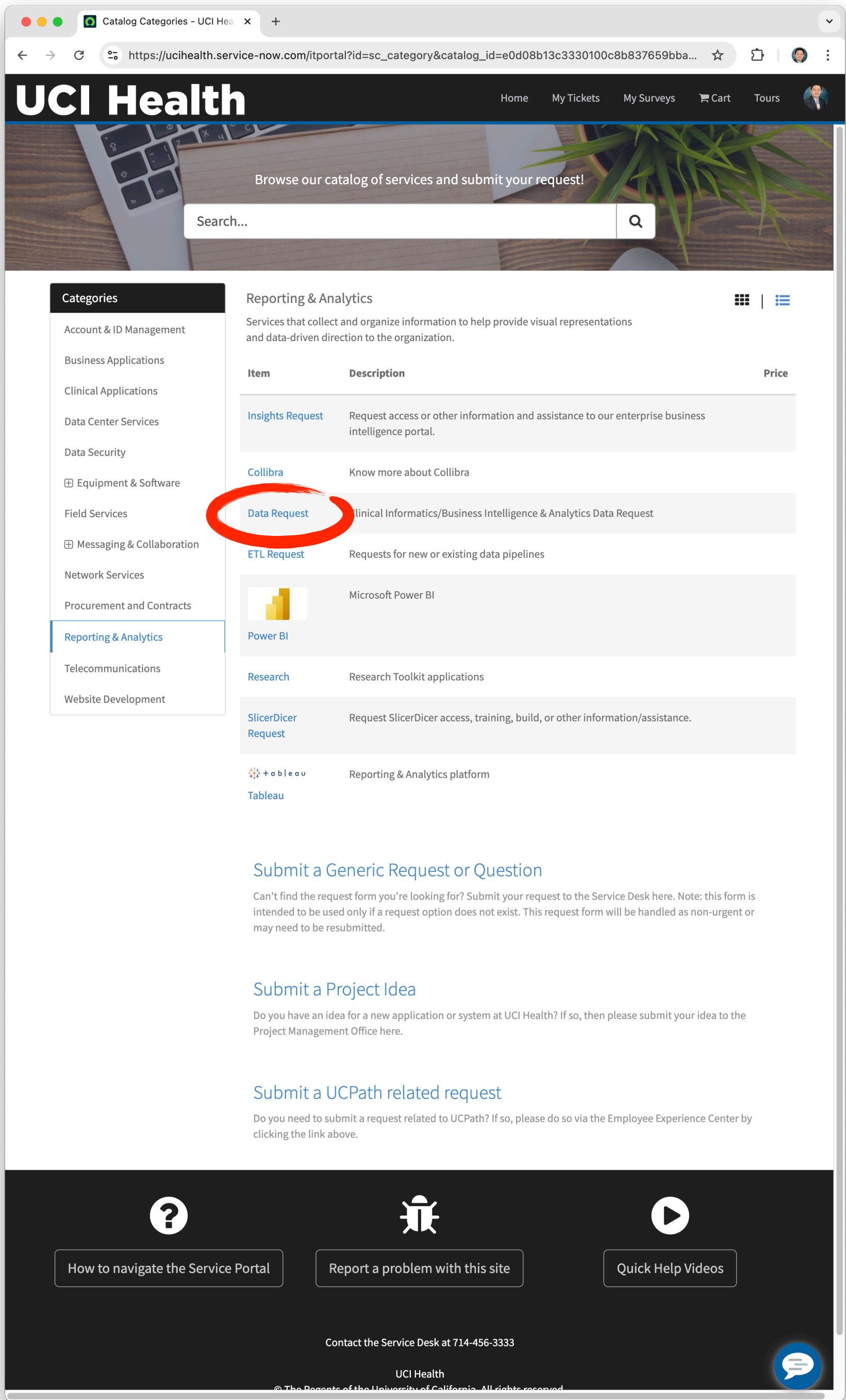
HCOs 1

Request TriNetX Dataset

Free to all UCI researchers unless you have a budget line in your grant for purchasing the data. Data fulfillment usually takes about 2 weeks.







UCI Health

Home All Catalogs Service Catalog Reporting & Analytics Data Request

Search

Data Request

Clinical Informatics/Business Intelligence & Analytics Data Request

By submitting this form you are requesting a report from the Epic EMR System
DO NOT use this request to submit urgent Patient Care issues; please contact the Service Desk.
For research data requests, please complete and attach the following worksheet: [Honest Broker Data Request Worksheet](#)

Order Now

Required information

- Department
- Requester supervisor (Manager or Director)
- Please describe the data request
- Is this a research request?

*Who is this for? Kai Zheng

* Callback number 949.824.6920

* Department

* Job title Professor

* Requester supervisor (Manager or Director)

* Please describe the data request

* Is this a research request?

* Did this data request originate based on the outcome of an adverse event?

* What is the problem that you are trying to solve?

* How will this data solve your problem?

* Risk/Value to organization

* Are there committed deadlines for delivering the information requested?

* Frequency/format frequency need

* Does the data need to be real time?

Intended report users/audience

Additional information

Add attachments

UCI Health Honest Broker Data Request Worksheet

Office of Research Information, UCI Health Affairs. Questions: please email ori@uci.edu.

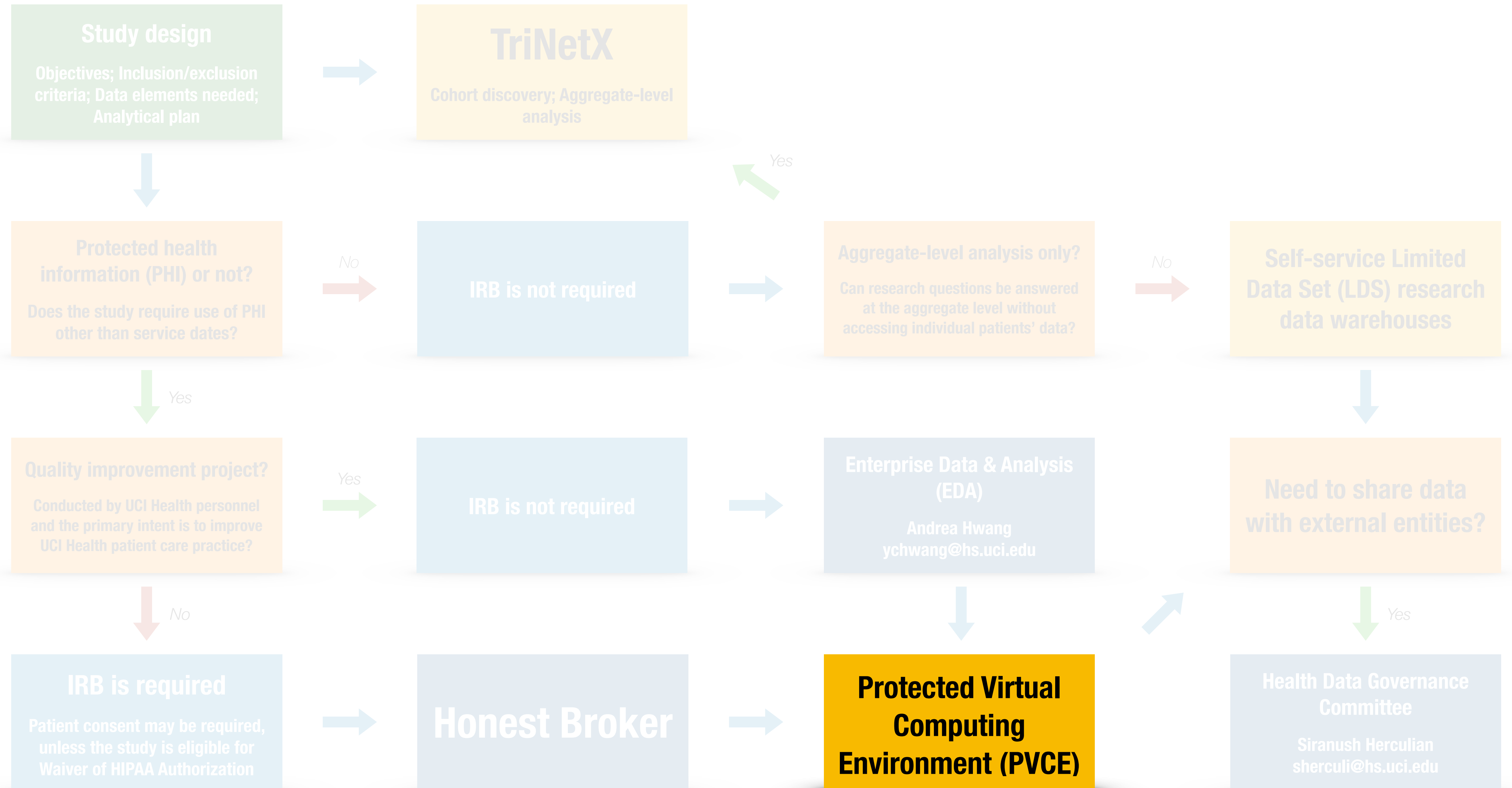
How to search for patients?

Timeframe(s) of the search

What types of data should be retrieved?

Timeframe(s) of the data to be retrieved

I. Research Team	
Title of the research project	
IRB #	
Name/email of primary project contact(s)	
Name/email of the Principal Investigator(s)	
Name/email of other IRB-authorized research personnel that should be provided access to the data	
II. Data Request Specifications	
For diagnoses, procedures, and laboratory tests, please provide standard codes whenever possible, such as ICD-9/10, CPT, LONIC, and SNOMED CT. If you don't know the code for a medical concept, please look it up at https://athena.ohdsi.org/ .	
A. Inclusion/Exclusion Criteria	
Please include timeframes and the data element that each timeframe should be applied on; see example. Please ensure consistency with the Eligibility Factors (Inclusion/Exclusion Criteria) as specified in the IRB protocol.	
☐ Check this box if the timeframes are listed in the Eligibility Section of the IRB protocol. Or <Type here>	Example: 1. Birth sex: female 2. Age: >= 18 years old as of 2020-01-01 3. The result of the most recent SARS-CoV-2 test (of any type), performed between 2020-01-01 and 2020-12-31, was positive 4. With any subclassification of type 2 diabetes (E11*) diagnosed at any date prior to 2020-01-01 5. Did <u>NOT</u> have an ED visit between 2020-01-01 and 2020-12-31
B. Data Elements to be Included	
Please include timeframes and the data element that each timeframe should be applied on; see example. Please ensure consistency with the Variables and/or Data Abstraction Sheet as specified in the IRB protocol.	
Indicate whether the timeframes are listed in the IRB protocol (check all that apply): ☐ Pre-screening Section ☐ Project Procedures Section ☐ Data Abstraction Sheet or Case Report Form ☐ Master Protocol Or <Type here>	Example: 1. MRN 2. Phone number(s) and Email address(es) 3. PCP name and email 4. Date and primary discharge diagnosis of all hospitalized admissions between 2020-01-01 and 2020-12-31 5. Date of the most recent ambulatory care visit of any time period



[illegible]

My Apps

- Recycle Bin
- Analytics Software
- Sign Out
- My Apps

Analytics Software

File Home Share View

Analytics Software

Name	Date modified	Type	Size
Access	9/10/2022 11:28 AM	Shortcut	3 KB
Excel	6/30/2021 9:13 AM	Shortcut	2 KB
IBM SPSS Statistics	2/24/2023 7:06 PM	Shortcut	3 KB
jupyter-notebook	1/18/2023 2:24 PM	Shortcut	2 KB
Microsoft SQL Server Management Studio 18	6/30/2021 4:25 PM	Shortcut	2 KB
R 4.2.2	1/5/2023 11:03 PM	Shortcut	2 KB
RStudio	10/10/2023 10:45 ...	Shortcut	1 KB
SAS 9.4 (English)	6/29/2021 11:44 PM	Shortcut	2 KB
StataMP 17 (64-bit)	9/10/2022 10:57 AM	Shortcut	2 KB
StataMP 18 (64-bit)	10/4/2023 9:44 AM	Shortcut	2 KB
Visual Studio Code	9/10/2022 12:00 PM	Shortcut	2 KB
Word	9/10/2022 11:27 AM	Shortcut	3 KB

12 items

UCI Health

SERVICE DESK

External - (714) 456-3333 | 24/7
Internal - x4563333 | 24/7

For non-urgent incidents and requests, please visit us at:
servicenow.hs.uci.edu

- Recycle Bin
- Analytics Software
- Sign Out
- My Apps

UCI Health

SERVICE DESK

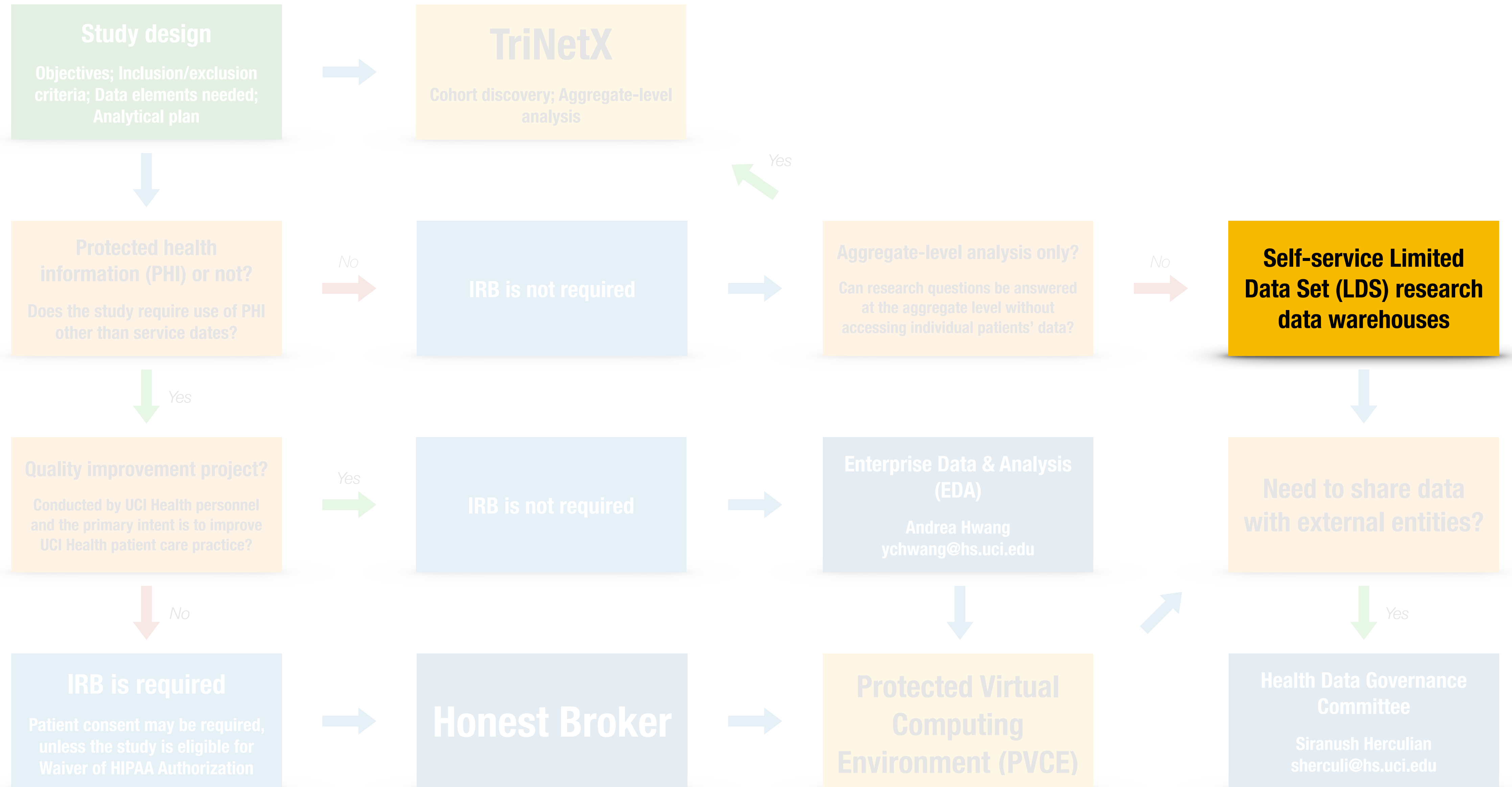
External - (714) 456-3333 | 24/7
Internal - x4563333 | 24/7

For non-urgent incidents and requests, please visit us at:
servicenow.hs.uci.edu

group (\omop.uci.edu\pvce_staging) (G:)

Name	Date modified	Type	Size
UCI IRB #1919	10/10/2023 11:17 ...	File folder	
UCI IRB #2002	10/10/2023 11:17 ...	File folder	
UCI IRB #2011	10/10/2023 11:17 ...	File folder	
UCI IRB #2039	10/10/2023 11:17 ...	File folder	
UCI IRB #2098	10/10/2023 11:17 ...	File folder	
UCI IRB #2126	10/10/2023 11:17 ...	File folder	
UCI IRB #2254	10/10/2023 11:17 ...	File folder	
UCI IRB #2271	10/10/2023 11:17 ...	File folder	
UCI IRB #2283	10/10/2023 11:17 ...	File folder	
UCI IRB #2326	10/10/2023 11:17 ...	File folder	
UCI IRB #2331	10/10/2023 11:17 ...	File folder	
UCI IRB #2349	10/10/2023 11:17 ...	File folder	
UCI IRB #2393	10/10/2023 11:17 ...	File folder	
UCI IRB #2487	10/10/2023 11:17 ...	File folder	
UCI IRB #2514	10/10/2023 11:17 ...	File folder	
UCI IRB #2636	10/10/2023 11:17 ...	File folder	
UCI IRB #2653	10/10/2023 11:17 ...	File folder	
UCI IRB #2708	10/10/2023 11:17 ...	File folder	
UCI IRB #2723	10/10/2023 11:17 ...	File folder	
UCI IRB #2764	10/10/2023 11:17 ...	File folder	
UCI IRB #2814	10/10/2023 11:17 ...	File folder	
UCI IRB #2908	10/10/2023 11:17 ...	File folder	
UCI IRB #3439	10/10/2023 11:17 ...	File folder	
UCI IRB #3501	10/10/2023 11:17 ...	File folder	

168 items



SQLQuery1.sql - omop.uci.edu.LDS_OMOP_RDW (HS\zhengkai (62)) - Microsoft SQL Server Management Studio (Administrator)

File Edit View Query Project Tools Window Help

LDS_OMOP_RDW

Execute

Object Explorer

Connect

omop.uci.edu (SQL Server 15.0.4249.2 - HS\zhengkai)

Databases

System Databases

Database Snapshots

AdminDB

LDS_OMOP_RDW

Database Diagrams

Tables

Views

External Resources

Synonyms

Programmability

Service Broker

Storage

Security

LDS_OMOP_RDW_EZ

LOGS_CBMI

OHDSI

UC_CORDS

Security

Server Objects

Replication

PolyBase

Always On High Availability

Management

Integration Services Catalogs

SQL Server Agent

XEvent Profiler

SQLQuery1.sql - om... (HS\zhengkai (62))

-- select the the database to use

use LDS_OMOP_RDW

-- drop temp table if it already exists

DROP TABLE IF EXISTS #patient_list

-- create list of patients with prostate cancer

SELECT DISTINCT person_id INTO #patient_list

FROM condition_occurrence

WHERE condition_concept_id = 200962

-- get age, sex, race, ethnicity

-- LDS OMOP only has year of birth. July 1st is appended as the middle point of the year to estimate age as of today

SELECT p.person_id, DATEDIFF(day, CONVERT(date, CONVERT(varchar, p.year_of_birth) + '-07-01'), SYSDATETIME(0)) AS age, sex.concept_name AS sex, race.concept_name AS race, ethnicity.concept_name AS ethnicity FROM person p

LEFT JOIN concept AS sex ON sex.concept_id = p.gender_concept_id

LEFT JOIN concept AS race ON race.concept_id = p.race_concept_id

LEFT JOIN concept AS ethnicity ON ethnicity.concept_id = p.ethnicity_concept_id

WHERE person_id IN (SELECT person_id FROM #patient_list)

-- get main language spoken (4267355)

-- we do not have valid data in observation table on gender (4135376) or race (4153313) sufficient

SELECT * FROM observation o

LEFT JOIN concept c_concept_id ON c_concept_id.concept_id = o.observation_concept_id

LEFT JOIN concept c_value ON c_value.concept_id = o.value_as_concept_id

WHERE person_id IN (SELECT person_id FROM #patient_list)

AND o.observation_concept_id = 4267355

100 %

Results Messages

	person_id	age	sex	race	ethnicity
1	329465	53.284052	MALE	Multirace	Not Hispanic or Latino
2	946732	74.283367	MALE	Multirace	Not Hispanic or Latino
3	1238346	64.284736	MALE	Multirace	Not Hispanic or Latino
4	251642	78.283367	MALE	Multirace	Not Hispanic or Latino
5	398696	66.283367	MALE	Multirace	Not Hispanic or Latino
6	1537738	87.282683	MALE	Multirace	Not Hispanic or Latino
7	241410	87.282683	MALE	American Indian or Alaska Native	Not Hispanic or Latino
8	849401	74.283367	MALE	American Indian or Alaska Native	Not Hispanic or Latino
9	1624443	78.283367	MALE	Native Hawaiian or Other Pacific Islander	Not Hispanic or Latino
10	98516	83.282683	MALE	Native Hawaiian or Other Pacific Islander	Not Hispanic or Latino
11	1143487	67.282683	MALE	American Indian or Alaska Native	Not Hispanic or Latino
12	798205	70.283367	MALE	American Indian or Alaska Native	Not Hispanic or Latino
13	944920	51.282683	MALE	American Indian or Alaska Native	Not Hispanic or Latino

	observation_id	person_id	observation_concept_id	observation_date	observation_datetime	observation_type_concept_id	value_as_string	value_as_concept_id	qualifier_concept_id	unit_concept_id	provider_id	visit_occurrence_id	observation_source_value	observation_source_concept_id	unit_source_value	qualifier_source_value	visit_detail_id	concept_id
1	8444503	1812027	4267355	2011-11-11	2011-11-11 22:19:00.0000000	44814721	NULL	NULL	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
2	17671340	1413372	4267355	2023-09-25	2023-09-25 06:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
3	10651422	174841	4267355	2023-09-25	2023-09-25 06:49:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
4	8737720	1759156	4267355	2023-09-27	2023-09-27 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
5	6592264	226268	4267355	2023-10-02	2023-10-02 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
6	13683196	52092	4267355	2023-10-04	2023-10-04 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
7	11237677	221371	4267355	2022-02-02	2022-02-02 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
8	4389566	775433	4267355	2023-10-02	2023-10-02 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
9	17743843	396459	4267355	2023-10-04	2023-10-04 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
10	5524513	1337613	4267355	2019-07-12	2019-07-12 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
11	12426970	436711	4267355	2023-10-02	2023-10-02 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
12	3052805	828945	4267355	2020-01-27	2020-01-27 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
13	17698217	416299	4267355	2022-07-07	2022-07-07 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355

Query executed successfully.

omop.uci.edu (15.0 RTM) | HS\zhengkai (62) | LDS_OMOP_RDW | 00:00:08 | 20,052 rows

Ready

Ln 26 Col 39 Ch 39 INS

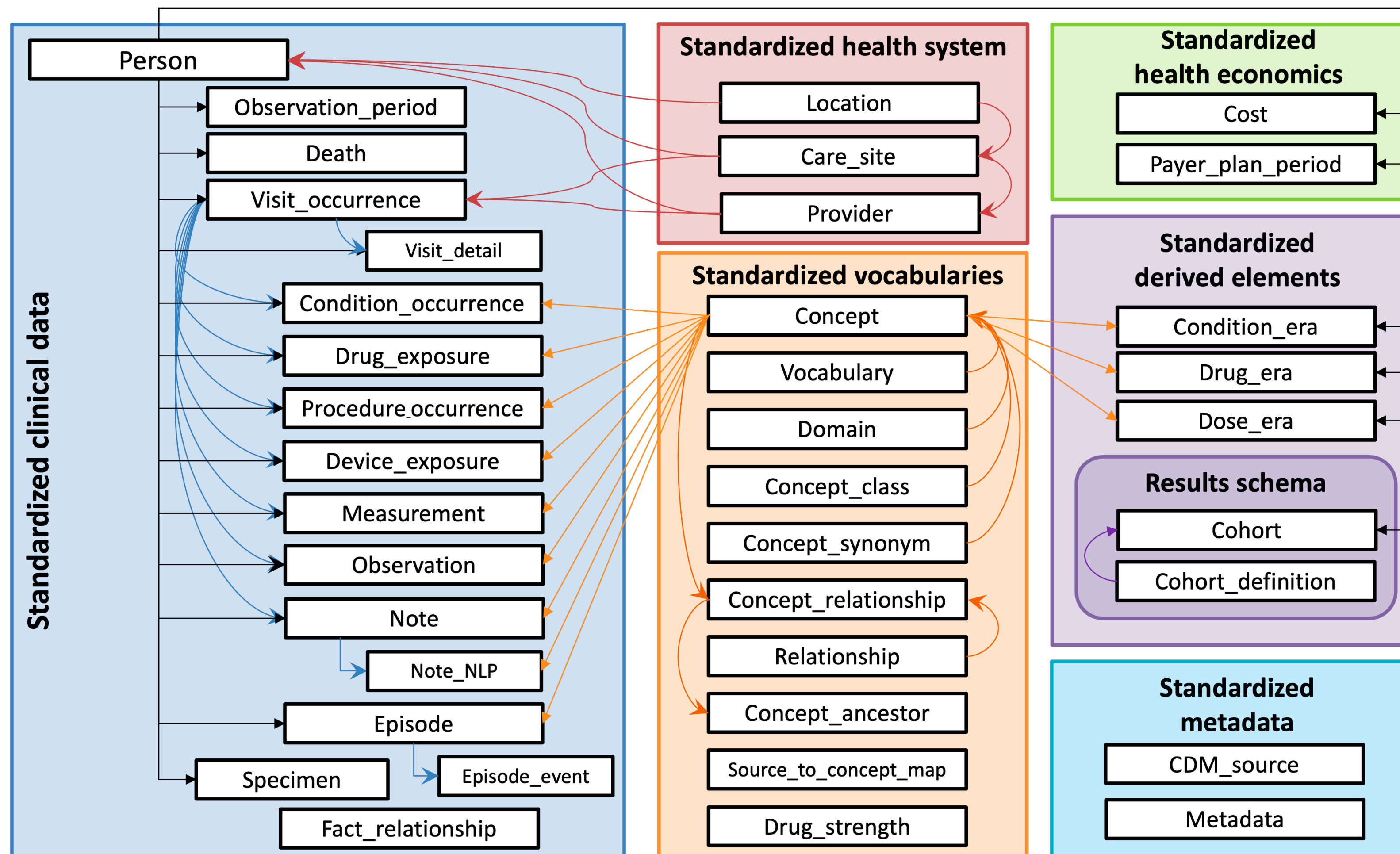
12:12 AM 10/13/2023

LDS OMOP

A Limited Data Set (LDS) version of the UCI Health medical records data for self-service use without IRB.

LDS: Limited Data Set; OMOP: Observational Medical Outcomes Partnership.

OMOP Common Data Model



SQLQuery1.sql - omop.uci.edu.LDS_OMOP_RDW (HS\zhengkai (62)) - Microsoft SQL Server Management Studio (Administrator)

File Edit View Query Project Tools Window Help

LDS_OMOP_RDW

Execute

Object Explorer

omop.uci.edu (SQL Server 15.0.4249.2 - HS\zhengkai)

Databases

System Databases

Database Snapshots

AdminDB

LDS_OMOP_RDW

Database Diagrams

Tables

Views

External Resources

Synonyms

Programmability

Service Broker

Storage

Security

LDS_OMOP_RDW_EZ

LOGS_CBM1

OHDSI

UC_CORDS

Security

Server Objects

Replication

PolyBase

Always On High Availability

Management

Integration Services Catalogs

SQL Server Agent

XEvent Profiler

SQLQuery1.sql - om... (HS\zhengkai (62))

-- select the the database to use
use LDS_OMOP_RDW

-- drop temp table if it already exists
DROP TABLE IF EXISTS #patient_list

-- create list of patients with prostate cancer
SELECT DISTINCT person_id INTO #patient_list
FROM condition_occurrence
WHERE condition_concept_id = 200962

-- get age, sex, race, ethnicity
-- LDS OMOP only has year of birth. July 1st is appended as the middle point of the year to estimate age as of today
SELECT p.person_id, DATEDIFF(day, CONVERT(date, CONVERT(varchar, p.year_of_birth) + '-07-01'), SYSDATETIME())/365.25 AS age, sex.concept_name AS sex, race.concept_name AS race, ethnicity.concept_name AS ethnicity FROM person p
LEFT JOIN concept AS sex ON sex.concept_id = p.gender_concept_id
LEFT JOIN concept AS race ON race.concept_id = p.race_concept_id
LEFT JOIN concept AS ethnicity ON ethnicity.concept_id = p.ethnicity_concept_id
WHERE person_id IN (SELECT person_id FROM #patient_list)

-- get main language spoken (4267355)
-- we do not have valid data in observation table on gender (4135376) or race (4013886), however the sex and race columns in person table above should be sufficient
SELECT * FROM observation o
LEFT JOIN concept c_concept_id ON c_concept_id.concept_id = o.observation_concept_id
LEFT JOIN concept c_value ON c_value.concept_id = o.value_as_concept_id
WHERE person_id IN (SELECT person_id FROM #patient_list)
AND o.observation_concept_id = 4267355

100 %

Results Messages

	person_id	age	sex	race	ethnicity
1	329465	53.284052	MALE	Multirace	Not Hispanic or Latino
2	946732	74.283367	MALE	Multirace	Not Hispanic or Latino
3	1238346	64.284736	MALE	Multirace	Not Hispanic or Latino
4	251642	78.283367	MALE	Multirace	Not Hispanic or Latino
5	398696	66.283367	MALE	Multirace	Not Hispanic or Latino
6	1537738	87.282683	MALE	Multirace	Not Hispanic or Latino
7	241410	87.282683	MALE	American Indian or Alaska Native	Not Hispanic or Latino
8	849401	74.283367	MALE	American Indian or Alaska Native	Not Hispanic or Latino
9	1624443	78.283367	MALE	Native Hawaiian or Other Pacific Islander	Not Hispanic or Latino
10	98516	83.282683	MALE	Native Hawaiian or Other Pacific Islander	Not Hispanic or Latino
11	1143487	67.282683	MALE	American Indian or Alaska Native	Not Hispanic or Latino
12	798205	70.283367	MALE	American Indian or Alaska Native	Not Hispanic or Latino
12	944920	51.282683	MALE	American Indian or Alaska Native	Not Hispanic or Latino

	observation_id	person_id	observation_concept_id	observation_date	observation_datetime	observation_type_concept_id	value_as_number	value_as_string	value_as_concept_id	qualifier_concept_id	unit_concept_id	provider_id	visit_occurrence_id	observation_source_value	observation_source_concept_id	unit_source_value	qualifier_source_value	visit_detail_id	concept_id
1	8444503	1812027	4267355	2011-11-11	2011-11-11 22:19:00.0000000	44814721	NULL	NULL	4151758	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
2	17671340	1413372	4267355	2023-09-25	2023-09-25 00:00:00.0000000	44814721	NULL	NULL	4151758	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
3	10651422	174841	4267355	2023-09-25	2023-09-25 06:49:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
4	8737720	1759156	4267355	2023-09-27	2023-09-27 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
5	6592264	226268	4267355	2023-10-02	2023-10-02 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
6	13683196	52092	4267355	2023-10-04	2023-10-04 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
7	11237677	221371	4267355	2022-02-02	2022-02-02 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
8	4389566	775433	4267355	2023-10-02	2023-10-02 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
9	17743843	396459	4267355	2023-10-04	2023-10-04 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
10	5524513	1337613	4267355	2019-07-12	2019-07-12 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
11	12426970	436711	4267355	2023-10-02	2023-10-02 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
12	3052805	828945	4267355	2020-01-27	2020-01-27 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
12	17688217	416299	4267355	2022-07-07	2022-07-07 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355

Query executed successfully.

omop.uci.edu (15.0 RTM) | HS\zhengkai (62) | LDS_OMOP_RDW | 00:00:08 | 20,052 rows

Ready Ln 26 Col 39 Ch 39 INS

12:12 AM 10/13/2023

SQLQuery1.sql - omop.uci.edu.LDS_OMOP_RDW (HS\zhengkai (62)) - Microsoft SQL Server Management Studio (Administrator)

File Edit View Query Project Tools Window Help

Object Explorer

Connect omop.uci.edu (SQL Server 15.0.4249.2 - HS\zhengkai)

Databases

System Databases

Database Snapshots

AdminDB

LDS_OMOP_RDW

Database Diagrams

Tables

Views

External Resources

Synonyms

Programmability

Service Broker

Storage

Security

LDS_OMOP_RDW_EZ

LOGS_CBM1

OHDSI

UC_CORDS

Security

Server Objects

Replication

PolyBase

Always On High Availability

Management

Integration Services Catalogs

SQL Server Agent

XEvent Profiler

SQLQuery1.sql - om... (HS\zhengkai (62))

-- select the the database to use
use LDS_OMOP_RDW

-- drop temp table if it already exists
DROP TABLE IF EXISTS #patient_list

-- create list of patients with prostate cancer
SELECT DISTINCT person_id INTO #patient_list
FROM condition_occurrence
WHERE condition_concept_id = 200962

-- get age, sex, race, ethnicity
-- LDS OMOP only has year of birth. July 1st is appended as the middle point of the year to estimate age as of today
SELECT p.person_id, DATEDIFF(day, CONVERT(date, CONVERT(varchar, p.year_of_birth) + '-07-01'), SYSDATETIME(0)) AS age, sex.concept_name AS sex, race.concept_name AS race, ethnicity.concept_name AS ethnicity FROM person p
LEFT JOIN concept AS sex ON sex.concept_id = p.gender_concept_id
LEFT JOIN concept AS race ON race.concept_id = p.race_concept_id
LEFT JOIN concept AS ethnicity ON ethnicity.concept_id = p.ethnicity_concept_id
WHERE person_id IN (SELECT person_id FROM #patient_list)

-- get main language spoken (4267355)
-- we do not have valid data in observation table on gender (4135376) or race (413800), however the sex and race columns in person table above should be sufficient
SELECT * FROM observation o
LEFT JOIN concept c_concept_id ON c_concept_id.concept_id = o.observation_concept_id
LEFT JOIN concept c_value ON c_value.concept_id = o.value_as_concept_id
WHERE person_id IN (SELECT person_id FROM #patient_list)
AND o.observation_concept_id = 4267355

100 %

Results Messages

	person_id	age	sex	race	ethnicity
1	329465	53.284052	MALE	Multirace	Not Hispanic or Latino
2	946732	74.283367	MALE	Multirace	Not Hispanic or Latino
3	1238346	64.284736	MALE	Multirace	Not Hispanic or Latino
4	251642	78.283367	MALE	Multirace	Not Hispanic or Latino
5	398696	66.283367	MALE	Multirace	Not Hispanic or Latino
6	1537738	87.282683	MALE	Multirace	Not Hispanic or Latino
7	241410	87.282683	MALE	American Indian or Alaska Native	Not Hispanic or Latino
8	849401	74.283367	MALE	American Indian or Alaska Native	Not Hispanic or Latino
9	1624443	78.283367	MALE	Native Hawaiian or Other Pacific Islander	Not Hispanic or Latino
10	98516	83.282683	MALE	Native Hawaiian or Other Pacific Islander	Not Hispanic or Latino
11	1143487	67.282683	MALE	American Indian or Alaska Native	Not Hispanic or Latino
12	798205	70.283367	MALE	American Indian or Alaska Native	Not Hispanic or Latino
13	944920	51.282683	MALE	American Indian or Alaska Native	Not Hispanic or Latino

	observation_id	person_id	observation_concept_id	observation_date	observation_datetime	observation_type_concept_id	value_as_number	value_as_string	value_as_datetime	qualifier_concept_id	unit_concept_id	provider_id	visit_occurrence_id	observation_source_value	observation_source_concept_id	unit_source_value	qualifier_source_value	visit_detail_id	concept_id
1	8444503	1812027	4267355	2011-11-11	2011-11-11 22:19:00.0000000	44814721	NULL	NULL	4151758	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
2	17671340	1413372	4267355	2023-09-25	2023-09-25 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
3	10651422	174841	4267355	2023-09-25	2023-09-25 06:49:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
4	8737720	1759156	4267355	2023-09-27	2023-09-27 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
5	6592264	226268	4267355	2023-10-02	2023-10-02 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
6	13683196	52092	4267355	2023-10-04	2023-10-04 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
7	11237677	221371	4267355	2022-02-02	2022-02-02 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
8	4389566	775433	4267355	2023-10-02	2023-10-02 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
9	17743843	396459	4267355	2023-10-04	2023-10-04 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
10	5524513	1337613	4267355	2019-07-12	2019-07-12 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
11	12426970	436711	4267355	2023-10-02	2023-10-02 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
12	3052805	828945	4267355	2020-01-27	2020-01-27 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
13	17698217	416299	4267355	2022-07-07	2022-07-07 00:00:00.0000000	44814721	NULL	NULL	4153313	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355

Query executed successfully.

omop.uci.edu (15.0 RTM) | HS\zhengkai (62) | LDS_OMOP_RDW | 00:00:08 | 20,052 rows

Ln 26 Col 39 Ch 39 INS

Ready

Windows Taskbar

12:12 AM 10/13/2023

LDS OMOP & UC CORDS

Access is free to all UCI employees and students. Email ori@uci.edu for an account.

Protected Virtual Computing Environment (PVCE)

What is PVCE?

PVCE is a HIPAA-compliant research computing environment for securely analyzing health data.

All datasets retrieved by the UCI Health Honest Broker based on IRB-approved research protocols are deposited in this environment.

Many commonly used analytical tools are readily available in PVCE. These include SAS, SPSS, Stata, R, Python, ArcGIS, and Excel. You may request new software, or new packages/libraries/procedures of the existing software, to be installed by sending an email to ori@uci.edu.

PVCE is also where researchers access the self-service research data warehouses, such as [LDS OMOP RDW](#), a partially de-identified research copy of the UCI Health patient care data.

How to request access

If have been granted access to the self-service research data warehouses such as [LDS OMOP RDW](#), you will be automatically granted access to PVCE.

If you want to use PVCE to access a dataset retrieved by the UCI Health Honest Broker, send an email to ori@uci.edu with the PI(s) of the study cc'd, if applicable.

The data folders for IRB-approved research projects are accessible to authorized users to PVCE, but cannot access a folder that you believe you should be able to access, send an email to ori@uci.edu with the PI(s) of the study cc'd, if applicable.

PREREQUISITES:

If you do not have a health system (HS) account, you will need to sign and attach the [UCI Health Confidentiality Agreement](#) in the email so we can request an HS account for you.

An HS account is needed because accessing the UCI Health network requires a separate login credential that is different from your general campus UCINetID account.

Note that for security reasons, your HS password may be forced to expire after a period of time. To reset it, use the [HS Self-Service Password Manager Portal](#).

Once your access is granted, you can connect to PVCE using the Microsoft Windows App. The software is generally built into Windows PCs. If not, you can obtain a copy [here](#). For macOS, you can download from the [Mac App Store](#).

Then, follow the [instructions](#) provided by the UCI Health Information Services to create a new remote desktop connection. The PC to connect to is **pvce5001.hs.uci.edu**.

As of November 11, 2024, all users are required to submit a one-time [Cisco Anyconnect VPN Access Request](#) prior to using HS Remote Desktop Gateway or HS VPN.

Troubleshooting for connection issues

If you get stuck at a "Please wait" blue screen, it is most likely because you have an inactive session that cannot be properly terminated by the server. If this happens, please email ori@uci.edu and we will manually log off your previous session.

If you cannot connect at all, and you are certain your access to PVCE has been granted, try the following steps:

1. Check in the [HS Self-Service Password Manager Portal](#) to see whether your HS password has expired. If so, reset it.

Self-Service Research Data Warehouses (RDW)

UCI researchers have access to several self-service research data warehouses. These include LDS OMOP RDW, a partially de-identified research copy of the UCI Health patient care data, and UCHDW, a UC-wide data warehouse that aggregates clinical data from all UC Health campuses. The use of these data warehouses is deemed by the IRB as non-human subjects research and thus does not require IRB approval, provided that the data are stored and analyzed in a secure environment.

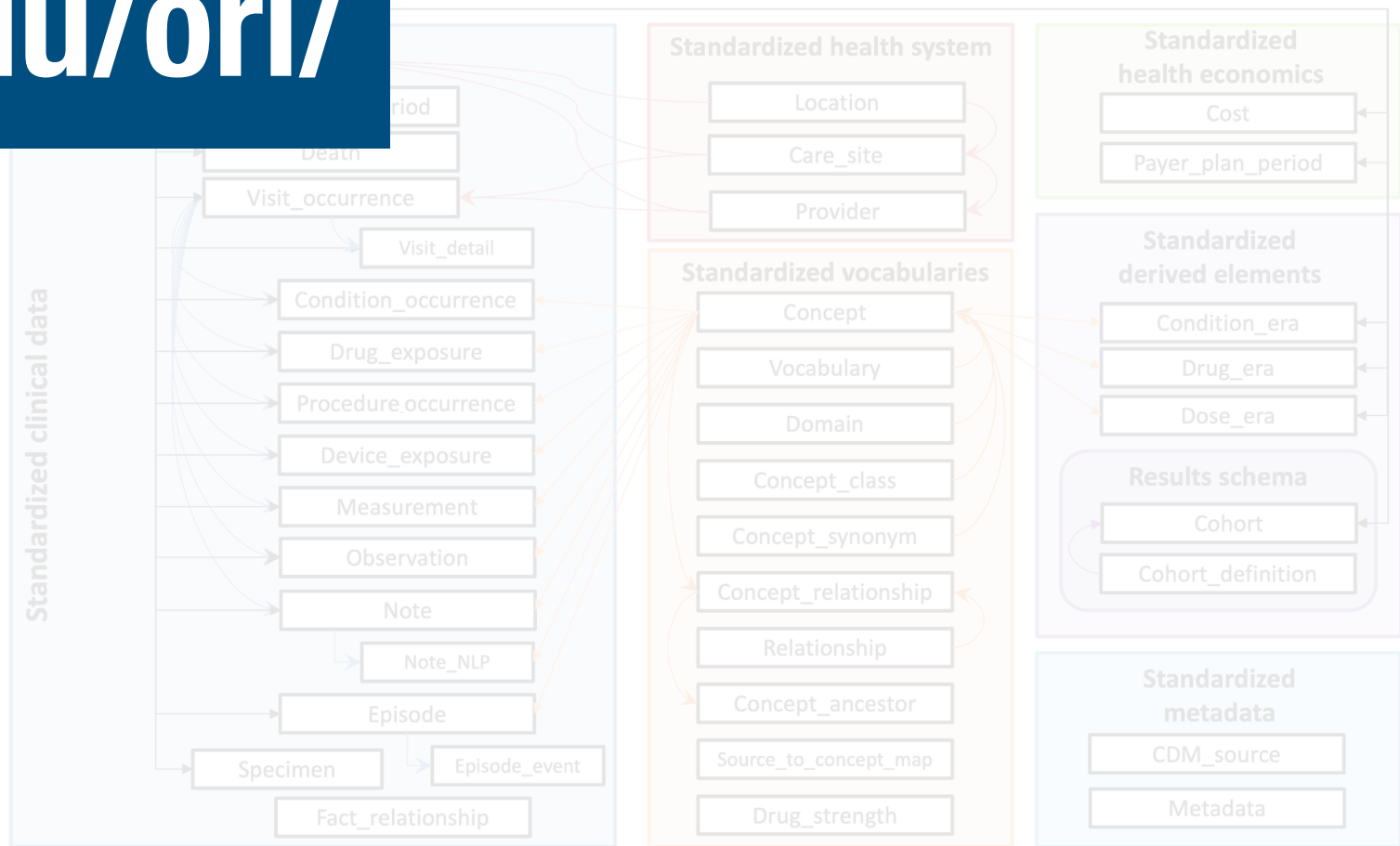
LDS OMOP RDW

LDS stands for Limited Data Set. LDS differs from fully anonymized data in that while a majority of patient identifiers such as names and phone numbers are removed, LDS retains certain data elements that HIPAA defines as patient identifiers, such as dates (e.g., admission date, discharge date, date of death).

In other words, LDS is only partially de-identified per HIPAA. However, because the likelihood of re-identifying patients from LDS is deemed reasonably low, it can be used for research and public health purposes without additional administrative or research ethnics approval, provided that adequate data protection procedures are in place.

Please note that while partial address information, such as city, state, five digit or more zip code, can be retained in an LDS, LDS OMOP RDW does not contain this information per UC policies.

OMOP standards for Observational Medical Outcomes Partnership. It is a common data model (CDM) that defines the structure (see figure below) in which data should be stored and the standard terminologies in which data are stored. OMOP standards are essential to ensure health data from different institutions are stored in a consistent manner across institutions.



How to request access

Any UCI employee or student can request access by sending an email to ori@uci.edu with the [LDS OMOP Data Use Agreement](#) signed and attached. For students, faculty supervisors' signature must be obtained.

Once access is granted, connect to [PVCE](#), and then launch Microsoft SQL Server Management Studio. If the Server Name field is not pre-populated, enter `omop.uci.edu`. Click Connect. LDS OMOP RDW can be now accessed on the Databases list to the left as `LDS_OMOP_RDW`.



Service Areas

1. **Honest broker for retrieving UCI Health patient data according to IRB-approved research protocols**
2. **Partially de-identified, self-service research data warehouses**
3. **Self-service cohort discovery tools**
4. **Advanced data analytics: artificial intelligence / machine learning (AI/ML), natural language processing (NLP)**
5. **Research data sharing networks**
6. **Research software development**
7. **Researcher engagement and enrichment**



Service Areas

1. Honest broker for retrieving UCI Health patient data according to IRB-approved research protocols
2. Partially de-identified, self-service research data warehouses
3. Self-service cohort discovery tools
4. **Advanced data analytics: artificial intelligence / machine learning (AI/ML), natural language processing (NLP)**
5. Research data sharing networks
6. Research software development
7. Researcher engagement and enrichment

Clamp Toolkit

New Project Run Tutorial A a

Component

- ML_components
- NLP_components

Pipeline

- MyPipeline
 - clamp-ner
 - Components
 - Data
 - Input
 - Output
 - sample_646.txt
 - sample_646.xmi
 - descriptor
 - PipelineLibrary

clamp-ner Data sample_646.xmi

1 Sample Type / Medical Specialty: Discharge Summary

2 Sample Name: Cardio/Pulmo Discharge Summary

3 Description: A 49-year-old man with **problem** respiratory distress, history of **problem** coronary artery disease with **problem** prior myocardial infarctions, and recently admitted with **problem** pneumonia and **problem** respiratory failure.

4 ADMISSION DIAGNOSIS:

5 1. **problem** Respiratory arrest.

6 2. **problem** End-stage chronic obstructive pulmonary disease.

7 3. **problem** Coronary artery disease.

8 4. **problem** History of hypertension.

9 DISCHARGE DIAGNOSIS:

10 1. **problem** Status post-respiratory arrest.

11 2. **problem** Chronic obstructive pulmonary disease.

12 3. **problem** Congestive heart failure.

13 4. **problem** History of coronary artery disease.

14 5. **problem** History of hypertension.

Semantic

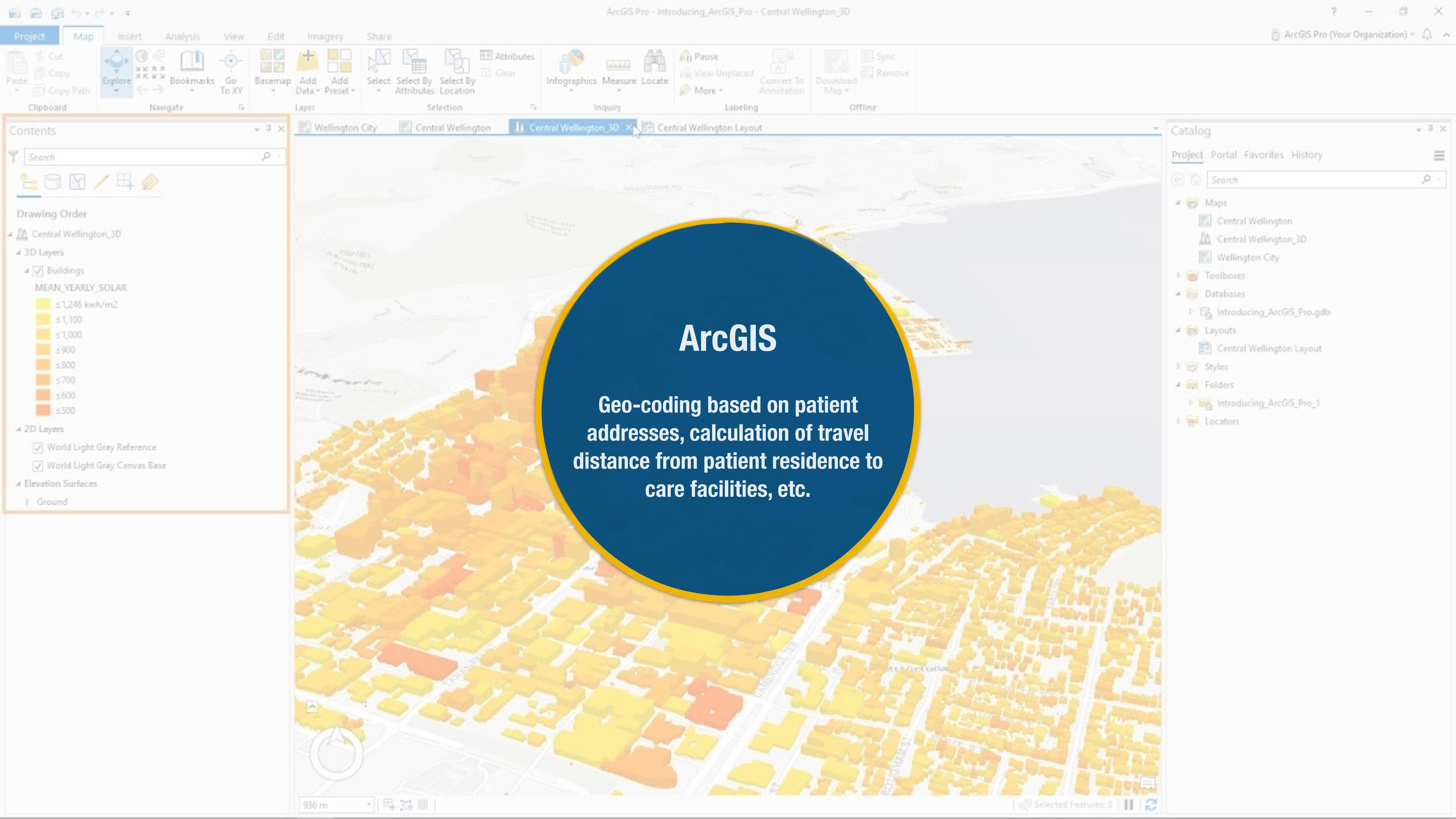
- ☒ Entity
- ☒ Relation
- ☐ Syntax

Console

/Users/kzheng/ClampMac_1.5.1/workspace/.log

Natural Language Processing / Large (and Small) Language Models / Clinical Notes De-Identification

Electronic Medical Record Search Engine (EMERSE)



ArcGIS

Geo-coding based on patient
addresses, calculation of travel
distance from patient residence to
care facilities, etc.



Service Areas

1. Honest broker for retrieving UCI Health patient data according to IRB-approved research protocols
2. Partially de-identified, self-service research data warehouses
3. Self-service cohort discovery tools
4. Advanced data analytics: artificial intelligence / machine learning (AI/ML), natural language processing (NLP)
5. **Research data sharing networks**
6. Research software development
7. Researcher engagement and enrichment



TRIAL DESIGN

DATA SETS & ANALYTICS

NON-US ONCOLOGY

RWE GENERATION

DRUG SAFETY

Explore. Discover. Collaborate.

TriNetX

A cohort discovery tool for constructing patient cohorts at UCI Health or globally, conducting aggregate-level analyses, or requesting de-identified patient-level data for research purposes.

Discover the power of TriNetX, the largest real-world data, as we revolutionize drug development and medical research. Beyond merely providing data, we offer end-to-end support throughout the drug development lifecycle. Our extensive research capabilities empower researchers to publish groundbreaking findings in prestigious medical journals while leveraging our platform technologies, integrated analytics, and downloadable data sets. From optimizing clinical trial designs to refining post-market safety measures, we offer comprehensive solutions that drive innovation and elevate healthcare standards worldwide. Join us in shaping a safer, healthier future, where every data point contributes to better patient outcomes and medical advancements.

Empowering Global Research: Leading the Way in Real-World Data for Comprehensive Drug Development

SQLQuery1.sql - omop.uci.edu.LDS_OMOP_RDW (HS\zhengkai (62)) - Microsoft SQL Server Management Studio (Administrator)

File Edit View Query Project Tools Window Help

LDS_OMOP_RDW

Execute

Object Explorer

Connect

omop.uci.edu (SQL Server 15.0.4249.2 - HS\zhengkai)

Databases

System Databases

Database Snapshots

AdminDB

LDS_OMOP_RDW

Database Diagrams

Tables

Views

External Resources

Synonyms

Programmability

Service Broker

Storage

Security

LDS_OMOP_RDW_EZ

LOGS_CBM1

OHDSI

UC_CORDS

Security

Server Objects

Replication

PolyBase

Always On High Availability

Management

Integration Services Catalogs

SQL Server Agent

XEvent Profiler

SQLQuery1.sql - om... (HS\zhengkai (62))

-- select the the database to use
use LDS_OMOP_RDW

-- drop temp table if it already exists
DROP TABLE IF EXISTS #patient_list

-- create list of patients with prostate cancer
SELECT DISTINCT person_id INTO #patient_list
FROM condition_occurrence
WHERE condition_concept_id = 200962

-- get age, sex, race, ethnicity
-- LDS OMOP only has year of birth. July 1st is appended as the middle point of the year to estimate age as of today
SELECT p.person_id, DATEDIFF(day, CONVERT(date, CONVERT(varchar, p.year_of_birth) + '-07-01'), SYSDATETIME(0)) AS age, sex.concept_name AS sex, race.concept_name AS race, ethnicity.concept_name AS ethnicity FROM person p
LEFT JOIN concept AS sex ON sex.concept_id = p.gender_concept_id
LEFT JOIN concept AS race ON race.concept_id = p.race_concept_id
LEFT JOIN concept AS ethnicity ON ethnicity.concept_id = p.ethnicity_concept_id
WHERE person_id IN (SELECT person_id FROM #patient_list)

-- get main language spoken (4267355)
-- we do not have valid data in observation table on gender (4135376) or race (4135376) sufficient
SELECT * FROM observation o
LEFT JOIN concept c_concept_id ON c_concept_id.concept_id = o.observation_concept_id
LEFT JOIN concept c_value ON c_value.concept_id = o.value_as_concept_id
WHERE person_id IN (SELECT person_id FROM #patient_list)
AND o.observation_concept_id = 4267355

100 %

Results Messages

	person_id	age	sex	race	ethnicity
1	329465	53.284052	MALE	Multirace	Not Hispanic or Latino
2	946732	74.283367	MALE	Multirace	Not Hispanic or Latino
3	1238346	64.284736	MALE	Multirace	Not Hispanic or Latino
4	251642	78.283367	MALE	Multirace	Not Hispanic or Latino
5	398696	66.283367	MALE	Multirace	Not Hispanic or Latino
6	1537738	87.282683	MALE	Multirace	Not Hispanic or Latino
7	241410	87.282683	MALE	American Indian or Alaska Native	Not Hispanic or Latino
8	849401	74.283367	MALE	American Indian or Alaska Native	Not Hispanic or Latino
9	1624443	78.283367	MALE	Native Hawaiian or Other Pacific Islander	Not Hispanic or Latino
10	98516	83.282683	MALE	Native Hawaiian or Other Pacific Islander	Not Hispanic or Latino
11	1143487	67.282683	MALE	American Indian or Alaska Native	Not Hispanic or Latino
12	798205	70.283367	MALE	American Indian or Alaska Native	Not Hispanic or Latino
13	944920	51.282683	MALE	American Indian or Alaska Native	Not Hispanic or Latino

	observation_id	person_id	observation_concept_id	observation_date	observation_datetime	observation_type_concept_id	observation_value	unit_concept_id	provider_id	visit_occurrence_id	observation_source_value	observation_source_concept_id	unit_source_value	qualifier_source_value	visit_detail_id	concept_id
1	8444503	1812027	4267355	2011-11-11	2011-11-11 22:19:00.0000000	44814721	NULL	NULL	NULL	-1	NULL	NULL	NULL	NULL	NULL	4267355
2	17671340	1413372	4267355	2023-09-25	2023-09-25 06:49:00.0000000	44814721	NULL	NULL	4153313	-1	NULL	NULL	NULL	NULL	NULL	4267355
3	10651422	174841	4267355	2023-09-25	2023-09-25 06:49:00.0000000	44814721	NULL	NULL	4153313	-1	NULL	NULL	NULL	NULL	NULL	4267355
4	8737720	1759156	4267355	2023-09-27	2023-09-27 00:00:00.0000000	44814721	NULL	NULL	4153313	-1	NULL	NULL	NULL	NULL	NULL	4267355
5	6592264	226268	4267355	2023-10-02	2023-10-02 00:00:00.0000000	44814721	NULL	NULL	4153313	-1	NULL	NULL	NULL	NULL	NULL	4267355
6	13683196	52092	4267355	2023-10-04	2023-10-04 00:00:00.0000000	44814721	NULL	NULL	4153313	-1	NULL	NULL	NULL	NULL	NULL	4267355
7	11237677	221371	4267355	2022-02-02	2022-02-02 00:00:00.0000000	44814721	NULL	NULL	4153313	-1	NULL	NULL	NULL	NULL	NULL	4267355
8	4389566	775433	4267355	2023-10-02	2023-10-02 00:00:00.0000000	44814721	NULL	NULL	4153313	-1	NULL	NULL	NULL	NULL	NULL	4267355
9	17743843	396459	4267355	2023-10-04	2023-10-04 00:00:00.0000000	44814721	NULL	NULL	4153313	-1	NULL	NULL	NULL	NULL	NULL	4267355
10	5524513	1337613	4267355	2019-07-12	2019-07-12 00:00:00.0000000	44814721	NULL	NULL	4153313	-1	NULL	NULL	NULL	NULL	NULL	4267355
11	12426970	436711	4267355	2023-10-02	2023-10-02 00:00:00.0000000	44814721	NULL	NULL	4153313	-1	NULL	NULL	NULL	NULL	NULL	4267355
12	3052805	828945	4267355	2020-01-27	2020-01-27 00:00:00.0000000	44814721	NULL	NULL	4153313	-1	NULL	NULL	NULL	NULL	NULL	4267355
13	17688217	416299	4267355	2022-07-07	2022-07-07 00:00:00.0000000	44814721	NULL	NULL	4153313	-1	NULL	NULL	NULL	NULL	NULL	4267355

Query executed successfully.

omop.uci.edu (15.0 RTM) | HS\zhengkai (62) | LDS_OMOP_RDW | 00:00:08 | 20,052 rows

Ready

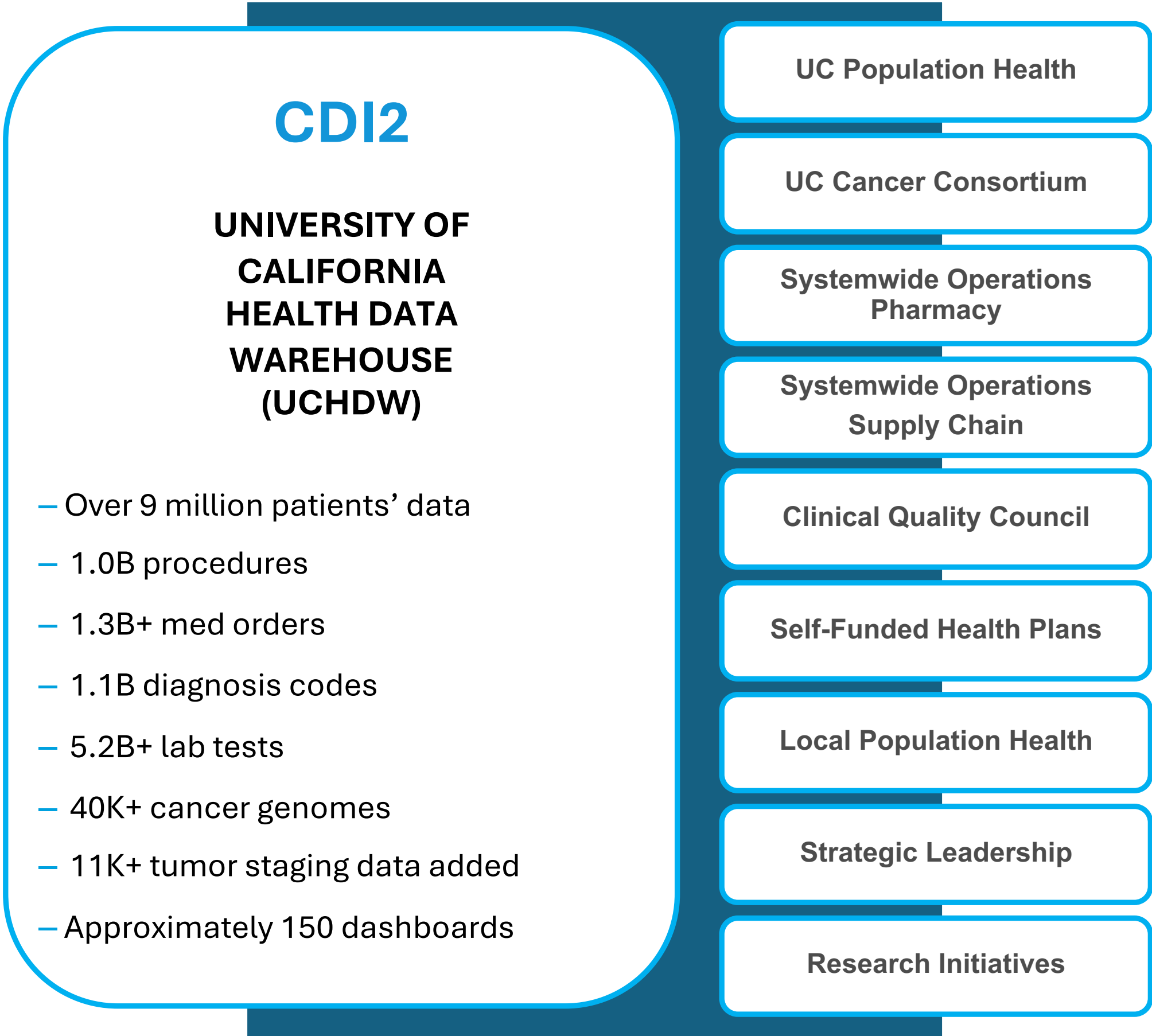
Ln 26 Col 39 Ch 39 INS

12:12 AM 10/13/2023

UC Health Data Warehouse (UCHDW)

Contains LDS OMOP data from all UCs. You must have used UCI LDS OMOP for a month before applying for an account (by sending an email to ori@uci.edu).

Robust Partnerships ... Dozens of Data-driven Projects



▼

Untitled Notebook 2024- x

Compute - Databricks x

location - Search - Databr x

+

←

→

↺

dbc-26b1f1de-b3df.cloud.databricks.com/compute?o=6458517280576964

☆

👤

⋮

☰ databricks

Q Search data, notebooks, recents, and more... CTRL + P

Prod 🛡️ ▼ ✨ Ⓢ

+ New

Workspace

Recents

Catalog

Workflows

Compute

SQL

SQL Editor

Queries

Dashboards

Alerts

Compute

All-purpose compute Job compute SQL warehouses Vector Search Pools Apps ⓘ

Q Filter compute you have access to

Created by ▼

☐ Only pinned

Create with Personal Compute ▼ ⓘ

State ⚙️	Name	Policy	Runtime	Active m...	Active co...	Active D...	Source	Creator	Notebooks	⚙️
📌	🟢									

▶ ▼

✓ Just now (3s)

1

```
%sql
select location_source_value as location, count(distinct p.person_id) as `No. of Patients`
from omop_prod.omop_deid.person p
join omop_prod.omop_deid.location l on l.location_id = p.location_id
group by l.location_source_value
```

▶ (4) Spark Jobs

▶ 📄 _sqldf: pyspark.sql.dataframe.DataFrame = [location: string, No. of Patients: long]

Table ▼ +

	📄 location	📄 No. of Patients
1	UCR	37797
2	UCSD	1316197
3	UCI	1117944
4	UCLA	3147235
5	UCSF	2419495
6	UCD	1408990

✦ Assistant

To find locations along with the number of distinct patients associated with each location from the `omop_prod.omop_deid.person` and `omop_prod.omop_deid.location` tables, you can use the following SQL query. This query joins the `person` table with the `location` table on the `location_id` column and groups the results by the `location_source_value` from the `location` table to count the distinct `person_id` values:

▶

```
%sql
SELECT
  l.location_source_value AS
  location,
  COUNT(DISTINCT p.person_id) AS `No.
  of Patients`
FROM
  omop_prod.omop_deid.person p
JOIN
  omop_prod.omop_deid.location l
  ON p.location_id = l.location_id
GROUP BY
  l.location_source_value
```

This query will give you a list of locations and the number of unique patients at each location.

on - Search - Databr x | +

76964

☆ | 👤 | ⋮

notebooks, recents, and more... CTRL + P

Prod 🛡️ | ✨ | Ⓢ

Vector Search Pools Apps ⓘ

🔍

☐ Only pinned

Create with Personal Compute ⌵ ⓘ

Policy	Runtime	Active m...	Active co...	Active D...	Source	Creator	Notebooks	⚙️
	15.1	48 GB	6 cores	4.14	UI	👤 David Gonza...	1	■ ⋮

as `No. of Patients`

✦ Assistant

To find locations along with the number of distinct patients associated with each location from the `omop_prod.omop_deid.person` and `omop_prod.omop_deid.location` tables, you can use the following SQL query. This query joins the `person` table with the `location` table on the `location_id` column and groups the results by the `location_source_value` from the `location` table to count the distinct `person_id` values:

```
%sql
SELECT
  l.location_source_value AS
  location,
  COUNT(DISTINCT p.person_id) AS `No.
  of Patients`
FROM
  omop_prod.omop_deid.person p
JOIN
  omop_prod.omop_deid.location l
  ON p.location_id = l.location_id
GROUP BY
  l.location_source_value
```

This query will give you a list of locations and the number of unique patients at each location.

LDS UCHDW

- Non-human subjects research determination by UC IRBs
- Currently contains only structured data
- De-identified free-text notes are being added and are expected to be available for research reuse in late 2026



Cosmos is a dataset created in collaboration with a community of health systems using Epic and is designed to improve patient care. By combining their data, participating organizations and Epic can make new discoveries and advance medicine. Cosmos also powers tools at the point of care, including clinical decision support systems that are tailored to the

Epic Cosmos

A dataset created by a community of health systems using Epic EHR system to “make new discoveries, advance medicine, and powers tools at the point of care.”

Cosmos Community



1,883
Hospitals



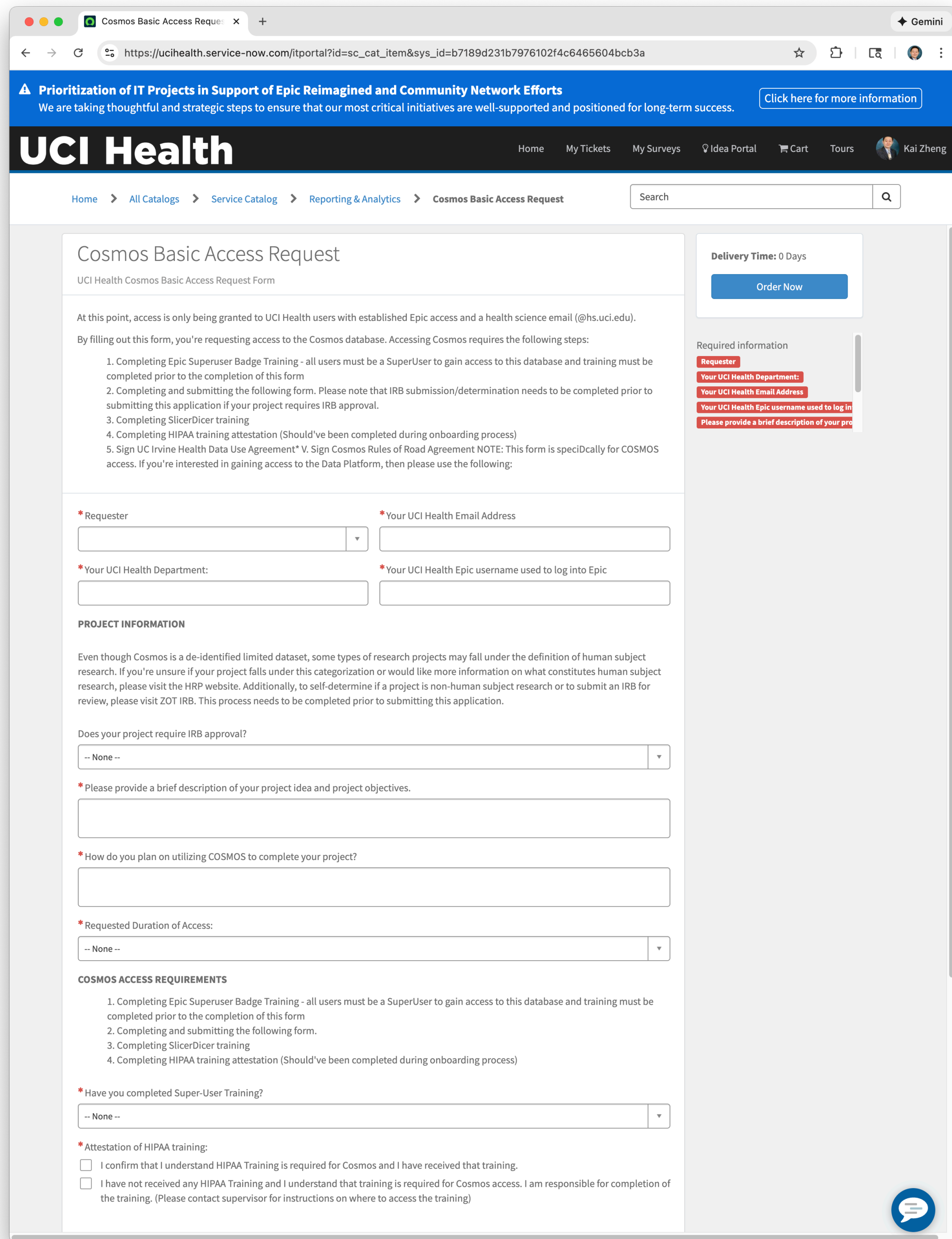
300 Million Patients



42.4K
Clinics

Meet a few of our Cosmos community members:





<https://reach.icts.uci.edu/ori/>



From promise to progress. The future of health research is now.

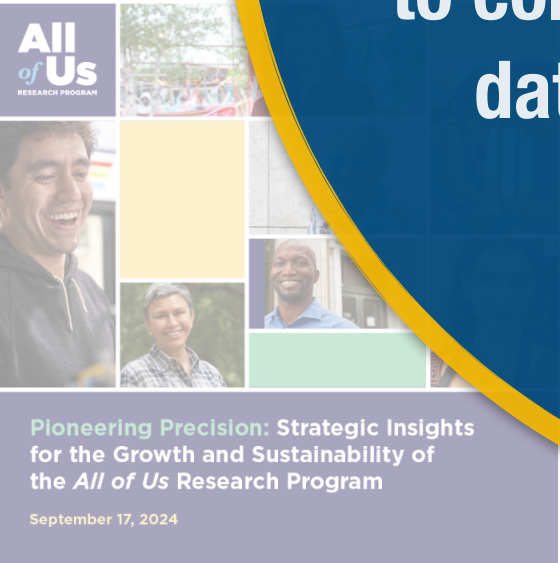
The *All of Us* Research Program is accelerating research to advance and precision medicine.

Our nationwide community of participants and researchers are partnering together to ensure that everyone is included in research. The information on this site is for informational purposes only, fueling thousands of studies that better understand health and disease and enable more personalized and equitable approaches to care.

JOIN NOW

All of Us

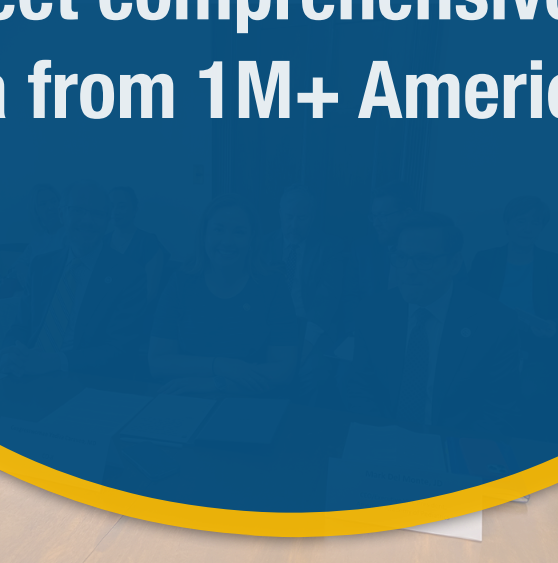
Largest cohort on the planet aiming to collect comprehensive health data from 1M+ Americans.



Pioneering Precision: Strategic Insights for the Growth and Sustainability of the *All of Us* Research Program
September 17, 2024

All of Us’ Advisory Panel Releases Report on the Future of the Program

A working group of the National Institutes of Health’s *All of Us* Research Program’s Advisory Panel has released a new report emphasizing the program’s successes and challenges in building a medical research platform that reflects the diversity of the U.S.



NIH Leaders Join Congressional Roundtable on Advancing Pediatric Medical Research

National Institutes of Health (NIH) leaders recently joined a congressional roundtable on Capitol Hill to discuss the importance of pediatric research to better understand the drivers of disease across the lifespan.



NIH Invites Experts and Leaders to Help Advance Environmental Justice

The National Institutes of Health (NIH) is initiating an Environmental Justice Scholars Program to advance knowledge on environmental factors impacting health disparities.

Interested in the *All of Us* Research Program?

GET INVOLVED

Diversity, transparency, and accessibility are core values that guide *All of Us*.

VIEW OUR CORE VALUES

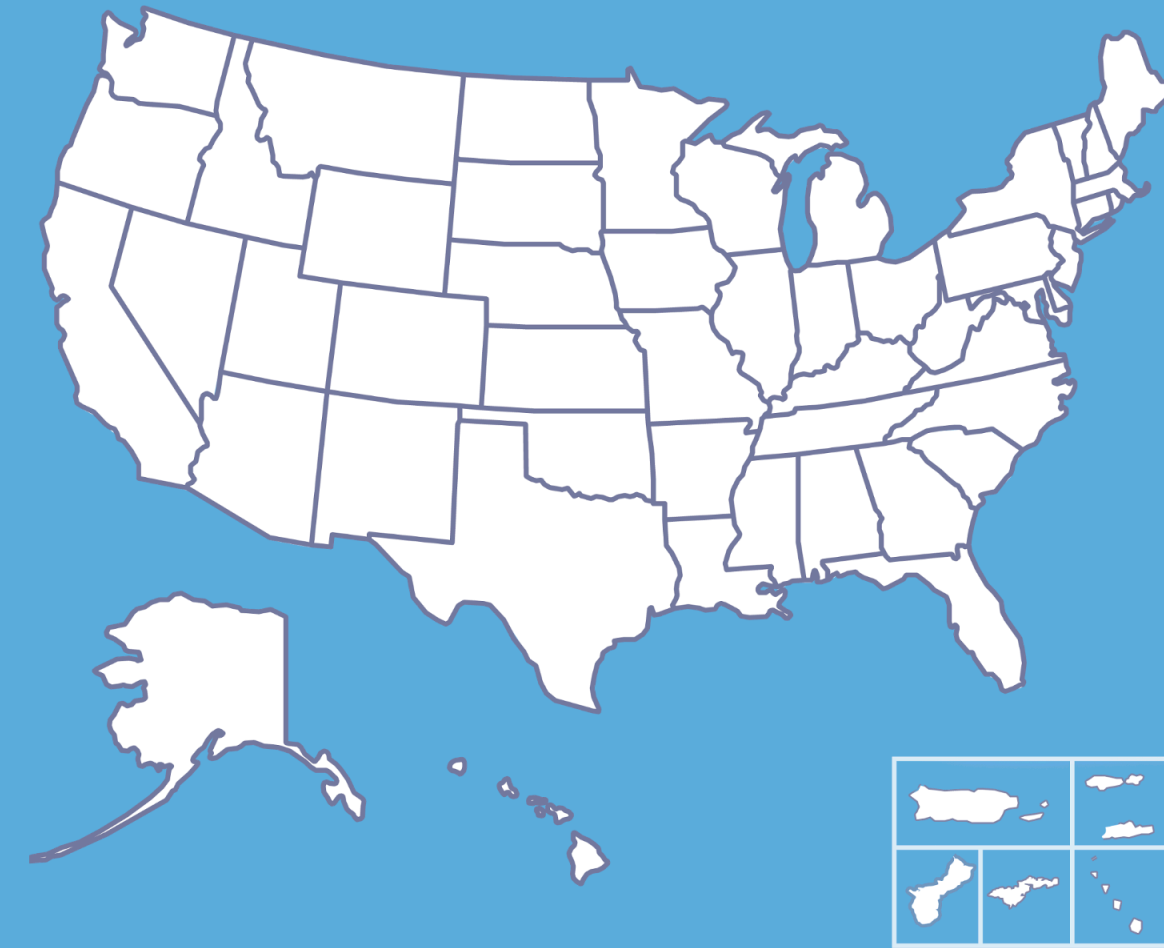
Sign up to be notified of announcements, events, funding news and more.

SUBSCRIBE

One of the world's largest, most diverse biomedical datasets of its kind

Inviting

1 Million or more people
from across
the United States



Data available from **873,000+** participants

75%

identify with
communities
underrepresented in
biomedical research



48%

are from racial
and ethnic
minority groups



Data types and sources

The *All of Us* Research Program’s Data and Research Center (DRC) curates a range of different data types as part of the data collection process. The numbers below reflect the number of participants with each data type available.



413,350+
Survey
Responses



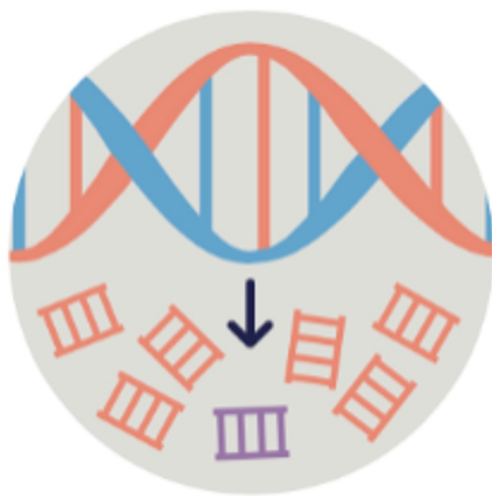
337,500+
Physical
Measurements



312,900+
Genotyping
Arrays



287,000+
Electronic Health
Records



245,350+
Whole Genome
Sequences



15,600+
Fitbit
Records
NEW! Device Data



97,900+
With structural variant data (More
than 1.5 million total structural
variants)



1,000+
Long-Read Sequences

Data as of April 2023 Curated Data Repository (CDR) v7 data release and CDRv7 off-cycle release in April and June 2024

Data access at ResearchAllofUs.org

Three Tiers



Public Tier

The Public Tier dataset is available at ResearchAllofUs.org. Visitors can explore aggregated overviews and interactive data previews—with participant identifiers removed—through the Data Browser. Registration is not required.



Registered Tier

Includes individual-level data from EHRs, wearables, and surveys, as well as physical measurements taken at the time of participant enrollment. Registration and a Responsible Conduct of Research training are required for access.

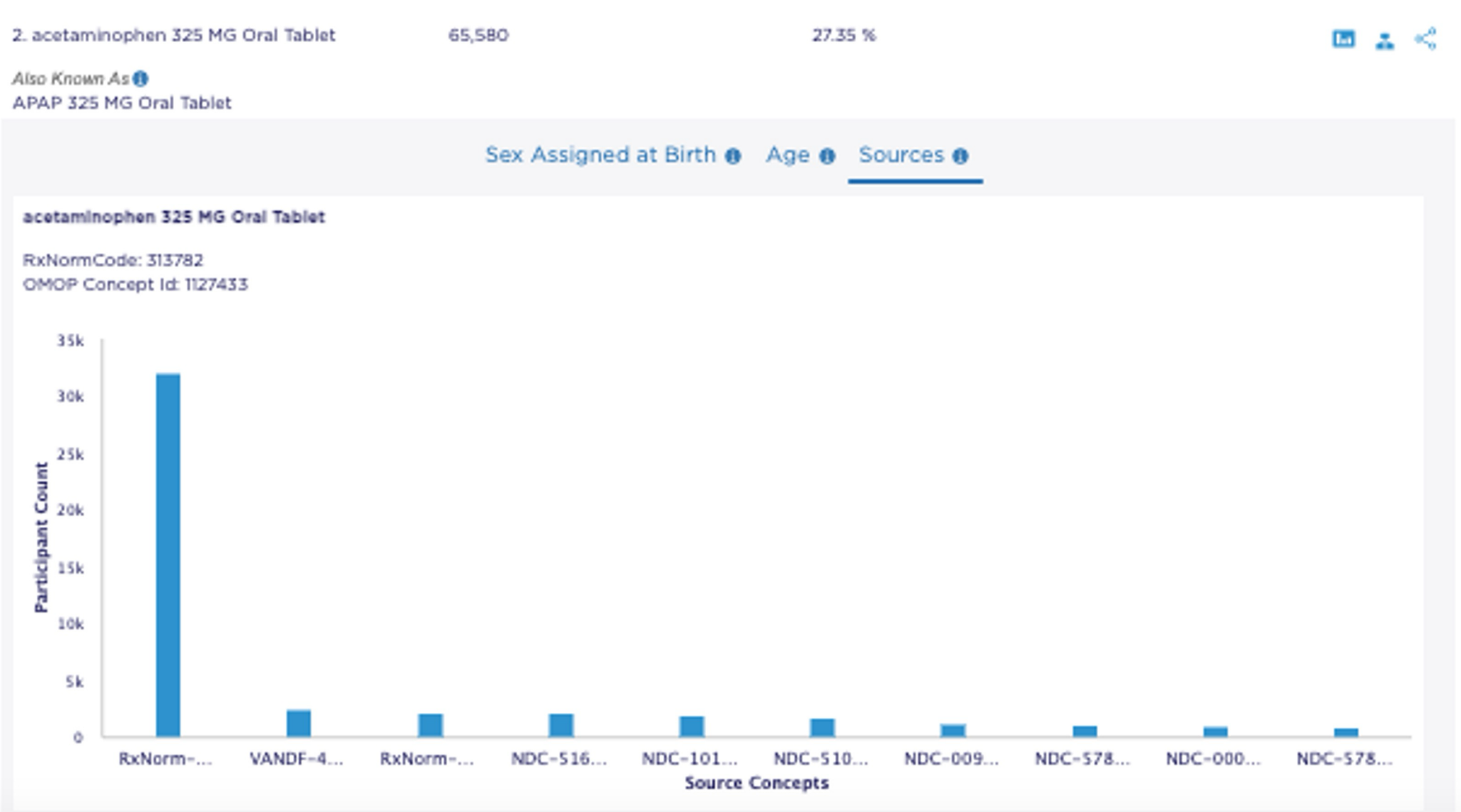
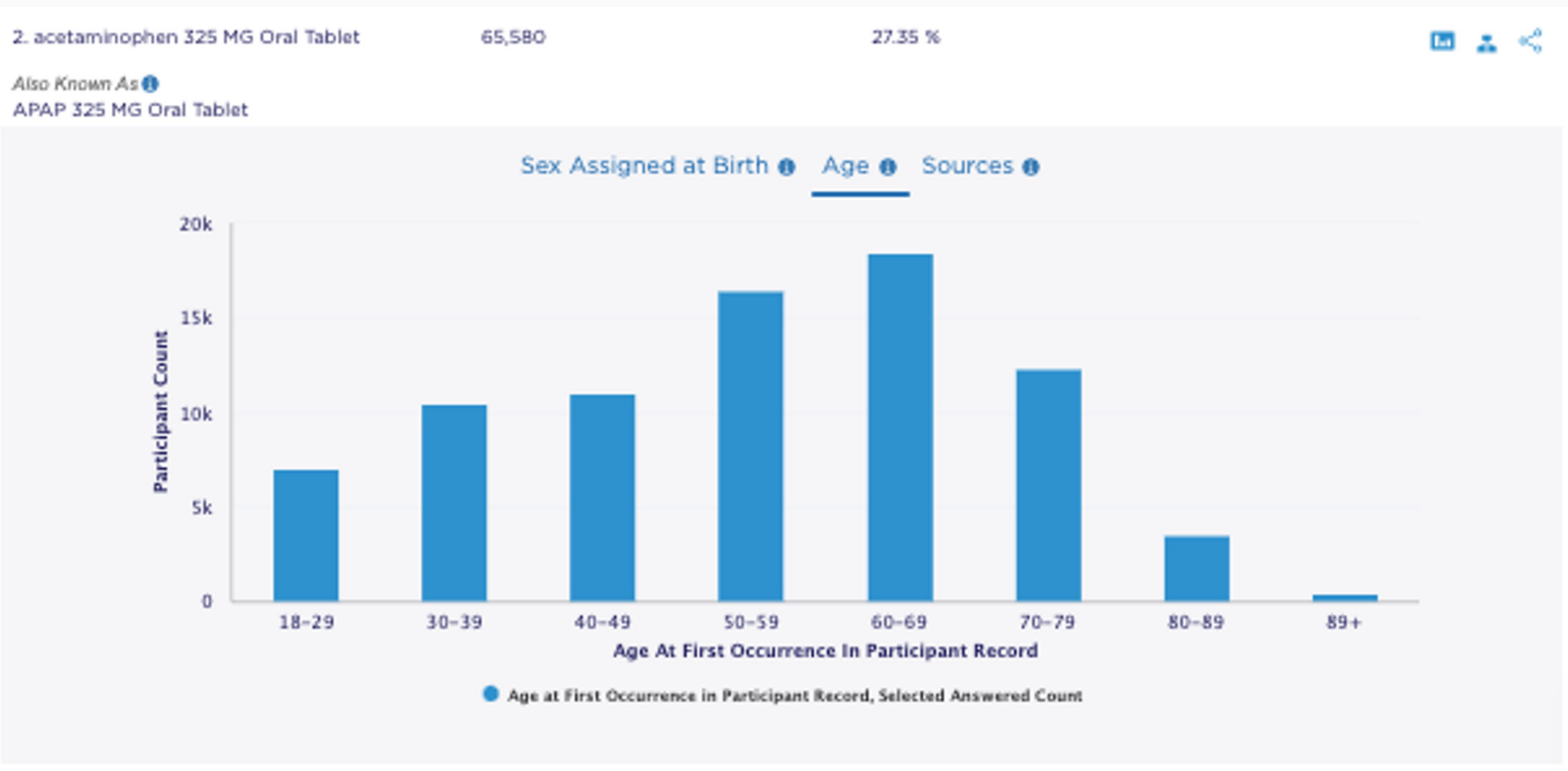
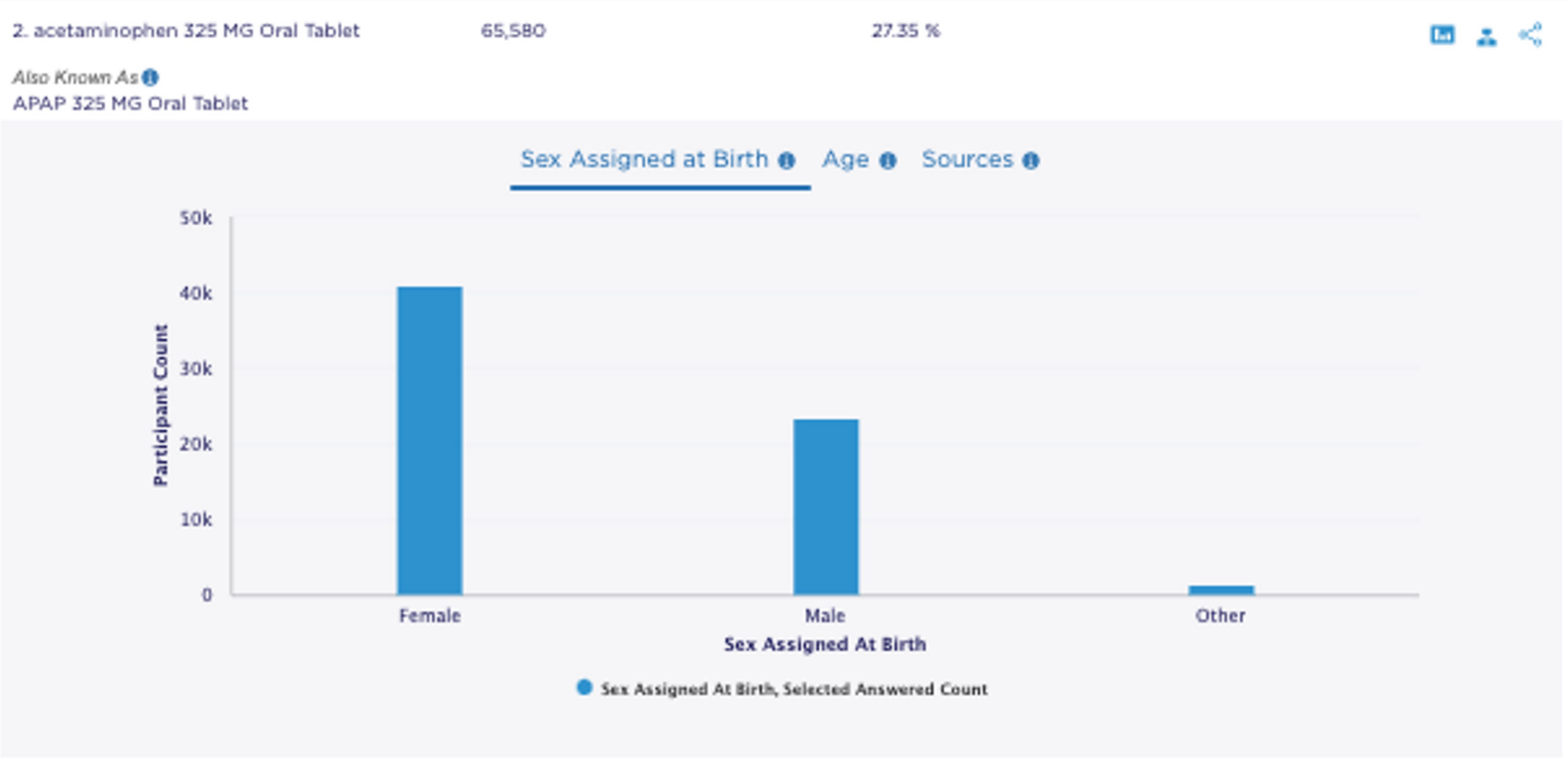


Controlled Tier

Includes data available on the Registered Tier, as well as genomic data and expanded demographic, survey, and EHR data. Genomic data includes whole genome sequences, long read sequences, genotyping arrays, and structural variants. Researchers must have amended institutional agreements and complete an additional training to access this tier.

Data Browser cont'd

Explore data breakdowns by sex assigned at birth & age.



Workspaces



Welcome to the
RESEARCHER WORKBENCH

The secure platform to analyze *All of Us* data



Workspaces

[See all Workspaces](#)

 Featured Workspace: Dem
entia

OWNER
Last Changed: 02/11/20, 07:32 PM

 All of Us Survey Codebook
and Frequency Distribu
t
ions

OWNER
Last Changed: 02/11/20, 07:51 PM

 Featured Workspace: Dep
ression

OWNER
Last Changed: 02/11/20, 07:50 PM

 Featured Workspace - Type
2 Diabetes

OWNER
Last Changed: 02/11/20, 07:50 PM

Recently Accessed Items

 Case 1 Noteboo
k

Last Modified: Mar 06 2020
Notebook

 Dementia Analy
sis from Cohort
Builder

Last Modified: Feb 04 2020
Notebook

 Ischemic Heart
Disease Analysis

Last Modified: Feb 04 2020
Notebook

 Dementia Analy
sis

Last Modified: Feb 04 2020
Notebook

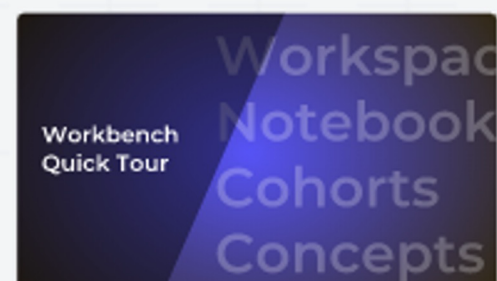
 Type 2 Diabetes
Analysis

Last Modified: Feb 04 2020
Notebook

 Ischemic Heart
Disease Analysis

Last Modified: Jan 31 2020
Notebook

Quick Tour and Videos



Workbench
Quick Tour

Workspace
Notebook
Cohorts
Concepts

3:01



Cohorts
Overview

3:07



Notebooks
Overview

3:07

Answers to common questions on the User Support Hub

No Login Required! Visit the Support Hub to learn more about becoming a registered researcher and using the *All of Us* dataset.



Find instructional materials about the *All of Us* Researcher Workbench

- Includes video tutorials, educational resources, release notes, and more

Search 320+ articles*

- Find information about data dictionaries, the Controlled Tier directory, how WGS and array data are organized, and more

Connect with experts during weekly LIVE office hours

- Explore our calendar of dedicated office hours where researchers can talk about data types and tools with experts



researchallofus.org/support

*Source: Figure as of August 2024

[Home](#) > [Research](#) > [Research Activities](#) > [National Clinical Cohort Collaborative \(N3C\)](#) > National Clinical Cohort Collaborative (N3C)

National Clinical Cohort Collaborative (N3C)

N3C translates health data into health solutions.

N3C Overview

N3C Overview

N3C is a national resource of real

With stewardship from NCATS, data derived from electronic health records can be analyzed. The enclaves offer spaces for collaboration and create a more complete picture of the patient.

Initially created in response to the need for additional health conditions.

We know that the data in the end
comprehensive approach to add



Pilot Enclaves

The N3C Pilot Enclaves leverage operational and governance aspects of the N3C COVID platform that are established already. The projects focus on using real-world clinical data to better understand chronic kidney disease and cancer.

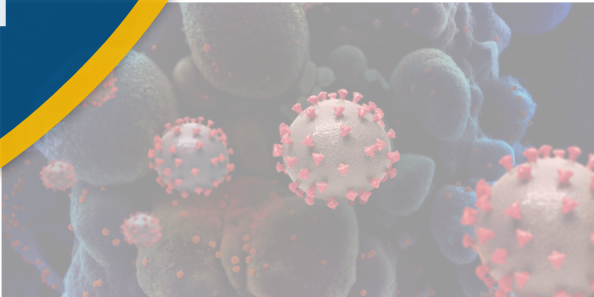
Learn About the Pilot Projects



Education Enclave

The N3C Education Enclave offers a space for researchers to develop and practice the skills needed to analyze real-world data. It includes data from 500,000 simulated patients that represent common elements found within electronic health records.

Learn About the N3C Education Enclave



COVID Enclave

The N3C COVID Enclave is the largest open U.S. database for ongoing COVID-19 research. Researchers and clinicians have used this national resource to reveal risk factors for Long COVID, help health care providers choose effective treatments, and so much more.

Learn About the N3C COVID Enclave

The N3C Data Enclave represents one of the largest secure collections of harmonized clinical health data in the United States.

**Sites:** 84

**Lab Results:** 16.3 billion

**Persons:** 22.8 million

**Medication Records:** 5.3 billion

**COVID+ Cases:** 8,914,402

**Procedures:** 1.2 billion

**# of Rows:** 33.9 billion

**Visits:** 2.0 billion

**Clinical Observations:** 3.3 billion

Data as of October 10, 2024

Explore Our Dashboards

Overview



Members



Tracking



Cases



Medications



Diseases & Risk Factors





Partners united as one community to improve research.

Trust is built. Answers are accelerated. Health improves.

A Decade of Patient-Centered Research Powered by

 **pcornet**
The National Patient-Centered Clinical Research Network

10

YEARS

PCORnet is a national resource, funded by PCORI, that enables insights from high quality health data, patient partnership, and research expertise deliver fast, trustworthy answers that advance health outcomes.

What Makes PCORnet Unique?

High Quality Health Data

Research Expertise

PCORnet

A national resource for enabling patient-centered comparative effectiveness research.



Event

Thursday, November 21st
Noon to 1pm ET

Join us for the webinar, *Lessons Learned from the Researching Adult to Advance Recovery RECOVER: PCORnet Adult EHR Hub*. In this webinar, Dr. Rainu Kaushal and Dr. Thomas Carton will review the strategic mobilization of 40 PCORnet sites to enable large scale real-world research, enhancement of data resources, research throughput, and best practices for collaboration.

Contact pmo@pcornet.org for a calendar invite.

Tools to speed study start-up

Use these [tools powered by PCORnet® Network Partners](#) to accelerate research.

Query Reports

[Telehealth](#)

[Intellectual and Developmental Disabilities](#)

[Maternal Morbidity and](#)

Latest News

2024 PCORI Annual Meeting Will Showcase How PCORnet® Is Powering Maternal Health Insights

Despite the U.S. spending more on childbirth-related healthcare than any other country, it has the highest maternal mortality rate among high-income nations at approximately 32.9 maternal deaths per 100,000 live births, with women of color disproportionately affected. At the upcoming 2024 PCORI Annual Meeting on October 22-23, 2024, researchers will demonstrate how

Shaping Research on Telehealth with PCORnet® Resources

A distinct and sudden need for virtual medical visits created by COVID-19 has led to an extraordinary expansion of telemedicine across nearly all medical specialties in the United States. Explore patient demographics, telehealth trends, and the influence telehealth has had in managing chronic conditions among underserved populations before and during the COVID-19 pandemic in this query of PCORnet data resources. This

About - The National Patient- x +

https://pcornet.org/about/

Newsroom Contact the Front Door



ABOUT PARTNERS RESEARCH DATA ENGAGEMENT RESOURCES

About PCORnet®

PCORnet: A Network of Networks

PCORnet is a national resource that offers the kind of research ecosystem that has long been pursued: a fully integrated network where vast, highly representative health data, research expertise, and patient insights are built-in and accessible from the very start. The infrastructure of PCORnet is well established, meaning that PCORnet® Network Partners know how to maximize the value of these connections to deliver fast, trustworthy answers that advance public health.

Research

PCORnet is uniquely networked to offer the real-world evidence essential for answering the [clinical questions](#) that impact peoples' lives and support studies addressing major public health issues.

Data

With access to [data from everyday healthcare encounters](#) with more than 47 million people annually across the U.S., users of PCORnet have potential insights at their fingertips. These data are accessible via a distributed research network model of eight large PCORnet® Clinical Research Networks (CRNs) and facilitated by the Coordinating Center for PCORnet®.

PCORnet: An Introduction

Get an overview of PCORnet in this [fact sheet](#) or learn more in this [overview presentation](#) for industry stakeholders, academic audiences, and health services researchers.

Learn about PCORnet milestones over the past decade in this [timeline](#).

Knock on the PCORnet® Front Door to access PCORnet



PCORnet is well suited to support the conduct of:

- Real-world evidence studies
- Comparative clinical effectiveness research
- Population health research
- Pragmatic research
- Health systems research
- And more

Providers and researchers

EHRs, claims, and patient-reported data



Front Door - The National Pati x +

https://pcornet.org/front-door/

Newsroom Contact the Front Door



ABOUT PARTNERS RESEARCH DATA ENGAGEMENT RESOURCES

PCORnet® Front Door

How do you access PCORnet?
PCORnet is a national resource available to everyone.

The PCORnet® Front Door is an access point for potential investigators, patient groups, healthcare organizations, clinicians and clinician groups, government, industry scientists, sponsors, and all stakeholders seeking to leverage PCORnet infrastructure and collaborate on patient-centered research.

CONTACT THE PCORNET® FRONT DOOR

Knock on the PCORnet® Front Door to access PCORnet



Tools to speed study start-up

Use these [tools powered by PCORnet® Network Partners](#) to accelerate research.

Common ways the PCORnet® Front Door team can help support your research needs:

Data Network Request

Data Network Requests generally provide preliminary data to help design proposals or assess study feasibility.

- The PCORnet® Front Door team contacts you to schedule a consultation and understand your data needs
- If appropriate, the team will partner to develop your query for the [PCORnet® Common Data Model](#), and securely distribute the query to PCORnet data partners. PCORnet data partners will choose whether or not to execute the query and will send back results.
- Once data are returned to the PCORnet® Front Door, the team will meet with you to review your data report and answer any questions

Network Collaborator Request

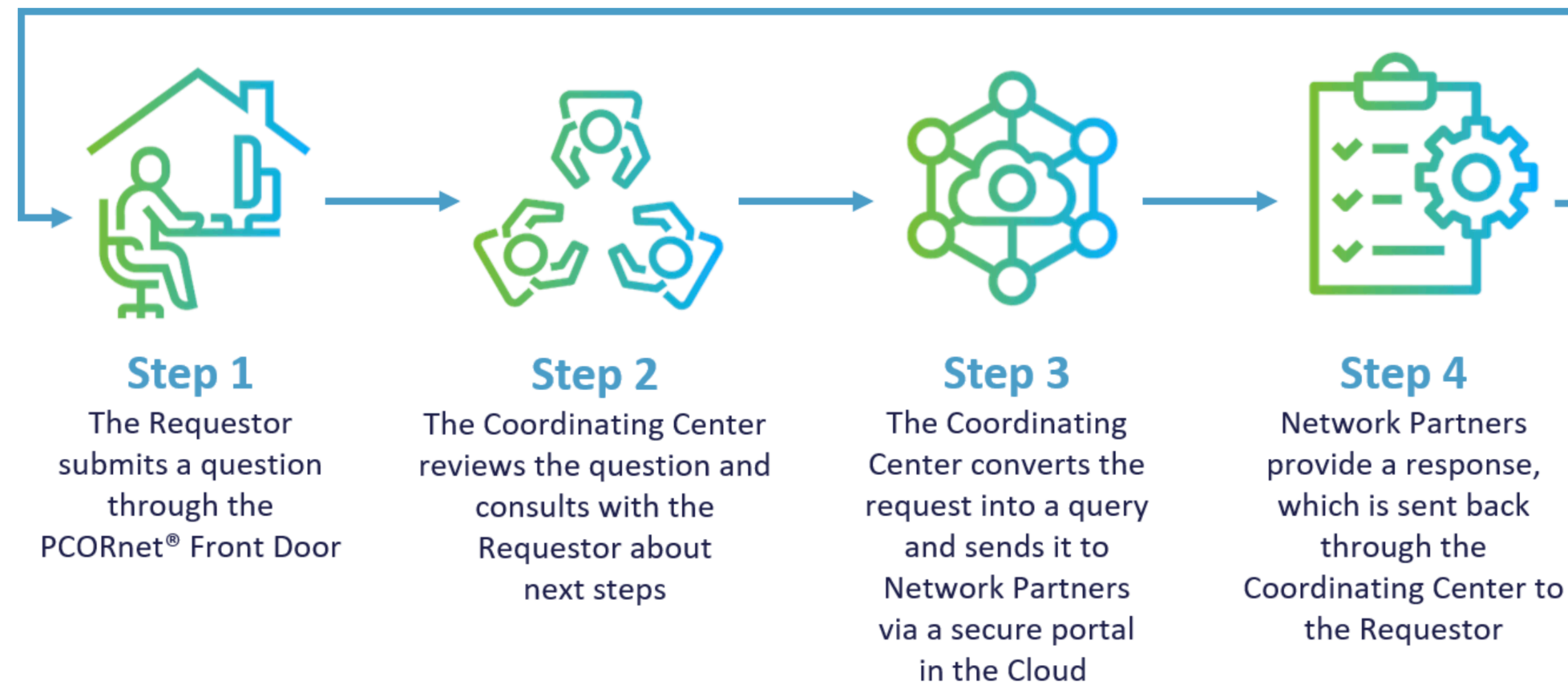
The PCORnet® Front Door team can connect you with collaborators within the Network. You can seek sites for funded studies, partners to co-design research leveraging PCORnet, or find people with specific expertise.

- The PCORnet® Front Door team shares your request by hosting an informational webinar to the Network
- The team sends information to each network; each network then disseminates internally
- The team collects questions and contact information and shares it with you.

The PCORnet® Front Door can advise you on site scope of work and budget parameters, however the requestor will generally execute

How It All Works

The [PCORnet® Front Door](#) is a central gateway point for anyone seeking to use PCORnet resources. Once users have reached out, the Coordinating Center for PCORnet® sets up a consultation to clarify the request and convert the question into a query to send out to the PCORnet® Network Partners for the best fit.





Service Areas

1. Honest broker for retrieving UCI Health patient data according to IRB-approved research protocols
2. Partially de-identified, self-service research data warehouses
3. Self-service cohort discovery tools
4. Advanced data analytics: artificial intelligence / machine learning (AI/ML), natural language processing (NLP)
5. Research data sharing networks
6. **Research software development**
7. Researcher engagement and enrichment

uci aou app demo (logout)

ReportsDataTraining

Search for participants

RECRUITMENT

Scheduled/Reminder (350)

Initial Contact (2,100)

Said Yes, Didn't Sign Up (5,150)

No Decision Yet (47,761)

Orphaned Said Yes (1,869)

Orphaned No Decision (16,907)

RETENTION

Scheduled/Reminder (136)

Said Yes, Not Yet Retained (1,055)

No Decision Yet (2,184)

Orphaned Said Yes (0)

Orphaned No Decision (0)

#1

Jane Smith👩🇬🇧

Study ID

DOB

Other name

Pilot

Text Pilot

000-000-0000

kai.zheng@uci.edu

Agreed 11/7/2019

Signed up 11/8/2019

P955797279

Became FULL 11/9/2019

Family Health

Care Access

Medical History

COPE (May)

COPE (June)

COPE (July)

DNA Consent

Last clinic visit: Santa Ana Family Medicine

Date: 11/8/2019

History (3 days ago)

9/15/2020 4:26 pm by

Type

Outcome

Alt. outcome

Response

Retention (CRC outreach)

No voicemail or full

Sent email

--

6/23/2020 10:12 am by

Type

Outcome

Alt. outcome

Response

Retention (CRC outreach)

--

Sent text

--

6/17/2020 1:54 pm by

Type

Outcome

Alt. outcome

Response

Retention (CRC outreach)

--

Sent text

--

5/18/2020 11:52 am by

Type

Outcome

Alt. outcome

Response

Retention (CRC outreach)

Left voicemail

--

--

5/11/2020 10:58 am by

Type

Outcome

Alt. outcome

Response

Retention (CRC outreach)

Spoke with participant

--

Agreed to do on own

still has not had time. Offered help on phone, but said she is busy but will get to surveys so help not needed

5/4/2020 11:52 am by

Type

Outcome

Alt. outcome

Response

Retention (CRC outreach)

Spoke with participant

--

Agreed to do on own

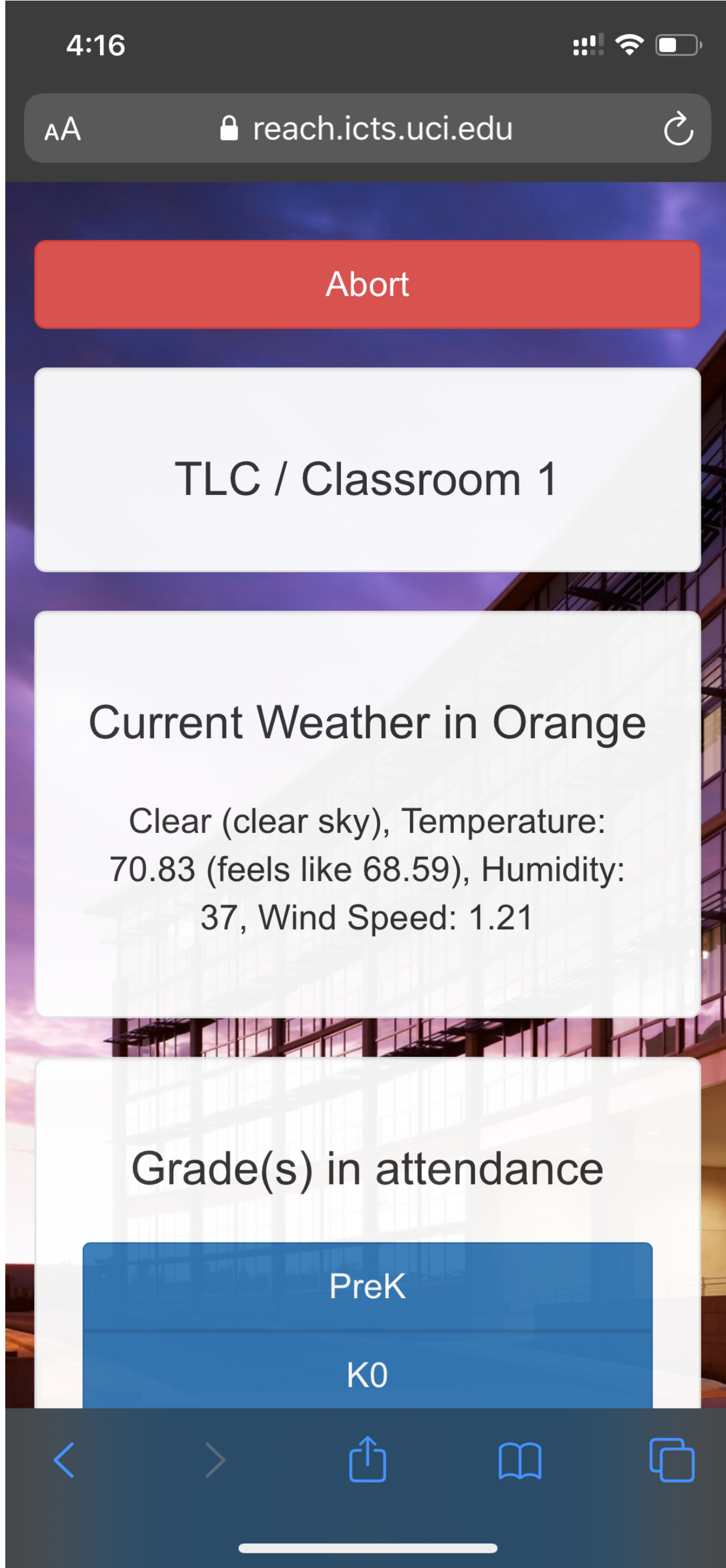
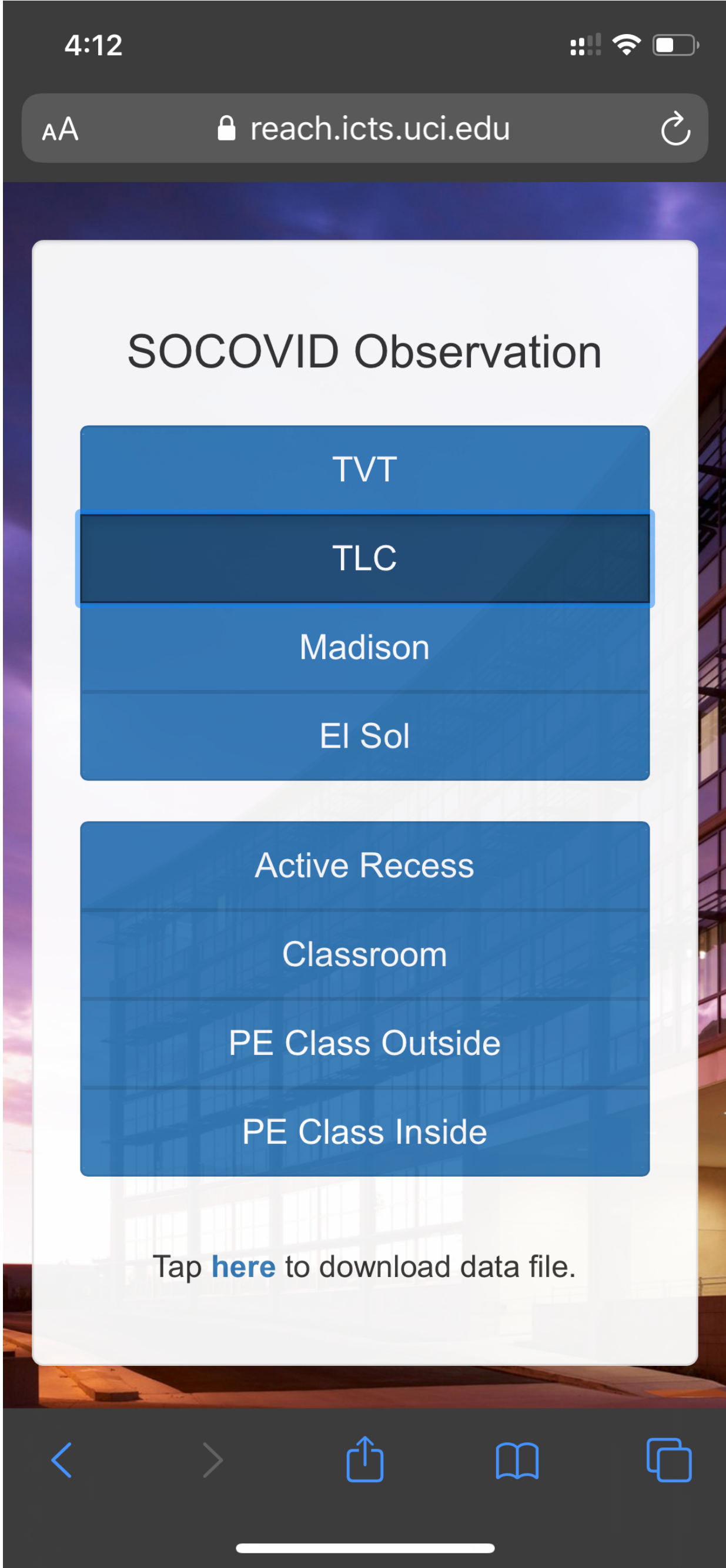
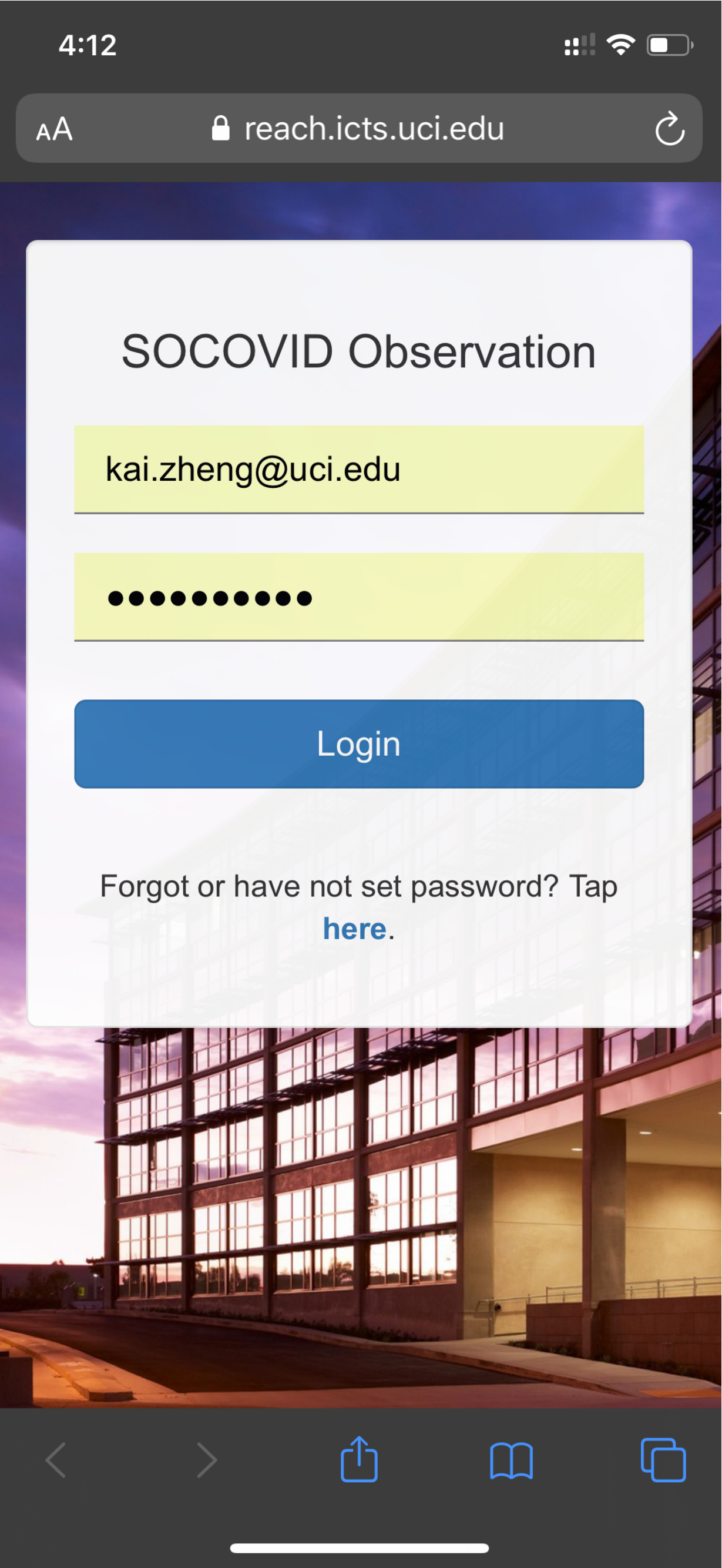
4/22/2020 11:21 am by

Type

Outcome

Retention (CRC outreach)

Left voicemail





Service Areas

1. Honest broker for retrieving UCI Health patient data according to IRB-approved research protocols
2. Partially de-identified, self-service research data warehouses
3. Self-service cohort discovery tools
4. Advanced data analytics: artificial intelligence / machine learning (AI/ML), natural language processing (NLP)
5. Research data sharing networks
6. Research software development
7. **Researcher engagement and enrichment**

Thanks!

kai.zheng@uci.edu

<https://reach.icts.uci.edu/ori/>