Precisely practicing medicine from 1.1 billion points of UC health data

Atul Butte, MD, PhD
Executive Director for Clinical Informatics, UC Health
Executive leadership at KP has approved significant funding of $45M to centralize and extend research biobanking across KP. KP offers an attractive environment for research bio-banking: one of the largest, most diverse, member bases in the world to recruit from, a fully integrated care system to interface with, and members who not only tend to stay with KP for years but also receive almost all of their care through the system, making KP’s EMR data an invaluable research resource. Further, the new entity will benefit from existing efforts, including KP’s landmark Research Program on Genes, Environment, and Health (RPGEH), which includes DNA and blood samples from over 220,000 KP members… cohorts in pregnancy and cancer. KP’s substantial investment will help to recruit a total of 500,000 members to participate… KP’s Research Bank will likely be the most robust of its kind in the US, with its inclusion of comprehensive longitudinal health data.
Cohorts of patients are extremely valuable for science and discovery, but we need to act as one.
Combining healthcare data from across the six UC medical schools and systems
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.
Thursday, October 13, 2016

NIH funds additional medical centers to expand national precision medicine research program

The awardees are:

- California Precision Medicine Consortium: University of California San Diego, with partners Cedars-Sinai Medical Center, Los Angeles; San Diego Blood Bank; University of California, Davis; University of California Health; University of California, Irvine; University of California, San Francisco; and University of Southern California, Los Angeles
data. With this collective potential to collect, aggregate, integrate, share and analyze vast and diverse datasets, these partnerships could enable discoveries to transform cancer prevention, detection and diagnosis, accelerate therapeutic development, and advance a patient-centered learning health care system. The overarching vision includes a virtual data ecosystem (drawing from the UHealth Data Warehouse, which brings together ~15 million patient records).
First stage (proof of concept) Completed

- Data warehouse includes all encounter based data elements
  - Built using Epic Cogito data model framework
  - Industry standard architecture to be used for data export to UCDW
  - Implement appropriate security and privacy controls and provide tools to assure data integrity and quality

- Success Metrics: Successfully integrating data from local Cogito tables across all 5 UCs into a central data warehouse
  - Final data set will include harmonized data with uniform coding patterns mapped across all local sites, with appropriate data quality and security checks in place

- Validation question: UC primary care on appropriate statin therapy

- Total count 15 million patients; ~600,000 primary care patients
  - Lab results: 800+ million
  - Medication Orders: 76+ million
  - Encounters: nearly 200 million
Why is a governance model needed now?

• Other parallel systems have been developed using data from across UC Health, but have been limited in scope - research projects, or in specific clinical domains
  – ATHENA, PCORI pScanner, NIH Precision Medicine

• The intent for BHDI is to view UC health data as an asset and create the ability to use the data for multiple aims
  – Identifiable data: improve operations and quality
  – De-identified data: clinical research, cost containment, patient aggregation, external research and clinical partnerships

• BHDI is starting with EHR data, but the intent is to grow to other sources of clinical data (e.g., images, patient-reported)
Designing the governance model

The BHDI presentation on the opportunity and potential uses of the data has been delivered broadly across the UC’s to administration, clinical & research communities.

The specific of the governance model has had input and vetting:
- Proof of Concept Steering Committee
- UC BRAID
- UC CIO/CMIO
- UC Faculty Senate
- UC Chief Medical Officers
- UC EVP Health Services & UC Health System CEOs

**Goal - Manage the UC Health clinical data asset and its use for clinical and research operations for the University of California**
Two-Level Governance Structure

• Executive Committee
  – Small, nimble group
  – Decision making; creates the consent agenda for the Board where appropriate
  – Diverse representation, rotating seats every 3 years

• Board
  – Larger (full) representation
  – Full ratification of critical decisions; important issues requiring broad input

• Task Forces
  – As needed
Governance Board Membership

- Goal of broad participation from all locations as well as from domain areas that have an interest and/or important perspective
- Health systems’ representatives from
  - Chief Medical Officers
  - Chief Nurse Officers
  - Chief Information Officers
  - Chief Medical Informatics Officers
  - Chief Data Officer/Equivalent
  - Chief Security Information Officers
  - Patient Advocate
- UC BRAID and REX Reps from each campus
  - Computational science expertise
- Faculty at-large representing the education mission
- Ethics community
- Non UC-Health Campus (placeholder as board evolves)
Governance: Executive Committee Membership (7 people)

- UC Clinical Informatics Director
- Three UC Health System Rep(s)
  - CMO
  - CIO
  - CMIO
- UC BRAID Rep
- Patient Advocate Rep
- UC Chief Information Officer
Governance: Structure

Executive Committee & Board
- Prioritization
- Staffing
- Compliance
- Program Monitoring
- Planning
- Budgeting
- Enterprise Communication
- Issue Resolution

Compliance & Security Task Force
- Data/Solution Access
- Auditing
- Security and Privacy
- Policies & Procedures
- IRB Coordination
- Collaboration Agreements
- Honest Broker Process
- Compliance Coordination & Monitoring

Analytics/Stewardship Task Force
- Needs Identification
- Metric Definition
- Business Rules
- Data Quality Program
- Enterprise Data Standards
- Common Data Definitions
- Quality Assurance

Architecture and Support
- Solution Design & Build
- Governance Support
- Stewardship Support
- Program Management
Next Steps

Short term: 6 months
• Update ETLs
• Run on periodic basis
• Include text notes
• Central data quality/analytics
• Drive encryption
• Keep month-to-month at SDSC until move

Medium term: 6 months
• OMOP backend
• De-identify data for clinical research
• Re-establish harmonization efforts

Long term
• Move server to Quincy
• Include payment data, claims data, third party data
• Include images and image archive
• Include send-out tests including genomics
• Link to biobanks
• Incorporate claims data for encounters outside UC Health
• Patients uploading and downloading their own data
• Include research data
• Enabling CRO-like activities across UC Health
What could we do with the 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 Modesto 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
UC Clinical Data Warehouse Team

Executive Team
• Atul Butte
• Joe Bengfort
• Michael Pfeffer
• Tom Andriola
• Chris Longhurst

Steering Committee
• Irfan Chaudhry
• Mohammed Mahbouba
• Lisa Dahm
• David Dobbs
• Kent Andersen
• Ralph James
• Jennifer Holland
• Eugene Lee

Epic
• Kevin Ames
• Ben Jenkins
• Steve Gesualdo

ETL Team
• Albert Dugan
• Tony Choe
• Michael Sweeney
• Timothy Satterwhite
• Ayan Patel
• Niranjan 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
• Jack Stobo
• Michael Blum
• Sam Hawgood