



University of California
San Francisco

Using Multi-Institutional Electronic Health Data

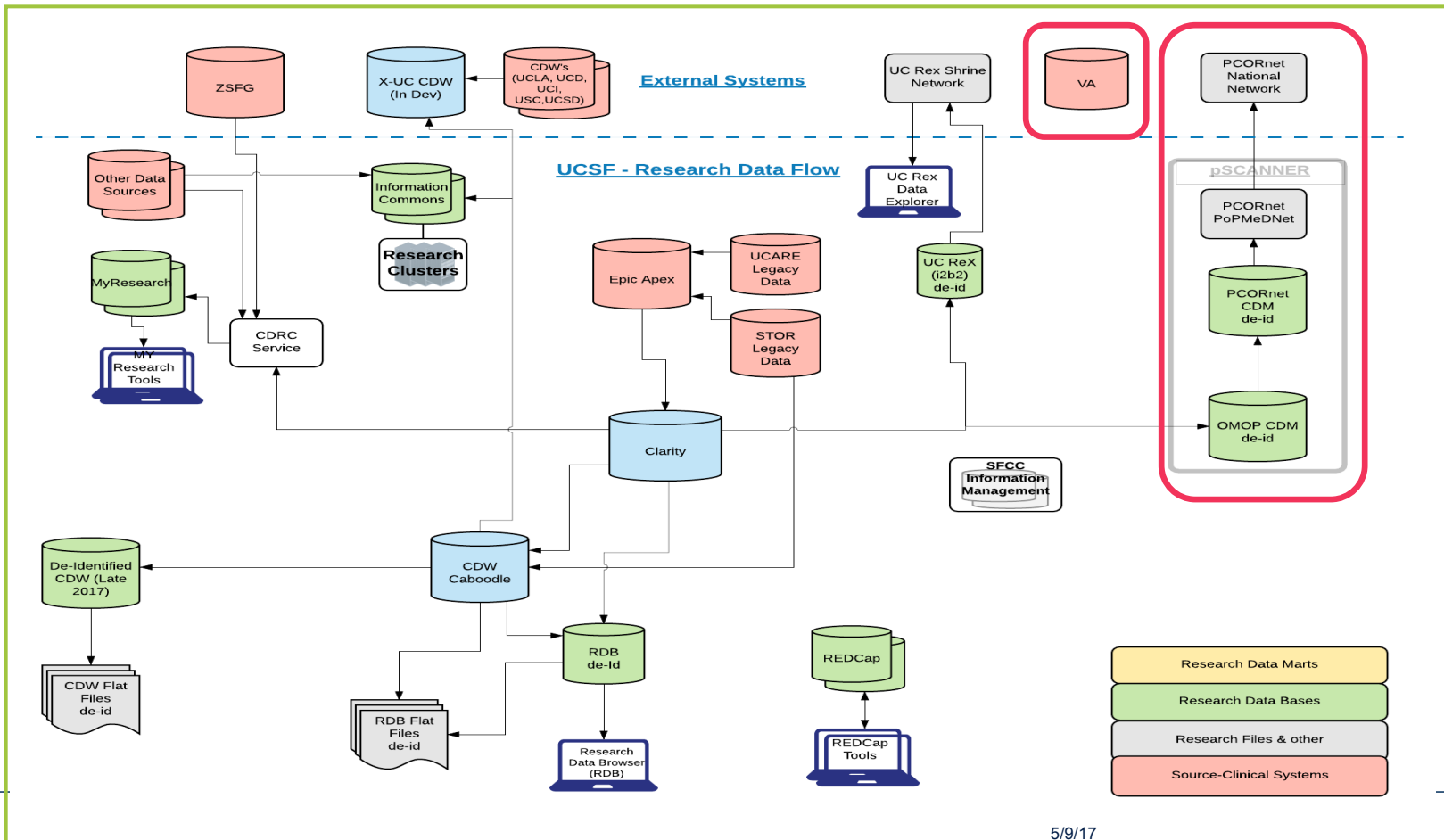
Mary Whooley, MD
Professor of Medicine, Epidemiology & Biostatistics

Multi-Institutional Electronic Health Data

Scenario	Self-serve (Free)	Consult Required (May have recharge)	IRB needed	Requires MyResearch account or other secure environment	Includes clinical notes	UC Health data available in addition to UCSF data
Counts	Yes	No	No	No	No	Yes
De-ided data	Yes	No	No	No	No	Yes *
Limited data (pSCANNER)	No	Yes	Yes	Yes	No	Yes *
Id'ed data (VA)	No	Yes	Yes	Yes	Yes	Yes *
Recruit	No	Yes	Yes	Yes	No	Yes *

* Process to obtain data from UC Health sites can be very lengthy

Multi-Institutional Electronic Health Data



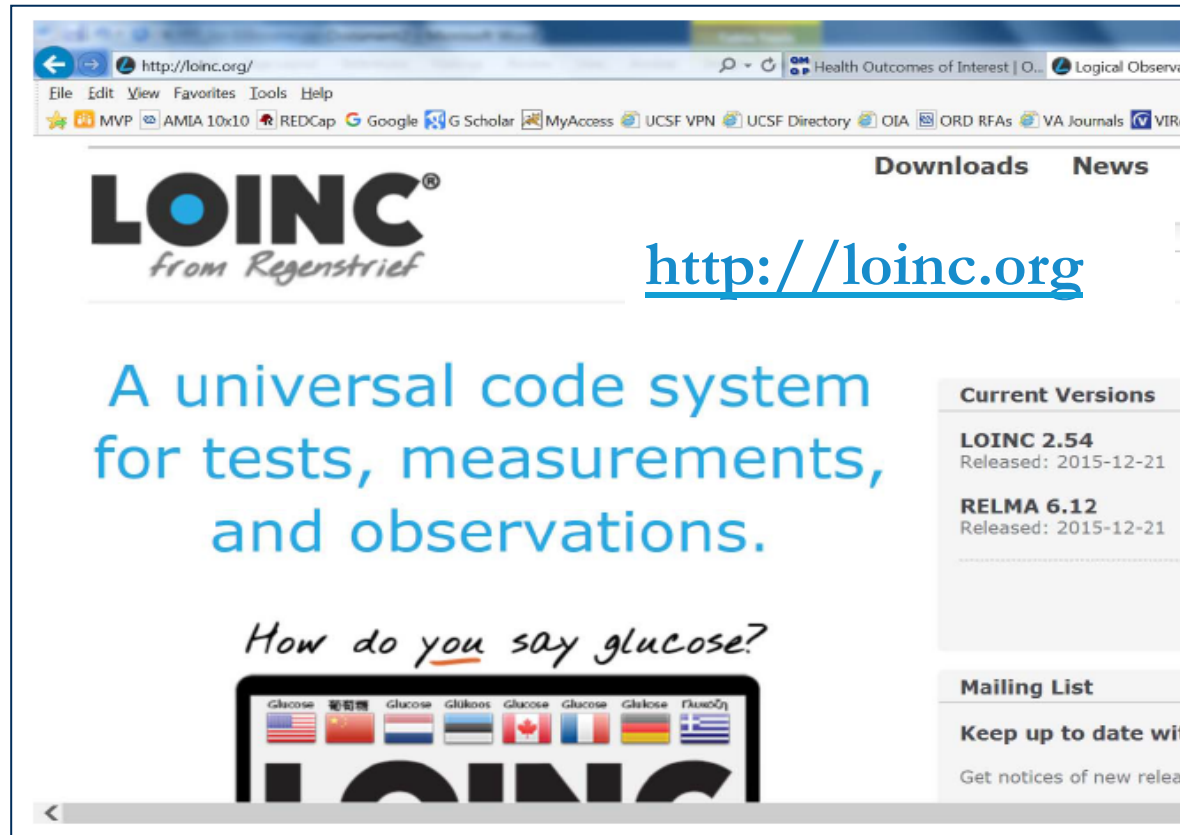
Outline

- Need for EHR standardization + interoperability
- UCSF and VA involvement in PCORnet
- Importance of OMOP Common Data Model
- Using VA data for health services research

What was the average white blood cell count?

Value	LabChemTestName	units
7.5	AUTO WBC	/UL
6.2	C-WBC-CBOC	K/CMM
5.9	HOPC-WBC	THOU/CMM
8.3	NAVY STAT WBC	/UL
9.2	NEW WBC	T/CMM
8.7	OB-WBC	K/CMM
4.5	Q-WBC DC'D	X10E9/L
5.7	TOTAL WBC COUNT (AML)	X10`9/L
6.4	TOTAL WHITE BLOOD COUNT	G/L
7.8	WBC	K/UL
7.5	WBC (AA)	K/CMM
6.8	WBC (AUTOMATED)	K/CUMM
4.5	WBC (AUTOMATED) WR	K/MM3

Logical Observation Identifiers Names and Codes



The screenshot shows the LOINC website homepage in a web browser. The browser's address bar displays <http://loinc.org/>. The page features the LOINC logo with the tagline "from Regenstrief". A large blue heading reads "A universal code system for tests, measurements, and observations." Below this, a handwritten-style question asks "How do you say glucose?" with the word "glucose" underlined. The answer is shown as a row of flags corresponding to different languages: American English (Glucose), Chinese (葡萄糖), Dutch (Glucose), German (Glukose), French (Glucose), Spanish (Glucose), and Greek (Γλυκόζη). To the right, a "Current Versions" section lists "LOINC 2.54" (Released: 2015-12-21) and "RELMA 6.12" (Released: 2015-12-21). A "Mailing List" section is partially visible at the bottom right.

What was the average white blood cell count?

Value	LabChemTestName	units	LOINC
7.5	AUTO WBC	/UL	6690-2
6.2	C-WBC-CBOC	K/CMM	6690-2
5.9	HOPC-WBC	THOU/CMM	6690-2
8.3	NAVY STAT WBC	/UL	6690-2
9.2	NEW WBC	T/CMM	6690-2
8.7	OB-WBC	K/CMM	6690-2
4.5	Q-WBC DC'D	X10E9/L	6690-2
5.7	TOTAL WBC COUNT (AML)	X10`9/L	6690-2
6.4	TOTAL WHITE BLOOD COUNT	G/L	6690-2
7.8	WBC	K/UL	6690-2
7.5	WBC (AA)	K/CMM	6690-2
6.8	WBC (AUTOMATED)	K/CUMM	6690-2
4.5	WBC (AUTOMATED) WR	K/MM3	6690-2

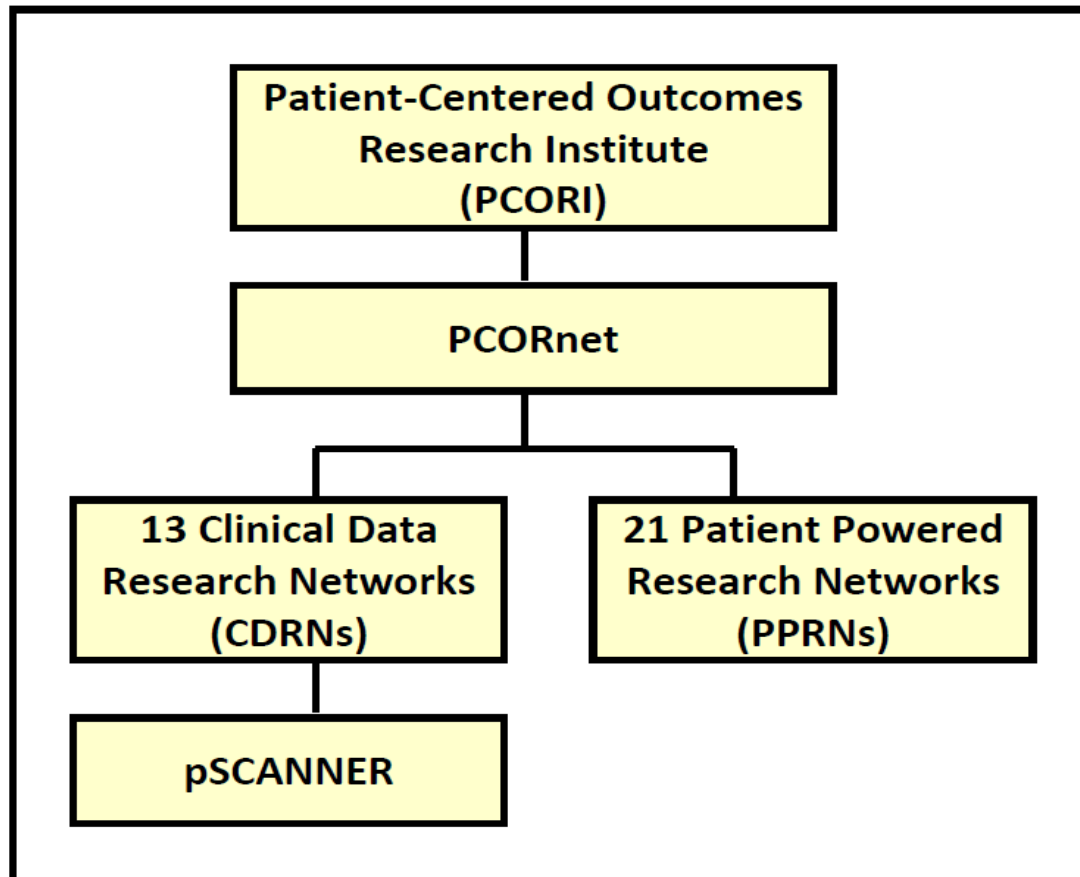
Outline

- Need for EHR standardization + interoperability
- UCSF and VA involvement in PCORnet
- Importance of OMOP Common Data Model
- Using VA data for health services research

PCORnet, the National Patient-Centered Clinical Research Network

PCORnet, the National Patient-Centered Clinical Research Network, is an innovative initiative of the Patient-Centered Outcomes Research Institute (PCORI). It is designed to make it faster, easier, and less costly to conduct clinical research than is now possible by harnessing the power of large amounts of health data and patient partnerships. In the process, it is transforming the culture of clinical research from one directed by researchers to one driven by the needs of patients and those who care for them.

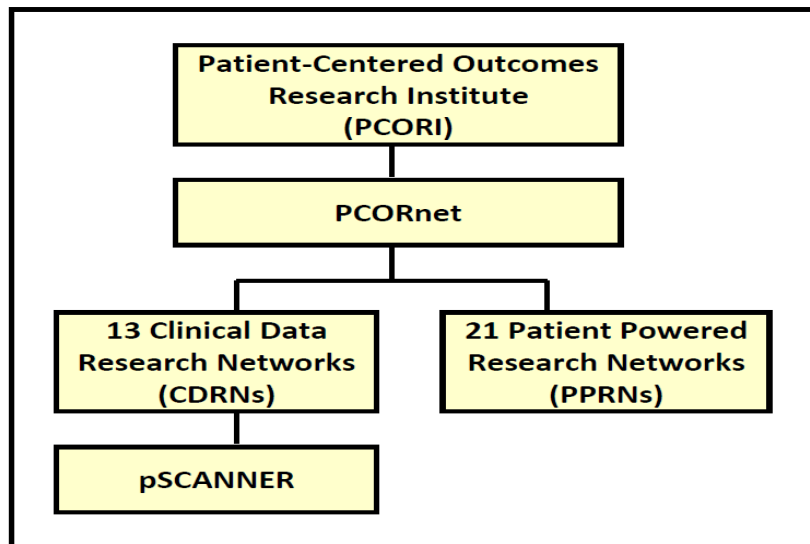
<http://www.pcornet.org/>



<http://pscanner.ucsd.edu/>



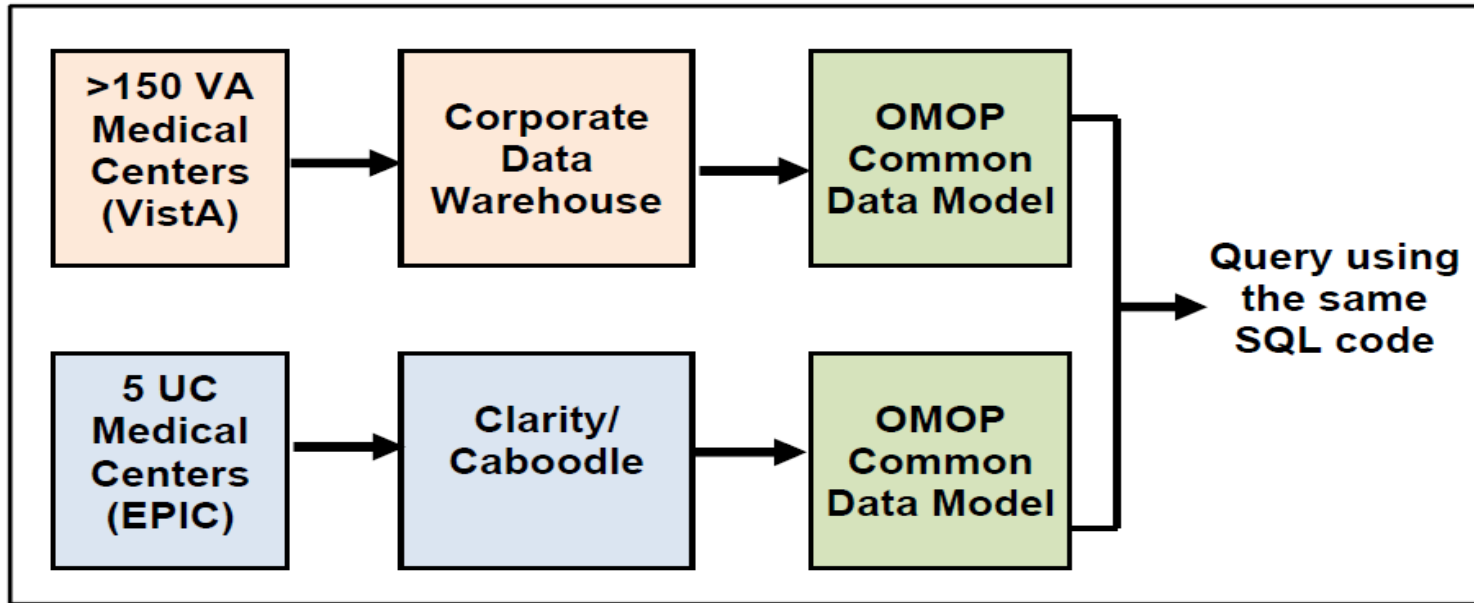
<http://pscanner.ucsd.edu/>



Organizations participating in pSCANNER

- (1) UC San Diego
- (2) UC San Francisco
- (3) UC Irvine
- (4) UC Davis
- (5) UC Los Angeles
- (6) VA (150 medical centers)
- (7) University of Southern California
- (8) University of Colorado
- (9) University of Washington
- (10) San Mateo Medical Center
- (11) Cedars-Sinai Medical Center

<http://pscanner.ucsd.edu/>



OMOP = Observational Medical Outcomes Partnership

SQL = Structured Query Language

Outline

- Need for EHR standardization + interoperability
- UCSF and VA involvement in PCORnet
- Importance of OMOP Common Data Model
- Using VA data for health services research

OMOP Implementations



> 600 Million Patients

Welcome to OHDSI!

The Observational Health Data Sciences and Informatics (or OHDSI, pronounced "Odyssey") program is a multi-stakeholder, interdisciplinary collaborative to bring out the value of health data through large-scale analytics. All our solutions are open-source.

OHDSI has established an international network of researchers and observational health databases with a central coordinating center housed at Columbia University.

Read more [about us](#), about [our goals](#), and how you can [help support the OHDSI community](#).

[Join the Journey](#)



www.ohdsi.org

<https://www.youtube.com/watch?v=wGdqGOQNkuM>

Standardized vocabularies (open source):

<http://vocabqueries.omop.org/>

- find a condition by keyword
- find a drug
- find a lab test
- find a procedure

Standardized queries (open source):

<http://cdmqueries.omop.org/>

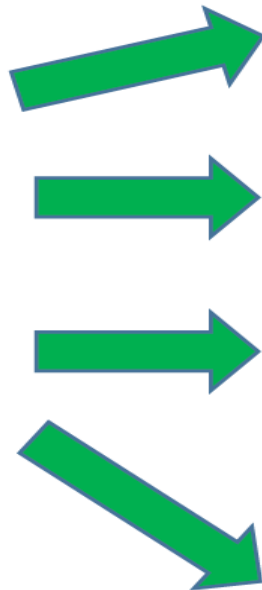
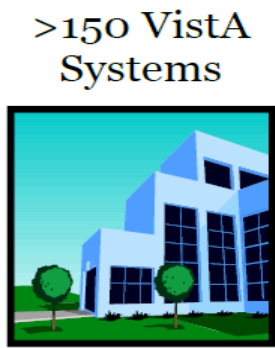
- find number of patients with condition
- determine mortality rate after diagnosis
- find number of patients on medication

Outline

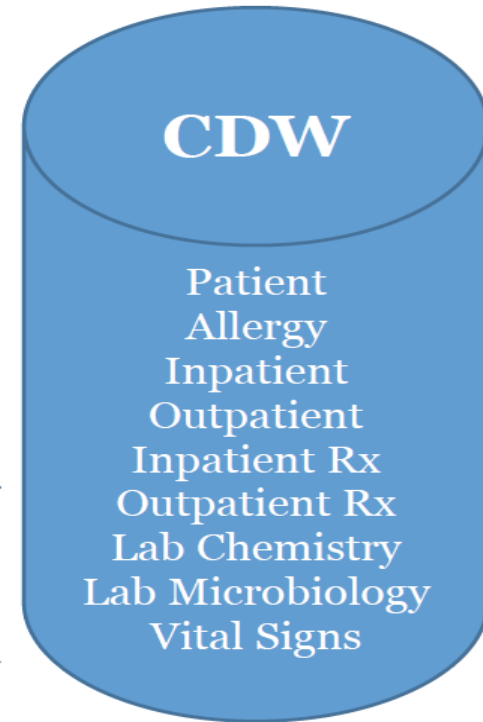
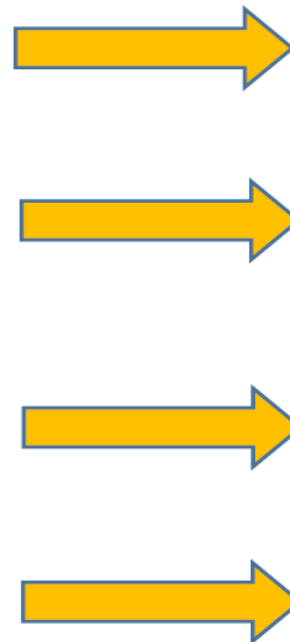
- Need for EHR standardization + interoperability
- UCSF and VA involvement in PCORnet
- Importance of OMOP Common Data Model
- **Using VA data for health services research**

CDW Origin & Content

<https://www.virec.research.va.gov/>



Extract,
Transform
and Load



Minute by minute transactions added – no changes

Nightly updates to architected database - no change to content

Transforming the National Department of Veterans Affairs Data Warehouse to the OMOP Common Data Model

Fern FitzHenry^{1,2}, Jesse Brannen¹, Jason Denton^{1,2}, Jonathan R. Nebeker^{3,4}, Scott L. DuVall^{3,4}, Freneka Minter^{1,2}, Jeffrey Scehnet³, Brian Sauer^{3,4}, Lucila Ohno-Machado⁵, Michael E. Matheny^{1,2}

¹Tennessee Valley Healthcare System, Veterans Affairs Medical Center, Nashville, TN;

²Vanderbilt University, Nashville, TN; ³VA Salt Lake City Health Care System, Salt Lake City, UT;

⁴University of Utah, Salt Lake City, UT; ⁵Bioinformatics and Systems Biology, University of California, San Diego, CA;

Abstract: To describe the conversion of the national Department of Veterans Affairs (VA) healthcare network's corporate data warehouse to the Observational Medical Outcomes Partnership (OMOP) common data model (CDM) suitable for distributed observational research. Observational outcomes from electronic medical record systems are becoming more important in comparative effectiveness research, particularly as post marketing surveillance research.¹

2000 to 2015

- 16 million unique patients
- 11 million w/ at least one encounter
- 5 million deaths
- 3 billion procedures
- 2.5 billion conditions
- 973,000 providers
- 9 million current enrollees

**Abstract presented Nov 2015
Am Medical Informatics Assoc**

Medicare Claims Files

- Institutional Files
 - Inpatient
 - Skilled Nursing Facility (SNF)
 - Hospice
 - Home Health Agency (HHA)
 - Outpatient
- Institutional Stay Level File
 - Medicare Provider Analysis and Review (MedPAR)
- Non-institutional Files
 - Carrier (Physician/Supplier)
 - Durable Medical Equipment (DME)

Home » Conducting Research » Research Consultation »

VA Data Core Consultation

The Veterans Health Administration (VHA) maintains a central data repository with health information from the electronic medical records of over 9 million Veterans in the United States. Data fields include (but are not limited to) age, gender, race, ethnicity, zip code, dates of outpatient and inpatient encounters, conditions, procedures, lab results, vital signs, medications, prescription refills, radiology reports, costs, vital status, location of care, and national provider identification number. Data files from the Center for Medicare and Medicaid Services (CMS), including the Medicare Chronic Conditions Warehouse, are also available. Access to these databases has substantially improved during the past several years making it possible for both VA and non-VA investigators to use them for clinical research projects.

The primary purpose of the **VA Data Core** is to facilitate access to VHA data for clinical research and improve the productivity of UCSF investigators. The VA Data Core:

1. Provides consultation regarding available data.
2. Facilitates necessary paperwork and approvals.
3. Ensures that use of VHA data meets all regulatory requirements.
4. Assists with identifying, extracting, and merging variables of interest.

Please note that obtaining access to VA data may require up to 40 hours of your time over a period of 3 to 6 months prior to starting any data analysis. In addition, you will need to perform your own data management and analysis. Although the VA data core can help you identify and extract data fields of interest (e.g., age, ICD-9 code, lab values, vital status), it does not provide data

Login to request consultation >

(UCSF only)

Data Resources

Access Electronic Health Record (EHR) Data:

- [UCSF APeX](#)
- [UCSF & ZSFGH](#)
- [VA Data Core](#)

Access Public Data:

- [Large Dataset Inventory](#)

Additional UCSF Data Resources:

- [REDCAP collect & manage data](#)
- [OnCore clinical trial management](#)
- [Library Research Support bioinformatics & data management](#)
- [Informatics Resources](#)

<https://accelerate.ucsf.edu/consult/vadata>

UCSF Accelerate
Access services to enable research

Powered by CTSI

Home ▾ Conducting Research ▾ Training ▾ Funding ▾ Search this Site

For Community Partners ▾

Home » Conducting Research » Research Consultation »

VA Data Core Consultation

Login to request consultation >

(UCSF only)

The Veterans Health Administration (VHA) maintains a central data repository with health information from the electronic medical records of over 9 million Veterans in the United States. Data fields include (but are not limited to) age, gender, race, ethnicity, zip code, dates of service, conditions, procedures, lab results, vital status, radiology reports, costs, vital status, location, identification number. Data files from the Veterans Affairs Medical Services (CMS), including the Medicare and Medicaid databases are available. Access to these databases has been available for over past several years making it possible for researchers to use them for clinical research projects.

The primary purpose of the VA Data Core is to support clinical research and improve the production of the VA Data Core:

1. Provides consultation regarding available data.
2. Facilitates necessary paperwork and approvals.
3. Ensures that use of VHA data meets all regulatory requirements.
4. Assists with identifying, extracting, and merging variables of interest.

REDCAP collect & manage data
OnCore clinical trial management
Library Research Support
bioinformatics & data management
Informatics Resources

Please note that obtaining access to VA data may require up to 40 hours of your time over a period of 3 to 6 months prior to starting any data analysis. In addition, you will need to perform your own data management and analysis. Although the VA data core can help you identify and extract data fields of interest (e.g., age, ICD-9 code, lab values, vital status), it does not provide data

Please note that obtaining access to VA data may require up to 40 hours of your time over a period of 3 to 6 months *prior* to starting any data analysis.

<https://accelerate.ucsf.edu/consult/vadata>

Closing

- EHR interoperability critical for healthcare
- UC Health and VA ahead of the curve
- National VA data available for clinical research