Integrating Genomic, Imaging, and Clinical Data for Precision Medicine

Shawn Murphy MD, Ph.D.
Partners Healthcare and Harvard Medical School
October 4th, 2017
Personalized Medicine and Genomic technology are critical to managing populations

- Managing a population involves improving health outcomes of the group as a whole by identifying, monitoring and addressing health needs of individuals through:
  - Subpopulation stratification
  - Targeted, evidence-based treatment protocols
  - Predictive analytics

Source: Personalized Medicine Coalition and innovation.org; Oliver Wyman
Example: PPARγ Pro12Ala and Diabetes

Estimated risk (Ala allele)

Overall $P$ value = 2 x $10^{-7}$
Odds ratio = 0.79 (0.72-0.86)

Ala is protective

Courtesy J. Hirschhorn
High Throughput Methods for supporting Translational Research

- Set of patients is selected from medical record data in a high throughput fashion

- Investigators explore phenotypes of these patients using i2b2 tools and a translational team developed to work specifically with medical record data

- Distributed networks cross institutional boundaries for phenotype selection, public health, and hypothesis testing

- Personalized medicine is delivered into clinical care
High Throughput Methods for supporting Translational Research

- Set of patients is selected from medical record data in a high throughput fashion

- Investigators explore phenotypes of these patients using i2b2 tools and a translational team developed to work specifically with medical record data

- Distributed networks cross institutional boundaries for phenotype selection, public health, and hypothesis testing

- Personalized medicine is delivered into clinical care
1) Queries for aggregate patient numbers

- Warehouse of in & outpatient clinical data
- 6.7 million Partners Healthcare patients
- 2.5 billion diagnoses, medications, procedures, laboratories, & physical findings coupled to demographic & visit data
- Authorized use by faculty status
- Clinicians can construct complex queries
- Queries cannot identify individuals, internally can produce identifiers for (2)

2) Returns identified patient data

- Start with list of specific patients, usually from (1)
- Authorized use by IRB Protocol
- Returns contact and PCP information, demographics, providers, visits, diagnoses, medications, procedures, laboratories, microbiology, reports (discharge, LMR, operative, radiology, pathology, cardiology, pulmonary, endoscopy), and images into a Microsoft Access database and text files.

Research Patient Data Registry exists at Partners Healthcare to find patient cohorts for clinical research
FINDING PATIENTS

Query items

Person who is using tool

Query construction

Results - broken down by number distinct of patients
Please enter your IRB protocol.

Partners IRB (required): 2002P000381
Title: Research Patient Data Registry (RPDR)
Status: Active - Ongoing

Newton Wellesley Hospital IRB:

Spaulding Rehabilitation Hospital IRB:

North Shore Medical Center IRB: NSM 2008-786 demo
Title: 
Status:

Options for returned set of patients:
- [ ] Exclude Partners Healthcare employees
- [ ] Create a static set of patients from this query that can be used in other RPDR queries
- [x] Rerun the base query shown above to obtain a fresh set of patients

Step 3
Detailed data is gathered from RPDR and other Partners sources. Output files are placed in a special directory. Data is gathered from RPDR and other Partners sources. Files include a Personal Database.
The Partners Biobank

➢ The Partners Biobank provides samples (plasma, serum, and DNA) collected from consented patients.

➢ 64,000 patients have consented to date

➢ Samples are available for distribution to Partners investigators* to help identify novel Personalized Medicine opportunities that reduce cost and provide better care

*with required approval from the Partners Institutional Review Board (IRB).

Research Discoveries

Improve Clinical Care for All Patients
Biobank Integrative Genomics Strategy

Partners BioBank Samples
(Whole Blood Extracted DNA/RNA)

Genotyping
- Illumina MEGArray:
  Multi-Ethnic GWAS/Exome SNP Array
  Array Cost: $59/sample

Transcriptome
- Whole Transcriptome Analysis:
  RNA-seq
  Array Cost: $40-50/sample

Epigenome Profiling
- Methylation Analysis:
  HumanM450K Array
  Array Cost: $150/sample

Genome/Transcriptome Analysis: ~$100/sample

Genome/Transcriptome/Epigenome Analysis: ~$260/sample
Partners Personalized Medicine Components

Partners Biobank
DNA, Plasma & Serum from 47,000 Consented Patients
Sample Processing Services

Phenotype Core/Research
Patient Data Repository
6.7 million Partners patients
Data from EMR
Additional research data
Validated phenotypes/controls
Genomic Data Repository (GDR)

Translational Genomics CORE
Sequencing
Genotyping
Microarray
RNA Seq.

Laboratory for Molecular Medicine
Sanger Sequencing
Targeted Next Gen Seq
Whole Genome Seq (WGS)
Medical Exome (WES)

PPM Information Technology /IT, Bioinformatics, Research, Administrative

Partners Biobank
DNA, Plasma & Serum from 47,000 Consented Patients
Sample Processing Services

Translational Genomics CORE
Sequencing
Genotyping
Microarray
RNA Seq.

Phenotype Core/Research
Patient Data Repository
6.7 million Partners patients
Data from EMR
Additional research data
Validated phenotypes/controls
Genomic Data Repository (GDR)

Laboratory for Molecular Medicine
Sanger Sequencing
Targeted Next Gen Seq
Whole Genome Seq (WGS)
Medical Exome (WES)

PPM Information Technology /IT, Bioinformatics, Research, Administrative
High Throughput Methods for supporting Translational Research

- Set of patients is selected from medical record data in a high throughput fashion

- Investigators explore phenotypes of these patients using i2b2 tools and a translational team developed to work specifically with medical record data

- Distributed networks cross institutional boundaries for phenotype selection, public health, and hypothesis testing

- Personalized medicine is delivered into clinical care
RPDR Evolved into international “Informatics for Integrating Biology and the Bedside (i2b2)” sponsored by the National Institutes of Health, what is it?

- Software for explicitly organizing and transforming person-oriented clinical data to a way that is optimized for clinical genomics research
  - Allows integration of clinical data, trials data, and genotypic data
- A portable and extensible application framework
  - Software is built in a modular pattern that allows additions without disturbing core parts
  - Available as open source at https://www.i2b2.org
I2b2 Community Software Modules contributed as “Cells”

New i2b2 Community Projects

i2b2 FHIR Cell
Built by Kavishwar Waghokar, this i2b2 addition allows SMART cells to communicate with the i2b2 core using the Fast Healthcare Interoperability Resources.

C3-PRO FHIR Uploading Cell
C3-PRO Research Framework is an iOS framework written in Swift. Combining FHIR and ResearchKit, usually for data storage into i2b2, this framework allows you to use FHIR Questionnaire resources directly with ResearchKit and will return FHIR QuestionnaireResponse that you can send to your server. In addition, a FHIR Contract resource can be used to carry trial eligibility requirements and define content to be shown during consenting. Subsequently, the contract can be “signed” with a FHIR Patient resource and returned to your server, indicating consent.

Current i2b2 Community Projects

Loyalty Cohorts
Because electronic health records are often missing information about patients, we developed and validated a tunable computer algorithm to identify subsets of patients whose data are relatively complete and therefore better suited for clinical research studies.

Workplace Items Sharing Enhancement
A collection of webclient plugins that Enhance the Sharing of Items within the Workplace pane

Ontology Tools
Tools to extract and organize ontologies. The tools are organized by Lori Phillips. Recent additions is a library of ontologies which can be downloaded using the i2b2 workbench.
i2b2 Cell: The Canonical Software Module

- Business Logic
- Data Access
- Data Objects

i2b2

HTTP XML
(minimum: RESTful)
An i2b2 Environment (the Hive) is built from i2b2 Cells

“Hive” of software services provided by i2b2 cells
I2b2 Software components are distributed as open source
Implementations

CTSA’s

- Boston University
- Case Western Reserve University *(including Cleveland Clinic)*
- Children's National Medical Center (GWU), Washington D.C.
- Duke University
- Emory University *(including Morehouse School of Medicine and Georgia Tech)*
- Harvard University *(including Beth Israel Deaconess Medical Center, Brigham and Women’s Hospital, Children's Hospital Boston, Dana Farber Cancer Center, Joslin Diabetes Center, Massachusetts General Hospital)*
- Medical University of South Carolina
- Medical College of Wisconsin
- Oregon Health & Science University
- Penn State Milton S. Hershey Medical Center
- Tufts University
- University of Alabama at Birmingham
- University of Arkansas for Medical Sciences
- University of California Davis
- University of California, Irvine
- University of California, Los Angeles*
- University of California, San Diego*
- University of California San Francisco
- University of Chicago
- University of Cincinnati *(including Cincinnati Children's Hospital Medical Center)*
- University of Colorado Denver *(including Children’s Hospital Colorado)*
- University of Florida
- University of Kansas Medical Center
- University of Kentucky Research Foundation
- University of Massachusetts Medical School, Worcester
- University of Michigan
- University of Pennsylvania *(including Children’s Hospital of Philadelphia)*
- University of Pittsburgh *(including their Cancer Institute)*
- University of Rochester School of Medicine and Dentistry
- University of Texas Health Sciences Center at Houston
- University of Texas Health Sciences Center at San Antonio
- University of Texas Medical Branch (Galveston)
- University of Texas Southwestern Medical Center at Dallas
- University of Utah
- University of Washington
- University of Wisconsin - Madison *(including Marshfield Clinic)*
- Virginia Commonwealth University
- Weill Cornell Medical College

Academic Health Centers (does not include AHCs that are part of a CTSA):

- Arizona State University
- City of Hope, Los Angeles
- Georgia Health Sciences University, Augusta
- Hartford Hospital, CN
- HealthShare Montana
- Massachusetts Veterans Epidemiology Research and Information Center (MAVERICK), Boston
- Nemours
- Phoenix Children's Hospital
- Regenstrief Institute
- Thomas Jefferson University
- University of Connecticut Health Center
- University of Missouri School of Medicine
- University of Tennessee Health Sciences Center
- Wake Forest University Baptist Medical Center

HMOs:

- Group Health Cooperative
- Kaiser Permanente

International:

- Georges Pompidou Hospital, Paris, France
- Hospital of the Free University of Brussels, Belgium
- Inserm U936, Rennes, France
- Institute for Data Technology and Informatics (IDI), NTNU, Norway
- Institute for Molecular Medicine Finland (FIMM)
- Karolinska Institute, Sweden
- Landspitali University Hospital, Reykjavik, Iceland
- Tokyo Medical and Dental University, Japan
- University of Bordeaux Segalen, France
- University of Erlangen-Nuremberg, Germany
- University of Goettingen, Goettingen, Germany
- University of Pavia, Pavia, Italy
- University of Seoul, Seoul, Korea

Companies:

- Johnson and Johnson (TransMART)
- GE Healthcare Clinical Data Services
Interrogation can occur through i2b2 web client.
I2b2 Workbench provides a detailed patient view for Investigator
Use Phenotyping Algorithms to define cohorts of treatment-resistant and treatment-responsive depression.

Need to Determine: Depressed or Well at Encounter

Must Improve Accuracy of Diagnoses from Electronic Health Record

<table>
<thead>
<tr>
<th>Clinical Status</th>
<th>Model</th>
<th>Specificity</th>
<th>Sensitivity</th>
<th>Precision</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depressed</td>
<td>Billing Codes</td>
<td>0.95</td>
<td>0.09 (0.03)</td>
<td>0.57 (0.14)</td>
<td>0.54 (0.02)</td>
</tr>
<tr>
<td>Depressed</td>
<td>NLP</td>
<td>0.95</td>
<td>0.42 (0.05)</td>
<td>0.78 (0.02)</td>
<td>0.88 (0.02)</td>
</tr>
<tr>
<td>Depressed</td>
<td>NLP + Billing Codes</td>
<td>0.95</td>
<td>0.39 (0.06)</td>
<td>0.78 (0.02)</td>
<td>0.87 (0.02)</td>
</tr>
<tr>
<td>Well</td>
<td>Billing Codes</td>
<td>0.95</td>
<td>0.06 (0.02)</td>
<td>0.26 (0.27)</td>
<td>0.55 (0.03)</td>
</tr>
<tr>
<td>Well</td>
<td>NLP</td>
<td>0.95</td>
<td>0.37 (0.06)</td>
<td>0.86 (0.02)</td>
<td>0.85 (0.02)</td>
</tr>
<tr>
<td>Well</td>
<td>NLP + Billing Codes</td>
<td>0.95</td>
<td>0.39 (0.07)</td>
<td>0.85 (0.02)</td>
<td>0.86 (0.02)</td>
</tr>
</tbody>
</table>

Initially: AUC = 0.54
Finally: AUC = 0.87
Data Integration | Phenotype Discovery Center

Electronic Medical Record (EMR) Data

- RPDR
  - Coded Data
    - Demographics
    - Diagnoses
    - Lab Results
  - Text Data (Notes/Reports)
    - Medications
    - Procedures
    - Visits
    - Physician Notes
    - Imaging Reports
    - Pathology Reports
    - Surgery Notes

Informatics Tools

- Calculated Controls (Charlson Index)
- Data Visualization
- Data Queries
- Annotation
- Extract Data
- Natural Language Processing

Additional Data

- Other Research Data
- Survey Data

Genetic Data

- GWAS

Biobank Data

- Samples
  - DNA
  - Serum
  - Plasma

- Consent
  - Recontact
  - Consent Status

Validated Phenotypes

- Type II Diabetes
- IBD
- Coronary Artery Disease
- Multiple Sclerosis
- Congestive Heart Failure
- Bipolar Disorder
- Rheumatoid Arthritis

Research
Curating a Disease Algorithm

1. Create a gold standard training set.

2. Create a comprehensive list of features from patient’s electronic data that describe the disease of interest.

3. Develop the classification algorithm. Using the data analysis file and the training set from step 1, assess the frequency of each variable. Remove variables with low prevalence. Apply adaptive LASSO penalized logistic regression to identify highly predictive variables for the algorithm.

4. Apply the algorithm to all subjects in the superset and assign each subject a probability of having the phenotype.
White matter abnormalities associated with treatment-resistant depression

- Scans collected as part of routine clinical care
- Diffusion tensor imaging in 150 pts
- Age-related decline in white matter integrity increases with treatment resistant depression

Medial fornix shows strongest effect

## Biobank Portal | Curated Diseases

<table>
<thead>
<tr>
<th>Validated Phenotype</th>
<th>Count*</th>
<th>Predictive Positive Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bipolar Disease</td>
<td>71</td>
<td>89%</td>
</tr>
<tr>
<td>Congestive Heart Failure</td>
<td>387</td>
<td>90%</td>
</tr>
<tr>
<td>Coronary Artery Disease</td>
<td>2,420</td>
<td>97%</td>
</tr>
<tr>
<td>Crohn’s Disease</td>
<td>453</td>
<td>90%</td>
</tr>
<tr>
<td>Multiple Sclerosis</td>
<td>94</td>
<td>90%</td>
</tr>
<tr>
<td>Rheumatoid Arthritis</td>
<td>550</td>
<td>90%</td>
</tr>
<tr>
<td>Type 2 Diabetes Mellitus</td>
<td>1,887</td>
<td>97%</td>
</tr>
<tr>
<td>Ulcerative Colitis</td>
<td>330</td>
<td>90%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Healthy Controls based on Charlson Index</th>
<th>Count**</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 10-year survival probability is &gt;98.3%</td>
<td>2,206</td>
</tr>
<tr>
<td>1 – 10-year survival probability is &gt;95.87%</td>
<td>4,343</td>
</tr>
<tr>
<td>2 – 10-year survival probability is &gt;90.15%</td>
<td>6,545</td>
</tr>
</tbody>
</table>

* Based on 15,880 patients
** Based on 21,300 patients
High Quality Phenotypes available for Genetic Studies
Genotype Data

- 3349 SNP or indels
- 1680 Homozygous
- 1336 subjects with protein altering (frameshift, missense, nonsense, start loss, stop loss) variant

TTN
Partners Biobank Portal – Request Genetic Data

Biobank Portal Genomics Request

The Biobank contains subjects who have consented to make their genomic data available for research. To request genomic data, please fill out the form below.

Once you have submitted this form, you will be contacted by the Biobank team to complete your request. Please note that no genomic data will be sent out without a discussion with the Biobank team.

For more assistance on making a genomic data request, please refer to the Biobank Wiki section "How to Make Genomic Requests Using the Biobank Portal" or contact us.

Step 1. Patient Selection:
Choose Your Patients: [Enter a list of Biobank Subject IDs] OR [Select a Previous Query]

Your Patient List:
Type or paste a list of Biobank Subject IDs, one per line. You can find Subject IDs by downloading the de-identified data set. Do not enter MRNs.

10021040
10023384
10018524
10018243
10018529
10020593

Step 2. Details of Request:
Request Type: [De-Identified] OR [Identified]

Contact Name: Wattanasin, Nich

Contact Email: nwattanasin@partners.org
High Throughput Methods for supporting Translational Research

- Set of patients is selected from medical record data in a high throughput fashion

- Investigators explore phenotypes of these patients using i2b2 tools and a translational team developed to work specifically with medical record data

- Distributed networks cross institutional boundaries for phenotype selection, public health, and hypothesis testing

- Personalized medicine is delivered into clinical care
emerge network
ELECTRONIC MEDICAL RECORDS AND GENOMICS

25,000 Network-wide

Genotyping

2500

Sequencing 100 high-priority genes

Discovery

Penetrance and Pleiotropy

PARTNERS BIOBANK

Outcomes:
- MD visits
- Labs/PLDL levels
- Health care costs
- Prescriptions
- Family screening

Receive LDLR LoF Results

Randomize

Weeks

52

Treatment as usual
Federated Queries in PCORNet

- Partners HealthCare System
- Boston Children’s Hospital
- BIDMC
- Boston Health Net (BMC and Community Health Centers)
- Columbia U. Medical Center and New York Presbyterian Hospital
- Wake Forest Baptist Medical Center
- Morehouse/Grady/RCMI
- University of California, Davis
- Washington University in St. Louis
- U Texas Health Science Center/Houston
Distributed Query System

1. Certify investigator
2. Compose Query
3. Broadcast Query
4. Verify user, adapt for local CRC
5. I2B2 CRC Query (no changes)
6. Send results to aggregator
7. Aggregate Results

CHB PM
CHB Composer
Partners Responder Adapter
CHB Responder Adapter
BIDMC Responder Adapter
CRC
CRC
CRC

CHB Broadcaster Aggregator

Partners Healthcare

Logo

Logo

Logo
Run Query Using SCILHS-SHRINE
Workflow at the sites to find patients for a clinical trial:

- After a query is run across the “SHRINE” network, the query is automatically saved at every site.

- The query saved at each site is transformed into a patient set.

- The patient set is studied and sorted for the specific patients eligible for the Clinical Trial.
**Review Patients at Sites**

![i2b2 Workbench for i2b2 Demo (Oracle)](image)

### Table: Review Patients at Sites

<table>
<thead>
<tr>
<th>Set #</th>
<th>Patient Set Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Patient Set for 'Asthma-Albuter@01:06:16'</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SMART</th>
<th>Patient ID</th>
<th>PSet #</th>
<th>Patient Name</th>
<th>Gender</th>
<th>Race</th>
<th>Date of Birth</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1000000052</td>
<td>1-1</td>
<td>xxxxx, xxxxx</td>
<td>F</td>
<td>white</td>
<td>1966-08-29T...</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>1000000053</td>
<td>1-1</td>
<td>xxxxx, xxxxx</td>
<td>M</td>
<td>asian</td>
<td>1969-03-07T...</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>1000000054</td>
<td>1-1</td>
<td>xxxxx, xxxxx</td>
<td>M</td>
<td>black</td>
<td>1976-08-13T...</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>1000000055</td>
<td>1-1</td>
<td>xxxxx, xxxxx</td>
<td>F</td>
<td>hispanic</td>
<td>1973-06-25T...</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>1000000056</td>
<td>1-1</td>
<td>xxxxx, xxxxx</td>
<td>F</td>
<td>black</td>
<td>1981-08-05T...</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>1000000057</td>
<td>1-1</td>
<td>xxxxx, xxxxx</td>
<td>M</td>
<td>asian</td>
<td>1981-07-10T...</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>1000000058</td>
<td>1-1</td>
<td>xxxxx, xxxxx</td>
<td>M</td>
<td>hispanic</td>
<td>1986-12-16T...</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>1000000059</td>
<td>1-1</td>
<td>xxxxx, xxxxx</td>
<td>F</td>
<td>hispanic</td>
<td>1989-01-11T...</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>1000000060</td>
<td>1-1</td>
<td>xxxxx, xxxxx</td>
<td>F</td>
<td>hispanic</td>
<td>1976-11-23T...</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>1000000061</td>
<td>1-1</td>
<td>xxxxx, xxxxx</td>
<td>M</td>
<td>black</td>
<td>1991-03-29T...</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>1000000062</td>
<td>1-1</td>
<td>xxxxx, xxxxx</td>
<td>F</td>
<td>black</td>
<td>2049-10-19T...</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>1000000063</td>
<td>1-1</td>
<td>xxxxx, xxxxx</td>
<td>F</td>
<td>white</td>
<td>2027-02-27T...</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>1000000064</td>
<td>1-1</td>
<td>xxxxx, xxxxx</td>
<td>M</td>
<td>white</td>
<td>1956-05-22T...</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>1000000065</td>
<td>1-1</td>
<td>xxxxx, xxxxx</td>
<td>M</td>
<td>hispanic</td>
<td>1978-09-11T...</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>1000000066</td>
<td>1-1</td>
<td>xxxxx, xxxxx</td>
<td>F</td>
<td>Indian</td>
<td>1983-11-19T...</td>
<td>23</td>
</tr>
</tbody>
</table>
High Throughput Methods for supporting Translational Research

- Set of patients is selected from medical record data in a high throughput fashion

- Investigators explore phenotypes of these patients using i2b2 tools and a translational team developed to work specifically with medical record data

- Distributed networks cross institutional boundaries for phenotype selection, public health, and hypothesis testing

- Personalized medicine is delivered into clinical care
Personalized Medicine for the Everyday Clinician - Finding similar patients

- Finding similar patients help us understand what is disease and what is normal, to distinguish between several disease states, help predict successful therapies, and to help determine next steps in potentially very expensive diagnostic pathways.

- This is an opportunity for combining the EHR, Big Data Queries, and SMART Apps.

- Apply the approaches we have used to conduct scientific research to Provider and Patient engaging visualizations.
Designing the App Store for Health

No Small Change for the Health Information Economy
Kenneth D. Mandl, M.D., M.P.H., and Isaac S. Kohane, M.D., Ph.D.

The economic stimulus package signed by President Barack Obama on February 17 included a $19 billion investment in health information technology. How can we best take advantage of this unprecedented opportunity to transform health care and stimulate the health information economy while also stimulating the U.S. economy? A health care system adapting to the effects of an aging population, growing expenditures, and a diminishing primary care workforce needs the support of a flexible information infrastructure that facilitates innovation in wellness, health care, and public health.

Flexibility is critical, since the system will have to function under new policies and in the service of new health care delivery mechanisms, and it will need to incorporate emerging information technologies on an ongoing basis. As we seek to design a system that will constantly evolve and encourage innovation, we can glean lessons from large-scale information-technology successes in other fields. An essential first lesson is that ideally, system components should be not only interoperable but also substitutable.

The Apple iPhone, for example, uses a software platform with a published interface that allows software developers outside Apple to create applications. There are now nearly 10,000 applications that consumers can download and use with the common phone interface. The platform separates the system from the functional-
State-of-the-Art ???

Cardio CRP

For Ages > 17 Years:

<table>
<thead>
<tr>
<th>CCRP mg/L</th>
<th>Risk According to AHA/CDC Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1.0</td>
<td>Lower Relative Cardiovascular Risk</td>
</tr>
<tr>
<td>1.0-3.0</td>
<td>Average Relative Cardiovascular Risk</td>
</tr>
<tr>
<td>3.1-10.0</td>
<td>Higher Relative Cardiovascular Risk</td>
</tr>
</tbody>
</table>

Consider retesting in 1 to 2 weeks to exclude a benign transient elevation in the baseline CRP value secondary to infection or inflammation.

Persistent elevations upon retesting, may be associated with infection and inflammation.
Bloodwork Cardiology Result

Patient info
NAME: John Doe
GENDER: M AGE: 49 DOB: 01/10/1961

About this test
This report evaluates your potential risk of heart disease, heart attack, and stroke.

Your results
CRP level test
3.3 your level of a specific protein in the blood linked to inflammation of blood vessels

Total cholesterol level
265

LDL “bad” cholesterol
232

HDL “good” cholesterol
32

Your risk
You show an elevated risk of cardiovascular disease

Your risk would be lowered to
12% if your blood pressure were 120mm/Hg
10% if you quit smoking
6% if you reduced cholesterol to 160mg/DL

Use your CRP results and cholesterol level to calculate your 10 year risk of a cardiovascular event at ReynoldsRisk.org

What now?
Diet & exercise can improve your cholesterol levels
Quitting smoking can decrease your heart disease risk by 50% or more
Ask your doctor about statins or other medications that can lower cholesterol
Consider retesting in 1 to 2 weeks to exclude a temporary spike in blood levels

An Inspired Design from Dave McCandless (cc license)
Out of the Box - SMART Apps can link Big Data to the EMR

- Substitutable Medical Application and Reusable Technology – Started with grant form the Office of the National Coordinator
- Paradigm is similar to Mobile Apps with a proposed standard interface using FHIR (Fast Healthcare Interoperable Resource)
1 SMART App in 3 SMART Systems
What Big Data can do for the Everyday Clinician
Finding Similar Patients

• Looking at similar patients can help predict:
  • Future outcomes and responses to therapy
  • Course of disease
  • Penetrance of genetic variants
  • Likelihood that a diagnostic pathway might be fruitful

• Big Data Commons is an opportunity for combining data from the Electronic Health Record, Specialized Health Databases, Analytics from Big Data Queries, and presentation in SMART Apps

• Presentation of results can be greatly enhanced with engaging visualizations for the provider making difficult, complex decisions
https://gallery.smarthealthit.org/boston-childrens-hospital/growth-chart
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>3y 6d</td>
<td>4y 6d</td>
<td>5y 5d</td>
<td>6y 6m</td>
<td>7y 5d</td>
<td>7y 7m</td>
<td>8y 5d</td>
<td>8y 7m</td>
<td>9y 1m</td>
<td>9y 6m</td>
</tr>
<tr>
<td>Annotation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>123.8cm</td>
</tr>
<tr>
<td>Percentile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Velocity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weight</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>24.4kg</td>
</tr>
<tr>
<td>Percentile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Velocity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Head C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Velocity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Paul Lutrell has a healthy weight of 24.4 kg (53 lb 13 oz).
Find Normal MRI’s at All Ages 0-6 y/o

Number of patients who had a brain MRI scan at a particular age in months from 0 to 6 years (A) and in weeks from 0 to 4 months (B)
Determining a Normal Child’s MRI
Atlases provide a visual guide for Radiology Decision Support, such as determining Perinatal Hypoxic Ischemic Encephalopathy

ADC map from 4 infants: Each statistically compared to age matched atlas yields visual guide to pathology

Quantitative analysis tools + large data sets = Great insights for practicing doctors
Tribute to…

- **I2b2 Core Team**
  - Isaac Kohane
  - Susanne Churchill
  - Michael Mendis
  -Christopher Herrick
  - Griffin Weber
  - Paul Avillach
  - Lori Phillips
  - Janice Donahoe
  - Nich Wattanasin
  - David Wang
  - Vivian Gainer
  - Victor Castro
  - Andrew Cagan
  - Wayne Chan

- **SMART Apps Team**
  - Calum MacRae
  - Sandy Aronson
  - Mike Oats
  - Layne Ainsworth
  - Kenneth Mandl
  - Joshua Mandel

- **Radiology (mi2b2) Team**
  - Randy Gollub
  - P Ellen Grant
  - Kathy Andriole
  - Kallirroi Retzepi
  - Rudolph Pienaar
  - Lilla Zollei
  - Yangming Ou
I2b2, SHRINE, and SMART Information and Software on the Web

i2b2 Homepage (https://www.i2b2.org)
i2b2 Software (https://www.i2b2.org/software)
i2b2 Community Site (https://community.i2b2.org)
SMART Platforms Homepage (http://smarthealthit.org)

Partners Healthcare, NIH/NCBC/BD2K; NIMH; NCATS; NIBIB; NHGRI

NIH R01 EB014947
NIH U54 LM008748
NIH U01 HG008685
PCORI 282364.5077585.0007

NIH U54 HG007963
NIH R01 AT006364
NIH R01 AT005280
NIH P01 AT006663
THANK YOU