UCLA Data Week

UC Health Data and Analytics Platform

May 26, 2021

Lisa Dahm, PhD
Director, UC Health Data Warehouse
UC Health Data Warehouse

- 6.8M Patients
- 213M Encounters
- 570M Procedures
- 700M Conditions
- 1.2B Lab Test Results
- 980M Vital Signs
- 780M Medication Orders

1/1/2012 - present
UC Health Data & Analytics Platform

**Operational Partnerships:**
- Quality & Population Health (QPH)
- Pharmacy
- UC Cancer Consortium (UCCC)
- Supply Chain Analytics (LSfV)

**Research Support:**
- Data Explorer Dashboard
- Data Discovery Platform
- UC CORDS

Diagram showing the relationships between UC Health entities and the University of California Health Data Warehouse.
UC Health Research Support

UC Data Explorer Dashboard - patient cohort counts, including
  o Patients, Procedures, Diagnoses, Medications from the UCHDW

UC CORDS – a HIPAA Limited Data Set (LDS) containing
  o SARS-CoV-2 testing results and inpatient COVID-19 treatment information for patients positive for the virus, collected from across UC Health

Data Discovery Platform – a secure, cloud-based data science environment
  o Utilizes Databricks which contains tools to allow researchers to query and analyze a legally de-identified version of the entire UCHDW
UC Health Analytics Website

https://analytics.uchealth.edu

UC Health Data Warehouse (UCHDW)

The Center for Data-driven Insights and Innovation (CDII) oversees the UC Health Data Warehouse (UCHDW), a unique data asset created to enhance operational improvements, promotion of quality patient care, and to enable the next generation of clinical research.

The UCHDW currently holds data on nearly 6 million patients seen at a UC facility since 2012. These patients received care from nearly 100,000 health care providers in over 200 million encounters, with nearly 200 million procedures, more than half a billion medication orders, and with over 2 billion vital signs measurements and test results. Over 600,000 of these patients are primary care patients.

This data is stored in the Observational Medical Outcomes Partnership (OMOP) data model, enabling a wide range of software tools and computational methods to be used consistently with other state and national efforts. Strategic dashboards showing aggregate data from UC Health sites are available to Leadership teams across UC Health. Please click on the About tab at the top of this page to learn more about select dashboards.

Access to most dashboards is currently limited to UC Health Leadership, including teams working on cross UC Health projects. Using your home UC login name and password, approved users may access the dashboards by clicking on the View Dashboards button below.

View Dashboards
ABOUT CDI2

With strategic direction from UC Health, the University of California, with 19 health professional schools and 6 health systems, has taken a collaborative approach to building the pre-eminent data-driven learning healthcare system that improves the human condition.

The Center for Data-driven Insights and Innovation (CDI2), established in January 2018, is a cornerstone of this collaborative approach. CDI2’s primary function is building and maintaining the data analytics capabilities and technical infrastructure for the UC Health Data Warehouse (UCHDW), a unique data asset created to enhance operational improvements, promote quality patient care, and enable the next generation of clinical research.

Using electronic health records (EHR) from the six health systems within the University of California, this system-level asset also includes claims data from UC’s self-funded health plans, as well as from external sources including Vizient and the California Office of Statewide Planning and Development (OSHPD).

ABOUT DASHBOARDS

Strategic dashboards showing aggregate data from UC Health sites are available to Leadership teams across UC Health. Access to most dashboards is currently limited to UC Health C-suite members, their designees, and individuals working on leadership-approved cross UC projects. However, UC researchers may request access to the UC Health Data Explorer dashboard by contacting their local site’s research analytics contact (listed by site in the CORDS Page of this Website).

Using your home UC login name and password, approved users may access the dashboards by clicking on the View Dashboards button located on the Home Page of this Website. Some information about select dashboards is included below.
Data Explorer Dashboard User Guide

The Data Explorer Dashboard provides researchers with the ability to discover patient cohort counts from system wide clinical data for preparatory to research purposes. The current version of the dashboard leverages a legally deidentified instance of our UC Health Data Warehouse which provides information such as patients, diagnoses, procedures, and medications. We update the Data Explorer monthly with the most recent two years of clinical data. This quick user guide covers the basics of Data Explorer use and to help you start interacting with the dashboard.
<table>
<thead>
<tr>
<th>Racial Category</th>
<th>Hispanic or Latino</th>
<th>Ethnic Category / Gender</th>
<th>Unknown</th>
<th>Grand Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Racial Category</td>
<td>FEMALE</td>
<td>MALE</td>
<td>OTHER</td>
<td>Unknown</td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>1,046</td>
<td>740</td>
<td>1</td>
<td>105</td>
</tr>
<tr>
<td>Asian</td>
<td>1,466</td>
<td>1,125</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Black or African American</td>
<td>1,808</td>
<td>1,299</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MultiRace</td>
<td>3,512</td>
<td>2,939</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>Native Hawaiian or Other Pacific Islander</td>
<td>337</td>
<td>261</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Other Race</td>
<td>41,725</td>
<td>33,241</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td>Unknown</td>
<td>56,074</td>
<td>47,160</td>
<td>8</td>
<td>44</td>
</tr>
<tr>
<td>White</td>
<td>68,775</td>
<td>49,509</td>
<td>4</td>
<td>17</td>
</tr>
<tr>
<td>Grand Total</td>
<td>174,543</td>
<td>136,274</td>
<td>16</td>
<td>83</td>
</tr>
</tbody>
</table>
University of California COVID Research Data Set (UC CORDS)

Overview

The University of California COVID Research Data Set (UC CORDS) is designed to be a timely data set for research purposes, containing SARS-CoV-2 testing results and inpatient COVID-19 treatment information (for those positive for the virus) collected from across UC Health. It is a HIPAA Limited Data Set (LDS) generated from the UC Health Data Warehouse (UCHDW), a UC-wide centralized database with data from all of the medical centers. Certain direct identifiers are removed, but dates of services are retained. The data set is distributed weekly to each UC Health site and will evolve over time as the UCHDW adds more patients and clinical depth.

UC CORDS, as well as the UCHDW, are using the OMOP Common Data Model version 5.3.1.
UC Data Discovery Portal

To further support systemwide research efforts, the Center for Data Driven Insights and Innovation has implemented the UC Data Discovery Platform (UCDDP), a secure cloud-based data science environment. The Data Discovery Platform utilizes Databricks and contains tools to allow researchers to query and analyze a legally de-identified version of the entire University of California Health Data Warehouse (UCHDW).

The Databricks environment enables UC Health researchers (and other UC researchers partnering with a health campus) to write code in R, Python, or Spark SQL that queries the deidentified clinical data. A basic familiarity with OMOP and database queries is required. De-identification of the data has already been completed to enable clinical research projects, under guidance from UC campus institutional reviews boards, privacy and compliance officers, and information security officers. Known identifiers have been removed in compliance with HIPAA, and all dates have been randomly shuffled.

Before gaining access here, researchers will have to:

- Work with their local campuses (or the UC Health campus they are partnering with if they are not affiliated with a UC Health campus) to develop and run their queries
- Undergo local review of the scientific merit of their project
- Undergo local review to confirm the value of running the query centrally
- Comply with all local processes and, upon access, sign an appropriate Data Use Agreement (DUA)

Once approved for access by your local site and provisioned by the UCHDW team, you may access UCDDP by clicking on the button below and following the login instructions you will receive via email.

Data Discovery Portal

UC Health Site Contacts for Access to UCDDP

Text of UC Health Electronic (Click-Through) Data Use Agreement:
Thank you!

Questions?