Updates on UC-wide Data Initiatives

BRAID Retreat
October 2017
A controversial electronic health records system that cost Sutter Health $1 billion went completely dark Monday at Sutter Medical Center, Sacramento. The system imposed an additional risk beyond problems reported with its predecessor, and the new system also faced pushback from employees.

For several years, Sutter, one of the largest health care systems in California, has been phasing out its $4.2 billion computer system over 11 years and replacing it with a system from Oracle that was expected to go live after the end of July.

Sutter soaked up $1 billion over 11 years to install the Oracle system, which was expected to be a big improvement over the old one. But the new system also faced pushback from employees and customer dissatisfaction with its performance.

Partners’ $1.2 billion EHR system
Aims for one file per person

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

InfoWorld | May 2, 2013

In July 1907, the first great breakthrough in medical IT took place at the Mayo Clinic in Rochester, Minn.: the paper medical record, dropped into a paper folder and stored in a file cabinet. Until then, information on patients was kept in a ledger that recorded all of a day’s patient visits, one after the other. Different departments kept separate ledgers, making it extremely difficult to track down patient information in a timely manner.
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 best-in-class electronic medical records (EMR) for each member. KP members not only tend to stay with KP for years, they also tend to 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.

High level of interest and investment in health system data
New value of using properly labeled or correctly treated patients for machine learning
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 before it was shared with Google.
- This is the latest in a series of research projects from Google to apply its machine learning expertise to health care.

Christina Farr | @chrissyfarr
Published 3:14 PM ET Wed, 17 May 2017 | Updated 7:43 PM ET Wed, 17 May 2017
106 STARTUPS TRANSFORMING HEALTHCARE WITH AI

PATIENT DATA & RISK ANALYTICS
- ROAM
- APIXIA
- Lumina
- Enosdata
- Medalogix

RESEARCH
- Sjane Healthcare Technologies
- Arteryx
- Enticic
- BayLabs
- Zebra

MEDICAL IMAGING & DIAGNOSTICS
- Maxwest
- Arteryx
- Enricic
- BAYLABS
- Zebra

LIFESTYLE MANAGEMENT & MONITORING
- AiCure
- PeerWell
- Wellframe
- Luma
- SkinVision
- Welltok
- Intendu
- Ovia
- GamersHealth

NUTRITION
- Nurtas
- Vitl

EMERGENCY ROOM & SURGERY
- Gauss Surgical
- Medica
- P3d

IN-PATIENT CARE & HOSPITAL MANAGEMENT
- Qualaris
- AEC
- analyticsMD
- Ivion

MISCELLANEOUS
- Digital Reasoning
- Saykara
- HealthBotics
- Metro
- NarrationDX
- AYASDI
- HealthData

WEARABLES
- Touchkin
- physIQ
- TinyKicks

VIRTUAL ASSISTANTS
- Buoy
- SenseLV
- Babylon
- Sophie
- medRX

MENTAL HEALTH
- TAO
- Ginger IO
- AnaLog

DRUG DISCOVERY
- Atomwise
- Good Pharma
- BenevolentAI

VIRTUAL ASSISTANTS
- Buoy
- SenseLV
- Babylon
- Sophie
- medRX

WEB EXPLORATION
- istock.com/hitch
Historical context

- **BRAID**: Collaborative foundation
- **UC ReX**: Enabled counting patients across the system, harmonization of data elements
- **Other UC-wide projects**: PCORI pScanner, ATHENA, NIH Precision Medicine
- **Health systems**: Launched LSfV
- **Nov-2014 retreat**: BRAID proposal & Precision Medicine: seeded the need to come together
- **Feb-2015 workshop**: Convened UC healthcare leaders
- **Mar-2015 retreat**: BHDI proposal to UC Health leadership to recognize the potential
- **Aug-2015 – May 2016**: Proof of Concept
  - $1.25 million funded through startup funds from UCSF
  - Resource contributions from all five sites
- **May-2016**: Update presentation to Health Retreat
- **Oct-2016**: Funding released for Phase 2
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.
Bringing UC’s clinical data together creates an asset with few peers

A Big UC Healthcare Data Analytics Platform

Confidential, Deliberative Process, CPRA Exempt – DRAFT Preliminary Results
UC-wide Clinical Data Warehouse (CDW) Proof of Concept (POC)

- Includes encounter-based data across five UC medical centers
- Built using Epic Cogito data model
- Implemented appropriate security and privacy controls
  - SDSC, Database-level encryption
- Total count 15 million patients; 600,000 primary care patients
  - Lab results: 803,060,636
  - Medication Orders: 76,866,799
  - Encounters: 196,631,625
- Validation question: UC primary care on appropriate statin therapy
Predicting Drug Efficacy in Type 2 Diabetes Patients at UCSF
### Healthy eating, weight control, increased physical activity, and diabetes education

<table>
<thead>
<tr>
<th>Monotherapy</th>
<th>Metformin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Efficacy</td>
<td>high</td>
</tr>
<tr>
<td>Hypo risk</td>
<td>low risk</td>
</tr>
<tr>
<td>Weight</td>
<td>neutral / loss</td>
</tr>
<tr>
<td>Side effects</td>
<td>GI / lactic acidosis</td>
</tr>
<tr>
<td>Costs</td>
<td>low</td>
</tr>
</tbody>
</table>

*If A1C target not achieved after ~3 months of monotherapy, proceed to 2-drug combination (order not meant to denote any specific preference—choice dependent on a variety of patient- and disease-specific factors).*

<table>
<thead>
<tr>
<th>Metformin +</th>
<th>Metformin +</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sulfonylurea</td>
<td>Thiazolidinedione</td>
</tr>
<tr>
<td>high</td>
<td>high</td>
</tr>
<tr>
<td>moderate risk</td>
<td>low risk</td>
</tr>
<tr>
<td>Weight</td>
<td>gain</td>
</tr>
<tr>
<td>Hypoglycemia</td>
<td>edema, HF, fx, s</td>
</tr>
<tr>
<td>Costs</td>
<td>low</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Metformin +</th>
<th>Metformin +</th>
</tr>
</thead>
<tbody>
<tr>
<td>DPP-4 inhibitor</td>
<td>SGLT2 inhibitor</td>
</tr>
<tr>
<td>intermediate</td>
<td>intermediate</td>
</tr>
<tr>
<td>Weight</td>
<td>low risk</td>
</tr>
<tr>
<td>Hypoglycemia</td>
<td>rare</td>
</tr>
<tr>
<td>Costs</td>
<td>high</td>
</tr>
</tbody>
</table>

*If A1C target not achieved after ~3 months of dual therapy, proceed to 3-drug combination (order not meant to denote any specific preference—choice dependent on a variety of patient- and disease-specific factors).*

<table>
<thead>
<tr>
<th>Metformin +</th>
<th>Metformin +</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sulfonylurea</td>
<td>Thiazolidinedione</td>
</tr>
<tr>
<td>or TZD</td>
<td>or DPP-4-i</td>
</tr>
<tr>
<td>or SGLT2-i</td>
<td>or GLP-1-RA</td>
</tr>
<tr>
<td>or Insulin</td>
<td>or GLP-1-RA</td>
</tr>
<tr>
<td>or Insulin</td>
<td>or GLP-1-RA</td>
</tr>
</tbody>
</table>

*If A1C target not achieved after ~3 months of triple therapy and patient (1) on oral combination, move to injectables; (2) on GLP-1-RA, add basal insulin; or (3) on optimally titrated basal insulin, add GLP-1-RA or mealtime insulin. In refractory patients consider adding TZD or SGLT2-i.*

<table>
<thead>
<tr>
<th>Metformin +</th>
<th>Metformin +</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basal insulin</td>
<td>Mealtime insulin</td>
</tr>
<tr>
<td>or GLP-1-RA</td>
<td>or GLP-1-RA</td>
</tr>
</tbody>
</table>

Source: American Diabetes Association Standards of Medical Care in Diabetes (2016)
UCSF Cohort Selection: Inclusion Criteria

• Must be >= 18 years old at the index point

• At least 90 days of follow-up after the index point
  • Must refill medication or have a new medication prescribed

• Must have an ICD10 code for Type 2 Diabetes (any time in medical history)

• No ICD10 codes for pregnancy in EHR (any time in medical history)
Medication Strategies for First-Time Type 2 Diabetes Patients

Type 2 Diabetes Medication Strategies
(11,185 First-Time Patients)
Medication Strategies for First-Time Type 2 Diabetes Patients

Type 2 Diabetes Medication Strategies (11,185 First-Time Patients)
There are 1,286 Unique Trajectories for UCSF T2D Patients
There are **1,286** Unique Trajectories for UCSF T2D Patients
Performance for Predicting Medication Class Increase Within 90 Days from Metformin

A. Model Comparison for Predicting Medication Class Increase Within 90 Days (Metformin Only)

B. Decision Tree Used In Classification (AUC 0.607)

Largest HbA1c
- Yes
  - Ever Measured < 8.8
  - No

Largest Fasting Glucose
- Yes
  - Ever Measured < 206
  - No

Odds Ratio: 2.14
n=371 14%

Odds Ratio: 0.988
n=2167 84%

Odds Ratio: 2.10
n=50 2%
Precision Medicine In Practice

Patient Needs New Medication → Predict Efficacy of Medications Being Considered for this Patient → Ranked List of Medications & Doses Likely To Be Effective For This Patient
F.D.A. Approves First Drug to Treat Severe Multiple Sclerosis

By KATIE THOMAS  MARCH 28, 2017

Jerrie Gullick, 51, was part of the clinical trial for Ocrevus and said the drug had significantly slowed the progression of her primary progressive multiple sclerosis. Sam Hodgson for The New York Times
Text De-identification

- Initial corpus: manually annotated 20 randomly selected UCSF notes from 10 departments
- Our de-identification method achieved excellent generalizable recall (UCSF 99.6%) while maintaining the precision necessary (UCSF 87%) to be useable for downstream research.
- Additionally, we are in the process of hiring 2 UCSF resident physicians to annotate an additional 200 of UCSF's clinical notes for PHI.
- We are ready to run the method on all 40 million of UCSF's notes and are simply awaiting approval and the use of machine resources.
What is an Enterprise Image?

The Information Commons

- To provide a shared repository of data and tools for integrative research, that allows for
  - Multifactor and multimodal data
  - Truly large scale (eg for machine learning models)
  - Compliance-ready (de-id, limited)
  - Integrates basic science, clinical, population and published data or information
Research Computing Capability for Integrative Research

Driving Research: Precision Medicine, Integrative Spanning Basic, Translational, Clinical, Population, …

Core Expertise

Common Core Knowledge Network
Graph theoretic representation of key biomedical pathways

Information Commons
“Data Lake,” Core Tools
Massively parallel, open source, federated where needed

Virtual Compute Cluster
Several Profiles: Big Data, Big Compute, Graphics ….
Our Vision of a Durable Information Commons for all UC

Information Commons

“Compute to the Data Lake” (Spark/Hadoop)
De-Identified & IRB-restricted
Premise based and Cloud pilot

Core Tools: De-ID, NLP, Annotation, Data Quality…

Analysis support: R, MLLib, DLLib, Tensorflow…

Radiology Imaging Sci
Tumor Studies
Infectious Diseases
Neurology - Oncology Spark Know Net

Pubs Textual Basic Research Clinical Specialty User reported Omics Therapeutics Billing Population Public

... Research disciplines