



UNIVERSITY
OF
CALIFORNIA

Precisely practicing medicine from ten trillion points of UC Health data

Atul Butte, MD, PhD

Chief Data Scientist, University of California Health (UC Health)

Director, Bakar Computational Health Sciences Institute, UCSF

Priscilla Chan and Mark Zuckerberg Distinguished Professor

Conflicts of Interest

- Scientific founder and advisory board membership
 - Genstruct
 - NuMedii
 - Personalis
 - Carmenta
- Honoraria for talks
 - Lilly
 - Pfizer
 - Siemens
 - Bristol Myers Squibb
 - AstraZeneca
 - Roche
 - Genentech
 - Warburg Pincus
- Past or present consultancy
 - Lilly
 - Johnson and Johnson
 - Roche
 - NuMedii
 - Genstruct
- Tercica
- Ecoeos
- Helix
- Ansh Labs
- Prevendia
- Samsung
- Assay Depot
- Regeneron
- Verinata
- Pathway Diagnostics
- Geisinger Health
- Covance
- Wilson Sonsini Goodrich & Rosati
- Orrick
- 10X Genomics
- Medgenics
- GNS Healthcare
- Gerson Lehman Group
- Coatue Management
- Corporate Relationships
 - Northrop Grumman
 - Aptalis
- Allergan
- Astellas
- Thomson Reuters
- Intel
- SAP
- SV Angel
- Progenity
- Illumina
- Speakers' bureau
 - None
- Companies started by students
 - Carmenta
 - Serendipity
 - Stimulomics
 - NunaHealth
 - Praedicat
 - MyTime
 - Flipora
 - Tumbl.in



UCSF Bakar Computational Health
Science Institute

Zuckerberg, Chan give UCSF \$10 million for health data research

By Catherine Ho, San Francisco Chronicle | July 28, 2017



University of California

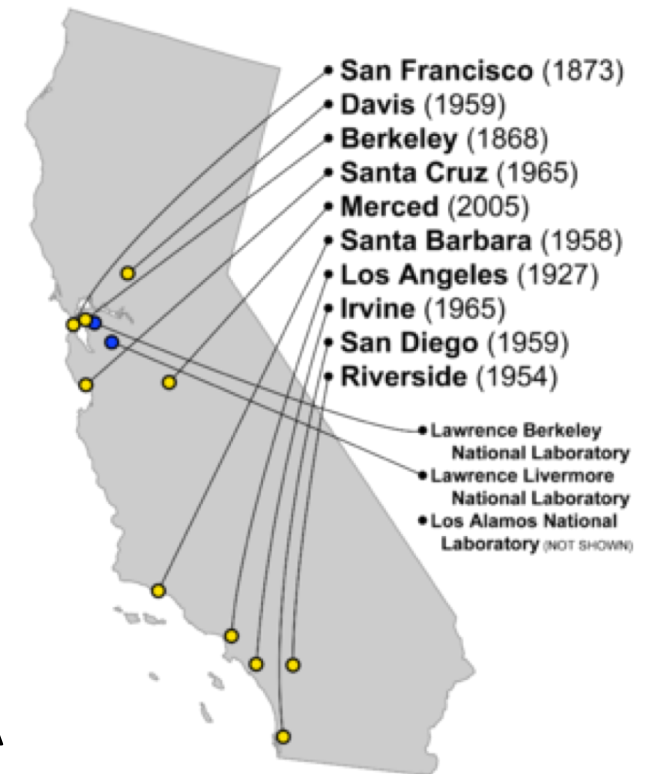
- 10 campuses and 3 national labs
- ~200,000 employees, ~250,000 students/yr



UC Health

UC Health

- 18 health professional schools (6 med schools)
- Train half the medical students and residents in California
- ~\$2 billion NIH funding
- \$11.4 billion clinical operating revenue
- 5000 faculty physicians, 12000 nurses
- UCSF and UCLA are in US News top 10
- 5 NCI Comprehensive Cancer Centers, 5 NIH CTSA



UC Health, United Healthcare Form New ACO & Clinically Integrated Networks

by Staff Writer

10/03/2016



0 Comments



UC Health

The University of California and UnitedHealth Group are teaming up to form a new accountable care organization (ACO) and clinically integrated network. As part of the 10-year strategic relationship, UC Health's five academic medical centers will expand use of Optum's clinically integrated network services and advanced data analytics services.

The United States is spending \$billions on electronic health records, and too few are using any of this data

Sutter's \$1 Billion Boondoggle-New Electronic Records System Goes Dark

California Nurses Association Press Release, 8/27/13

[Contact Information](#) | [Media Center](#)

Yet Another Risk

A controversial e
billion went comp
additional risk be

For several mont
care delivery tha
over 100 reports
Oakland, docum

Sutter managem

Partners' \$1.2b patient data system seen as key to future

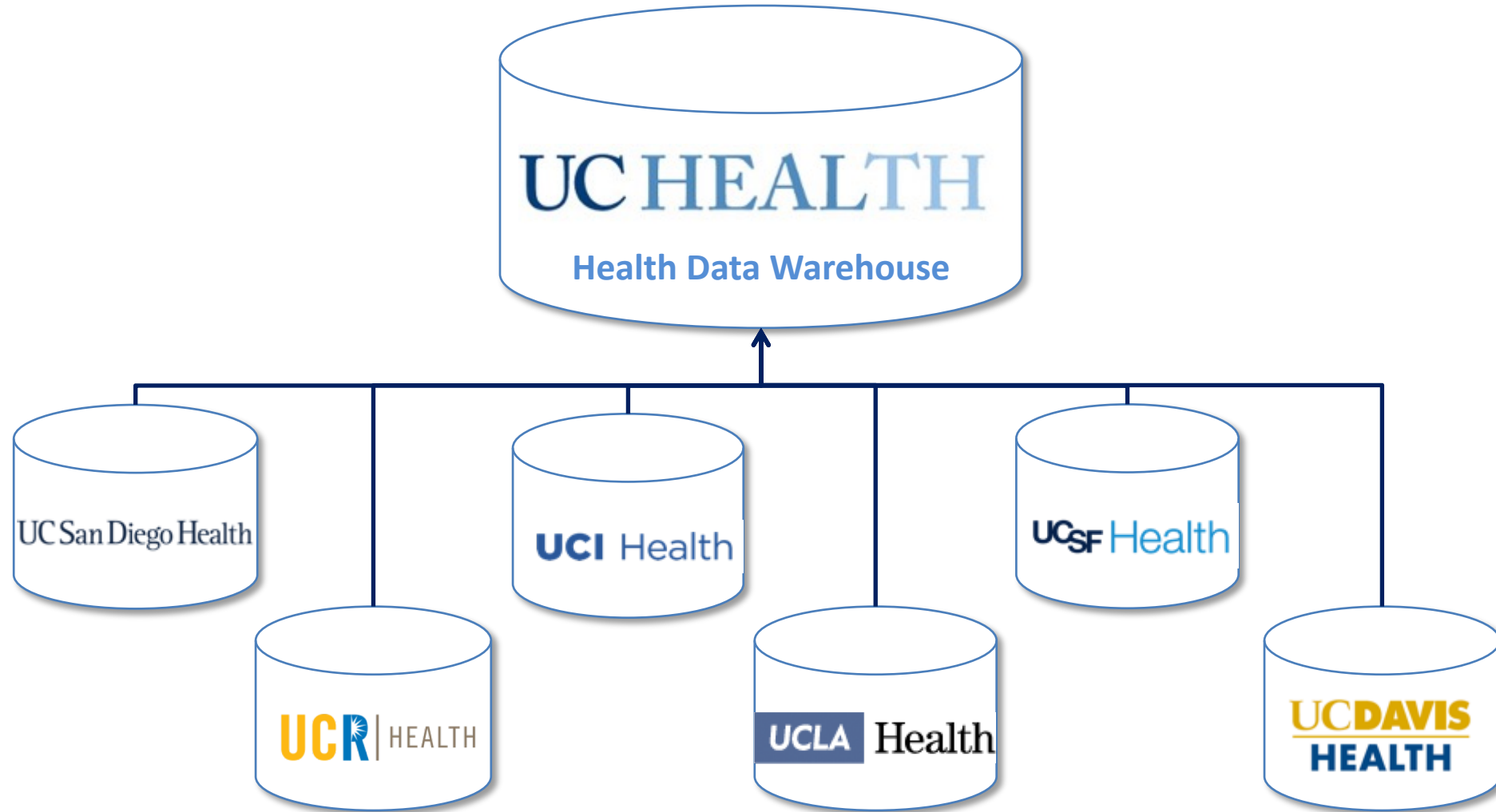
Aims for one file per person, fewer errors



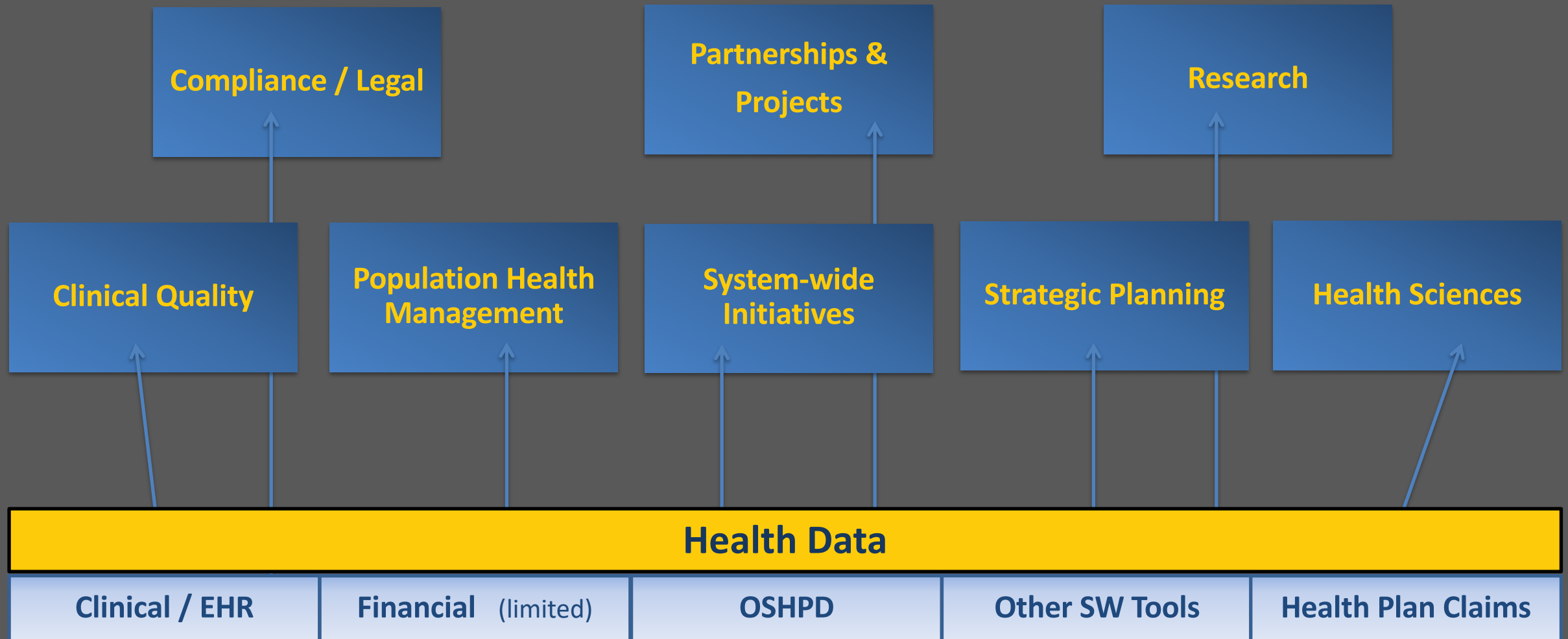
How Kaiser bet \$4 billion on electronic health records -- and won

Kaiser Permanente CIO Philip Fasano explains how electronic records have paid off and the health care giant's embrace of mobile technology

Combining healthcare data from across the six University of California medical schools and systems



Center for Data-Driven Insights and Innovation



Center for Data-Driven Insights and Innovation

UC HEALTH

Governance

Chief Data Scientist

Executive Director / Health Data Officer

Building / Maintaining Technical Infrastructure

- Supports efforts within clinical enterprise, ACOs, self funded plans:
- Compiling, structuring data (including from third parties)
 - Generating reports, visualizations
 - Interoperability with other campus systems

Data Science (Analytics)

- Working with business, clinicians, scientists – determine how to use the data strategically to derive the desired information in useful and accessible form
- Pioneer and develop novel data analytics and computational capabilities to drive value and insights (clinical, operational and research)

Data Governance

- Policies / Guidelines
- Lead Health Data Governance Committee to assess appropriateness of proposed projects
- Analyze positive and negative outcomes of projects UC has undertaken
- Share learnings, refine policies and practices

Strategy and Support

- Collaborate with campuses to develop strategy to inform, measure, drive our efforts
- Identify, accelerate projects and partnerships opportunities
- Facilitate resolution of internal conflicts and legal, ethical, reputational issues
- Project management

Data Platform

- Tableau – Data Visualization
 - Dashboards, Data Exploration, Trending
- Collibra – Data Governance
 - Data Definitions, Terminology Mapping
 - Value Set Management with Quality Measures
- OHDSI (Observational Health Data Science & Informatics)
 - OMOP Data Model
 - Standards-based model to harmonize data
 - Consortium of industry, government, and academia
- Microsoft SQL Server – Database
- Azure – Cloud-based user platform



UC has an unprecedented view of the medical system

- Combined data from UCSF, UCLA, UC Irvine, UC Davis, UC San Diego, and UC Riverside
- Total 15 million patients with an MRN, basic diagnosis data
- Central database built using OMOP (not Epic) as a data backend
 - Structured data from 2012 to the present day
 - **5.2 million patients with “modern” data**
 - 101M encounters, 307M procedures, 283M med orders, 684M vital signs, 453M lab test results, 422M diagnosis codes
 - Claims data from our self-funded plans now included
 - Continually harmonizing elements
- Quality and performance dashboards

University of California Cancer Consortium Takes on California's \$14 Billion Killer

By [Elizabeth Fernandez](#) on September 11, 2017



University of California President Janet Napolitano announced the formation of a UC Cancer Consortium to include five of the nation's leading academic cancer centers during a press conference in Genentech Hall at Mission Bay. *Photo by Susan Merrell*

What do actual cancer genetics reports look like?

UCSF Clinical Genomics

Patient: [REDACTED]

MRN: [REDACTED]

Ordering Provider: [REDACTED]

Surgical Pathologist: [REDACTED]

Electronically Signed-Out by: [REDACTED]

Pathogenic

VARIANT

NF1 c.205-1G>C, homozygosity resulting from germline mutation with accompanying somatic copy-neutral loss of heterozygosity

NF1 p.G629R

*Reads indicate the number of unique DNA molecules affected by the degree of normal cell contamination

Pathogenic

VARIANT

NF1 c.205-1G>C (splice site)

*Germline variants are reported if classified as pathogenic in ClinVar, truncating variants in well-established tumor databases of 1000G or ESP6500 datasets are reviewed only if they are small indels in gene coding regions.

INTERPRETATION

Identified in the submitted peripheral blood. This specific variant is not present in dbSNP ID, nor are there any classified as pathogenic in a functionally altered or inactivated gene. This variant is associated with neurofibromatosis type 1 (OMIM #105100), an autosomal recessive disorder characterized by café-au-lait freckling, Lisch nodules of the iris, and malignant neoplasms including diffuse astrocytomas, medullary thyroid

FOUNDATION ONE Home

Date of Birth [REDACTED]

Sex Male

FMI Case # [REDACTED]

Medical Record # [REDACTED]

Specimen ID Not Given

ABOUT THE TEST:

FoundationOne Heme™ is a next-generation sequencing test for hematologic malignancies.

PATIENT RESULTS¹

3 genomic findings

1 therapy associated with potential clinical benefit

0 therapies associated with lack of response

10 clinical trials

¹ Reduced sensitivity due to sample quality. Performance Specifications for details.

THERAPEUTIC IMPLICATIONS

Genomic Findings Detected

CCND3 amplification - equivocal

CDK4 amplification

MYC amplification

None

None

None

Note: Genomic alterations detected may be associated with a clinical trial, but have little or no evidence in the patient's tumor tissue. Predicted efficacy for this patient, nor are there any clinical trials identified.

For more comprehensive information please log on to the Interactive Cancer Explorer™

To set up your Interactive Cancer Explorer account, contact your sales representative or call (888) 988-3639.

Electronically Signed by [REDACTED], M.D. | [REDACTED]

Foundation Medicine, Inc. / 1-888-988-3639

FOUNDATION ONE Home

GENOMIC ALTERATIONS

GENE ALTERATION

INTERPRETATION

● CCND3 amplification - equivocal

● CDK4 amplification

● MYC amplification

Gene and Altered with and regulation inactivation overexpression

Frequency of osteosarcoma osteosarcoma amplification study⁷. Increased with leiomyosarcoma

Potential Target CDK4/6 inhibitor which are under investigation 2013; ASCO

Gene and Altered to cell cycle, promote cell cycle progression

Frequency of osteosarcoma osteosarcoma conventional osteosarcoma subtypes CDK4, and osteosarcoma

Potential Target such as LEE011 investigator Abstract 250

Gene and Altered to cell cycle, cancers²⁵. Malignant number of d mRNAs and p

Frequency of osteosarcoma osteosarcoma

THERAPIES

There are no therapies FDA-approved

ADDITIONAL THERAPIES – FDA-APPROVED

THERAPY

SUMMARY

Palbociclib

Approved in combination with endocrine therapy for ER-positive, HER2-negative breast cancer^{24,48,49}. Phase 2 trial of small cell lung cancer (8/16) and ASCO Abstract 400. Hepatocellular carcinoma patients, including ASCO GI Abstract 400. IGF1R^{50,51,52} resistance to

Genomic alterations detected may be associated with a clinical trial, but have little or no evidence in the patient's tumor tissue.

For more comprehensive information please log on to the Interactive Cancer Explorer™

To set up your Interactive Cancer Explorer account, contact your sales representative or call (888) 988-3639.

Electronically Signed by [REDACTED], M.D. | [REDACTED]

Foundation Medicine, Inc. / 1-888-988-3639

FOUNDATION ONE Home

Patient Name [REDACTED]

Patient, Test

Report Date [REDACTED]

Tumor Type PEDIATRIC Bone osteosarcoma

CLINICAL TRIALS TO CONSIDER

IMPORTANT: While every effort is made to ensure the accuracy of the information contained below, the information available in the public domain is continually updated and should be investigated by the physician or research staff. This is not meant to be a complete list of available trials. In order to conduct a more thorough search, please go to www.clinicaltrials.gov and use the search terms provided below. For more information about a specific clinical trial, type the NCT ID of the trial indicated below into the search bar.

GENE

RATIONALE FOR POTENTIAL CLINICAL TRIALS

● CCND3 amplification - equivocal

CCND3 amplification may predict sensitivity to CDK4/6 inhibitors. Due to the limited number of clinical trials available for pediatric patients, clinical trials recruiting adults only are listed below for informational purposes. Examples of clinical trials that may be appropriate for this patient are listed below. These trials were identified through a search of the trial website clinicaltrials.gov using keyword terms such as "CDK4", "CDK6", "PD 0332991", "LEE011", "LY2835219", "palbociclib", "child", "osteosarcoma", "solid tumor", and/or "advanced cancer".

TITLE

PHASE

TARGETS

LOCATIONS

NCT ID

Targeted Agent and Profiling Utilization Registry (TAPUR) Study

Phase 2

CDK4, CDK6, Others

Michigan, North Carolina

NCT02693535

Phase II Trial of the Cyclin-Dependent Kinase Inhibitor PD 0332991 in Patients With Cancer

Phase 2

CDK4, CDK6

Pennsylvania

NCT01037790

Abemaciclib in Children With Newly Diagnosed Diffuse Intrinsic Pontine Glioma, and in Children With Recurrent and Refractory Solid Tumors Including Malignant Brain Tumors

Phase 1

CDK4, CDK6

Georgia

NCT02644460

A Phase I Study of the CDK4/6 Inhibitor PD-0332991, 5-Fluorouracil, and Oxaliplatin in Patients With Advanced Solid Tumor Malignancies

Phase 1

CDK4, CDK6

District of Columbia

NCT01522989

Modular Phase II Study to Link Targeted Therapy to Patients With Pathway Activated Tumors: Module 8 - LEE011 for Patients With CDK4/6 Pathway Activated Tumors

Phase 2

CDK4, CDK6

Texas

NCT02187783

For more comprehensive information please log on to the Interactive Cancer Explorer™

To set up your Interactive Cancer Explorer account, contact your sales representative or call (888) 988-3639.

Electronically Signed by [REDACTED], M.D. | [REDACTED], M.D., Medical Director

Foundation Medicine, Inc. / 1-888-988-3639

Sample Preparation: 150 Second St., 1st Floor, Cambridge, MA 02141 / CLIA:22D2027531

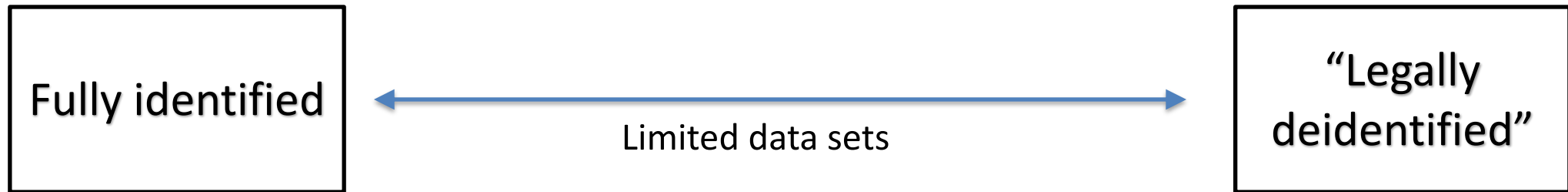
Sample Analysis: 150 Second St., 1st Floor, Cambridge, MA 02141 / CLIA:22D2027531

page 5 of 16

What could we do with clinical data?

- Clinical researcher at UCLA could run a genome wide association study across UC Health
- Mobile health researcher at UCSD can enable patients to contribute data for research
- Community activist and researcher UC Merced can study environmental factors contributing to health and disease
- Transplant patient at UC Irvine can download all their data across UC Health
- Data scientist at UC Santa Barbara can model development of Alzheimer's disease and build a multi-modal predictor
- App designer at UC Riverside can show patients their choices with chronic disease
- CMO at UCSF can build predictive models for readmission, test, share across UC Health
- AI researcher at UC Berkeley can build deep-learning models for image-based diagnostics
- Health services researcher at UC Davis can build predictive models for drug efficacy, and maybe enable pay-for-performance
- Cancer genomics researcher at UCSC can study all our clinical cancer genomes

How will research access work?



- Researchers should first write and optimize OMOP SQL queries locally, on their own campuses
- When ready to scale (and authorized), we spin up a virtual machine for the researcher, populated with common tools
 - Docusign a UC Health data use agreement
 - R, Tableau, Jupyter Notebooks, Julia, SQL, Windows or Linux
- Upload your scripts and run, but cannot download data

Campus OMOP Resources

- UCSF
- UC Davis
- UCLA
- UC Irvine
- UCSD
- Others



UC San Diego
Altman Clinical and Translational Research Institute

Altman Clinical and Translational Research Institute

HOME ABOUT CLINICAL RESEARCH INFORMATICS & STATS LABORATORIES EDUCATION & TRAINING COMMUNITY

School of Medicine / Research / Altman Clinical and Translational Research Institute / Informatics & Stats / Databases and Clinical Records

Informatics & Stats

- ACTRI Biostatistics
- Databases and Clinical Records
- ▶ Velos/REDCap
- Center for Computational Biology and Bioinformatics
- Accrual to Clinical Trials (ACT)
- ▶ UC-ReX
- Informatics Recharge Rates
- Biostatistics Recharge Rates

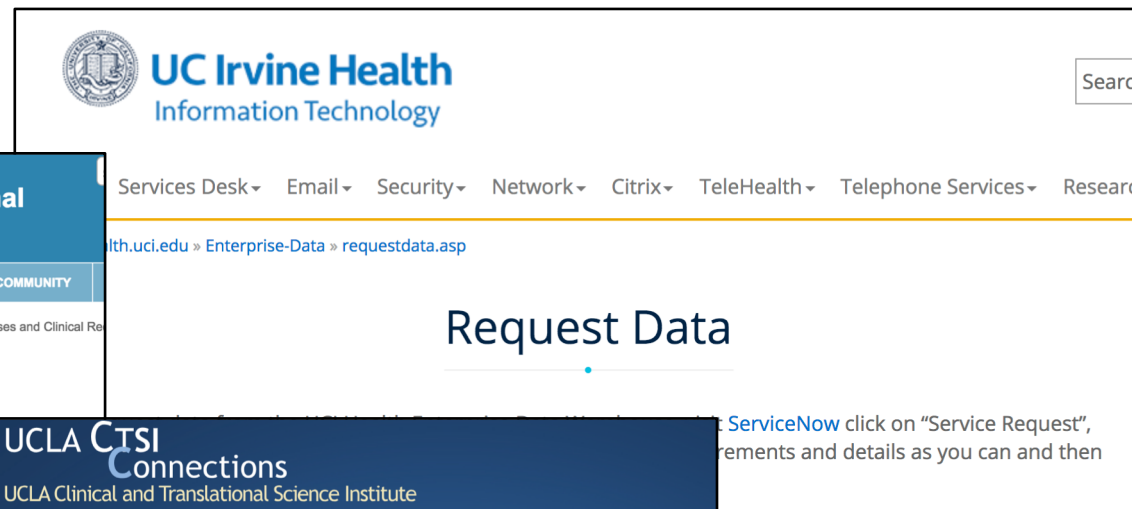
Databases and Clinical Records

[Request ACTRI Services](#)

Clinical Trials and Sample Repository

Velos eResearch is an integrated software system...
Electronic Medical Record System to provide improved...
module within this platform, called eSample, will...
robust support team assists investigators in implementing...
a web-based system that supports

- Creation of budgets
- Creation of protocols with calendars
- Scheduling of subjects
- Electronic case report forms



UC Irvine Health
Information Technology

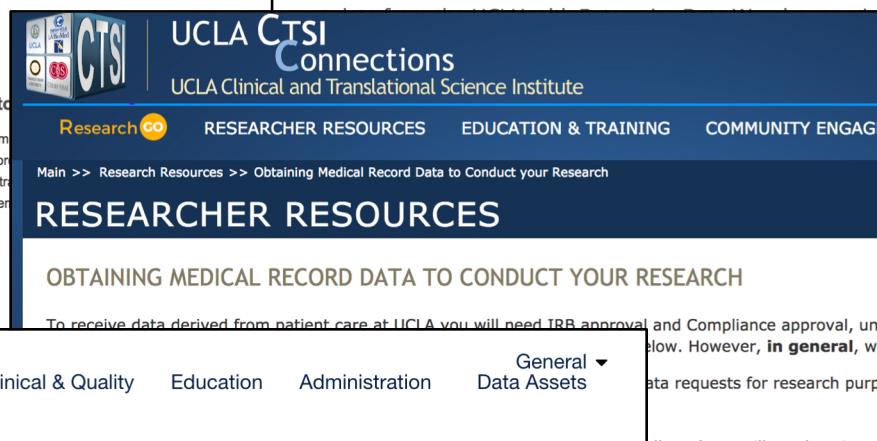
Search

Services Desk Email Security Network Citrix TeleHealth Telephone Services Research

Health.uci.edu » Enterprise-Data » requestdata.asp

Request Data

ServiceNow click on "Service Request",
requirements and details as you can and then



UCLA CTSI
Connections
UCLA Clinical and Translational Science Institute

Research GO RESEARCHER RESOURCES EDUCATION & TRAINING COMMUNITY ENGAGEMENT

Main >> Research Resources >> Obtaining Medical Record Data to Conduct your Research

RESEARCHER RESOURCES

OBTAINING MEDICAL RECORD DATA TO CONDUCT YOUR RESEARCH

To receive data derived from patient care at UCLA you will need IRB approval and Compliance approval, unless otherwise specified. However, **in general**, we will process data requests for research purposes.



UCSF Data Resources

Research Clinical & Quality Education Administration General Data Assets

UCDAVIS HEALTH Clinical and Translational Science Center

Home CTSC Programs Request Services Investigator Resources News About Us Contact Us

UC Davis Health / Clinical and Translational Science Center / Biomedical Informatics / EMR Data Retrieval for Research

Biomedical Informatics

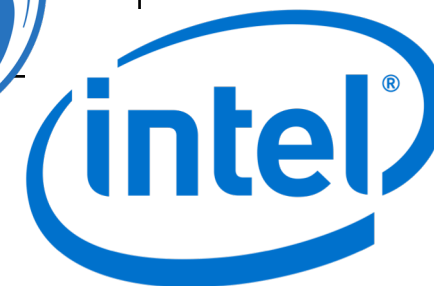
- Consultation
- EMR Data Retrieval for Research**

EMR Data Retrieval for Research

UC Davis researchers can acquire patient data for research from the electronic medical record. The type of data needed dictate the process to access the data. It is necessary to understand the sources of data, and process to gain access to data.

Data powers UCSF work
research, education and a

UCSF, GE Healthcare Launch Deep Learning Partnership to Advance Care Globally



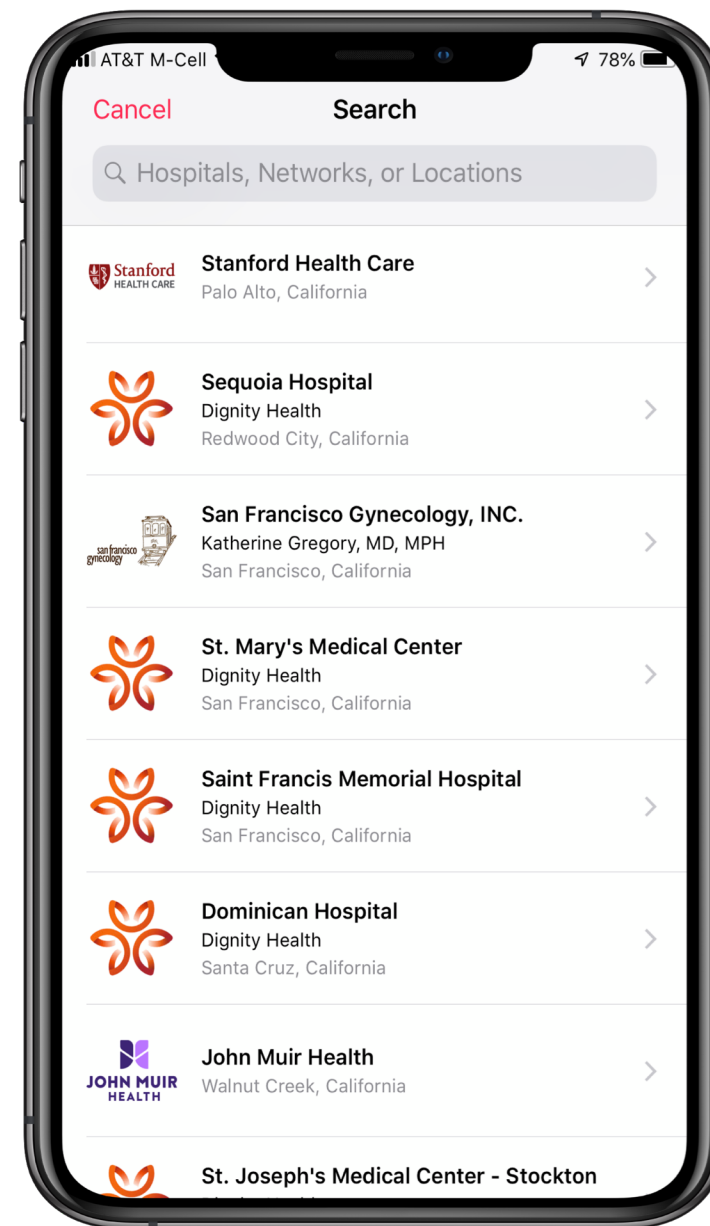
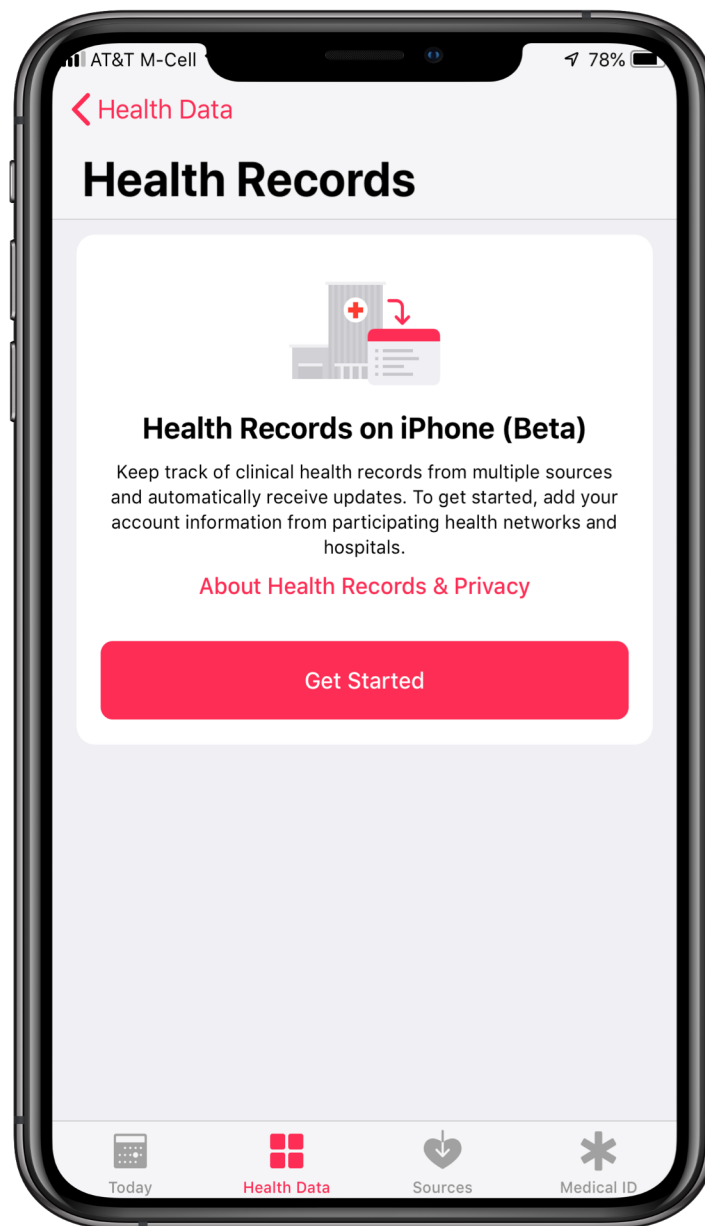
UCSF, Intel Join Forces to Develop Deep Learning Analytics for Health Care

Hello Kitty: How Pictures of Cats Help Computers Read Chest X-Rays

Google is training computers to predict when you might get sick

- Google Brain is working with top hospitals to predict health outcomes from medical data.
- That data was stripped of personally-identifiable information and shared with Google.
- This is the latest in a series of research projects applying Google's machine learning expertise to health care.





UC Clinical Data Warehouse Team

Executive Team

- Atul Butte
- Joe Bengfort
- Michael Pfeffer
- Tom Andriola
- Chris Longhurst

Steering Committee

- Lisa Dahm
- Mohammed Mahboub
- David Dobbs
- Kent Andersen
- Ralph James
- Jennifer Holland
- Eugene Lee

ETL Team

- Albert Dugan
- Tony Choe
- Michael Sweeney
- Timothy Satterwhite
- Ayan Patel
- Niranjana Wagle
- Ralph James
- Joseph Dalton

Data Harmonization

- Dana Ludwig
- Daniella Meeker

Data Quality

- Momeena Ali
- Jodie Nygaard

Business Analyst

- Ankeeta Shukla

Hardware

- Sandeep Chandra
- Jeff Love
- Scott Bailey
- Kwong Law
- Pallav Saxena

Support

- Elizabeth Engel
- Jack Stobo
- Michael Blum
- Sam Hawgood

Collaborators

- Alejandro Sweet-Cordero, Julien Sage / Pediatric Oncology
- Elizabeth Thomson, Patrick Dunn / Northrop Grumman
- Gabe Rosenfeld, Quan Chen / NIAID
- Andrei Goga / UCSF Oncology
- Mallar Bhattacharya / UCSF Pulmonary
- Minnie Sarwal / UCSF Nephrology
- Geoffrey Gurtner / UCSF Surgery
- Roberta Diaz Brinton / Arizona
- Carol Bult / Jackson Labs
- David Stevenson, Gary Shaw / Stanford Neonatology
- Takashi Kadowaki, Momoko Horikoshi, Kazuo Hara, Hiroshi Ohtsu / U Tokyo
- Kyoko Toda, Satoru Yamada, Junichiro Irie / Kitasato Univ and Hospital
- Shiro Maeda / RIKEN
- Mark Davis, C. Garrison Fathman / Immunology
- Russ Altman, Steve Quake / Stanford Bioengineering
- Euan Ashley, Joseph Wu / Stanford Cardiology
- Mike Snyder, Carlos Bustamante, Anne Brunet / Stanford Genetics
- Jay Pasricha / Stanford Gastroenterology
- Rob Tibshirani, Brad Efron / Stanford Statistics
- Hannah Valantine, Kiran Khush / Stanford Cardiology
- Mark Musen, Nigam Shah / National Center for Biomedical Ontology
- Sam So, Ken Weinberg, David Miklos / Stanford Oncology

Support

- University of California, San Francisco
 - Priscilla Chan and Mark Zuckerberg
 - NIH: NIAID, NLM, NIGMS, NCI, NHLBI, OD; NIDDK, NHGRI, NIA, NCATS, NICHD
 - California Governor's Office of Planning and Research
 - March of Dimes
 - Juvenile Diabetes Research Foundation
 - Howard Hughes Medical Institute
 - California Institute for Regenerative Medicine
 - Hewlett Packard, L'Oreal, Progenity
 - Luke Evnin and Deann Wright (Scleroderma Research Foundation)
 - Clayville Research Fund, PhRMA Foundation
 - Stanford Cancer Center, Bio-X, SPARK
-
- Tarangini Deshpande
 - Kimayani Butte
 - Sam Hawgood, Keith Yamamoto, Jack Stobo, Talmadge King
 - Isaac Kohane

Admin and Tech Staff

- Mary Lyall
- Mounira Kenaani
- Kevin Kaier
- Boris Oskotsky
- Mae Moredo
- Ada Chen