How Australia can lead the way to precision medicine

Isaac S. Kohane, MD, PhD
Toward Precision Medicine: Building a Knowledge Network for Biomedical Research and a New Taxonomy of Disease
Report from National academy of science, USA, 2011
Afferent and Efferent Arms of Healthcare Information
Cusum of Hospitalizations due to Myocardial Infarction

Rofecoxib introduced

Rofecoxib withdrawn

Reis, et al., BMJ 2009

Jane Q. Doe
ED Profile
June, 1992 – Feb, 1999

JURY RISK: HIGH
### Survival 3 Years After a WBC Test

(White, Male, 50-69 Years; Using Last WBC Between 7/28/05 and 7/27/06)

<table>
<thead>
<tr>
<th>Repeat Interval</th>
<th>Result Time</th>
<th>WBC Value</th>
<th>Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Low</td>
<td>Normal</td>
</tr>
<tr>
<td><strong>&lt; 1 Day</strong></td>
<td>12a-8a</td>
<td>43.33%</td>
<td>84.68%</td>
</tr>
<tr>
<td></td>
<td>8a-4p</td>
<td>54.55%</td>
<td>86.61%</td>
</tr>
<tr>
<td></td>
<td>4p-12a</td>
<td></td>
<td>77.30%</td>
</tr>
<tr>
<td><strong>&lt; 1 Year</strong></td>
<td>12a-8a</td>
<td>47.83%</td>
<td>79.58%</td>
</tr>
<tr>
<td></td>
<td>8a-4p</td>
<td>76.96%</td>
<td>90.73%</td>
</tr>
<tr>
<td></td>
<td>4p-12a</td>
<td>81.65%</td>
<td>92.99%</td>
</tr>
<tr>
<td><strong>&gt; 1 Year</strong></td>
<td>12a-8a</td>
<td></td>
<td>95.65%</td>
</tr>
<tr>
<td></td>
<td>8a-4p</td>
<td>97.30%</td>
<td>98.13%</td>
</tr>
<tr>
<td></td>
<td>4p-12a</td>
<td>92.68%</td>
<td>97.35%</td>
</tr>
<tr>
<td><strong>Any</strong></td>
<td></td>
<td>73.17%</td>
<td>91.79%</td>
</tr>
<tr>
<td><strong>Patients</strong></td>
<td>1286</td>
<td>18775</td>
<td>3052</td>
</tr>
</tbody>
</table>
Figure. The Tapestry of Potentially High-Value Information Sources That May Be Linked to an Individual for Use in Health Care

Types of Data

- **Structured Data**
  - Electronic pill dispensers
  - Medication filled
  - Dose
  - Route
  - NDC
  - RxNorm
  - HL7
- **Unstructured Data**
  - Medication instructions
  - Allergies
  - Out-of-pocket expenses
  - Medication taken
  - Diaries
  - Herbal remedies
  - Alternative therapies

Types of Data Continued:

- **Demographics**
- **Encounters**
- **Diagnoses**
- **Procedures**
- **Diagnostics (ordered)**
  - Personal health records
  - Home treatments, monitors, tests
- **Diagnostics (results)**
  - PatientsLikeMe.com
  - 23andMe.com
  - Lab values, vital signs
  - SNPs, arrays
- **Genetics**
  - Ancestry.com
- **Social history**
  - Police records
  - Tobacco/alcohol use
- **Family history**
- **Symptoms**
  - Indirect from OTC purchases
- **Lifestyle**
  - Fitness club memberships, grocery store purchases
- **Socioeconomic**
  - Census records, Zillow, LinkedIn
- **Social network**
  - Facebook friends, Twitter hashtags
- **Environment**
  - Climate, weather, public health databases, HealthMap.org, GIS maps, EPA, phone GPS

Probabilistic Linkage to Validate Existing Data or Fill in Missing Data

Examples of Biomedical Data:

- **Pharmacy data**
- **Health care center (electronic health record) data**
- **Claims data**
- **Registry or clinical trial data**
- **Data outside of health care system**

Ability to Link Data to an Individual:

- Easier to link to individuals
- Harder to link to individuals
- Only aggregate data exists

Data Quantity:

- More
- Less

Observation Identifiers Names and Codes:

- NDC, National Drug Code
- OTC, over-the-counter
- SNOMED, Systematized Nomenclature of Medicine
- SNP, single-nucleotide polymorphism

CPT indicates current procedural terminology; ECG, electrocardiography; EPA, US Environmental Protection Agency; GIS, geographic information systems; GPS, global positioning system; HL7, Health Level 7 coding standard; ICD-9, Institutional Classification of Diseases, Ninth Revision; LOINC, Logical
An Aspiration

**Ranking**

- Ranking = **Scoring** + **Sorting** + **Filtering** bags of movies for presentation to a user
- Key algorithm, sorts titles in most contexts
- Goal: Find the best possible ordering of a set of *videos* for a *user* within a specific *context* in real-time
- Objective: maximize consumption & "enjoyment"

**Factors**

- Accuracy
- Novelty
- Diversity
- Freshness
- Scalability
A question that has withstood the test of time:

If a test to detect a disease whose prevalence is 1/1000 has a false positive rate of 5%, what is the chance that a person found to have a positive result actually has the disease, assuming you know nothing about the person's symptoms or signs?
RESEARCH LETTER

Medicine’s Uncomfortable Relationship With Math: Calculating Positive Predictive Value

JAMA Internal Medicine  Published online April 21, 2014

Survey Responses

$n = 61$

Most common answer – 95%

$n = 27$

Correct answer – 2%

$n = 14$

Count

Response

Students  $n = 10$
House Staff  $n = 26$
Attendings  $n = 25$
Finding rare events of interest.

Cardiac angiogenic imbalance leads to peripartum cardiomyopathy

samples from subjects with PPCM have been previously described\textsuperscript{a}. Patients in both studies were predominantly Caucasian. Retrospective analyses of PPCM and pre-eclampsia in the Harvard teaching hospitals were performed using the Harvard Shared Health Research Information Network (SHRINE)\textsuperscript{b}, a de-identified repository of aggregate patient information.
Skin/fibroblast -> NPC
Clinical phenotyping
WGS
Coding/NC RNA
RNA + perturbations
RBBS + selected H marks
Differentiated neuron ± depol.
CEGS: NIMH/NHGRI
RNA + selected H marks
Probabilistic Modeling

Regulatory Dependencies

Classification & Prediction
Patterns across tens of thousands of patients...
Autism or Autisms?
Opportunities for Australia

• Create a national educational program for clinicians that bakes quantitative reasoning into ALL medical decision making
• Create a genuine nation-wide health monitoring network.
• Define ‘diseasome’ trajectories.
• Bank samples for near-term cost-effective molecular & genomic characterization
Afferent and Efferent Arms of Healthcare Information
“A woman could be delivered on Monday, happy and well with her newborn baby on Tuesday, feverish and ill by Wednesday evening, delirious and in agony with peritonitis on Thursday, and dead on Friday or Saturday.”

Cancer’s Super-Survivors
How the Promise of Immunotherapy Is Transforming Oncology

BY RON WINSLOW
Adam’s Story
Titin
Breakthrough
A Tale of Two Children

• Los Angeles

• London

“Concordia”
IBD Expression Profiles: Whole Blood

Can we identify an existing drug that will move these patients towards the healthy region?
Predicted PCA of IBD Patients' Gene Expression Profiles Treated with Azathioprine

Tissue: Blood
Drug: Azathioprine
IBD Expression Profiles: Colon
Patient population likely to benefit from Prednisone

Tissue: Colon
Drug: Prednisone
Applications

• Drug Repositioning
• Drug Off-Target Effect Screening
• Directed Screen Development
• Identify classes of action beyond ostensible target
With a little help from PMI
We can all be as effective as Matt
Informed Cohort

Kohane, Science, 2007
UDN Data Process Overview

Data Mgt. Step #1:
Patient Application & initial response

UDN Gateway

Public summaries

Metrics outcomes

UDN Record Archive and Electronic Trial Mgt.

Data Mgt. Step #2:
referral and data transfer to a UDN clinical center

Data Mgt. Step #3:
inpatient evaluation

Data Mgt. Step #4:
post visit reporting and review

Digital files, workflow, admin reports, phenotype, genotype

UDN Data Warehouse
Small Change for the Health Information Economy

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 computerize 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 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 functiona
Designing the App Store for Health
‘API’

Can EMRs behave like iPhones or Androids in that innovators readily create and widely distribute substitutable apps across thousands of installs?

$15M
Apps Gallery

Launched April 2015

https://gallery.smarthealthit.org/
Inspired by a

WIREDC

18.12 Issue

Design Challenge
State-of-the-Art ???

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. Consider retesting in 1 to 2 weeks to exclude a benign transient elevation in the baseline CRP value secondary to infection or inflammation.</td>
</tr>
<tr>
<td>&gt;10.0</td>
<td>Persistent elevations upon retesting, may be associated with infection and inflammation.</td>
</tr>
</tbody>
</table>
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.

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

Low risk          Average          High risk of cardiovascular disease
0 - 0.5 mg/L     0.5 - 2.0 mg/L     2.0 - 10 mg/L

Total cholesterol level

Desirable: 0
Borderline: 200 - 239
High: 240 - 249+

LDL “bad” cholesterol

Optimal: 60 - 159
Near optimal: 160 - 189
Borderline: 190 - 239
High: 240 - 190+

HDL “good” cholesterol

Low: 35 - 49
Normal: 50 - 64
High: 65 - 70+

3 Your risk
You show an elevated risk of cardiovascular disease

If you’re a smoker with normal blood pressure (120 mm/Hg) but family history of heart attack before age 60 (one or both parents) your risk over 10 years is: 15%

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 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
1 Design + 1 Developer + 1 Week
1 SMART App in 3 SMART Systems
### Meducation

**Coumadin Tablet 5 mg**

**How to take medicine**

Take the medicine by mouth once a day.

<table>
<thead>
<tr>
<th>Time</th>
<th>Pill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breakfast</td>
<td>1</td>
</tr>
<tr>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>Dinner</td>
<td></td>
</tr>
<tr>
<td>Before Bedtime</td>
<td></td>
</tr>
</tbody>
</table>

**Instructions**

Swallow the medicine without crushing or chewing it. This medicine may be taken with or without food. It is very important that you take the medicine at about the same time every day. It will work best if you do this.

---

### Indivohealth

**Coumadin Tablet 5 mg**

**Cómo tomar el medicamento**

Tomar el medicamento por la boca una vez al día.

<table>
<thead>
<tr>
<th>Time</th>
<th>Pill</th>
</tr>
</thead>
<tbody>
<tr>
<td>Desayuno</td>
<td>1</td>
</tr>
<tr>
<td>Almuerzo</td>
<td></td>
</tr>
<tr>
<td>Comida</td>
<td></td>
</tr>
<tr>
<td>Al acostarse</td>
<td></td>
</tr>
</tbody>
</table>

**Instrucciones**

Trague el medicamento sin aplastarlo o masticar. Este medicamento se puede tomar con o sin alimentos.

---

**Harvard Medical School**

Department of Biomedical Informatics
Precision Medicine App

Patient Search

(MRN: )

Male,
Diagnosis: Lung Cancer
Mutation: KRAS G12C

Mutated Genes Observed in Lung Cancer

Observed Variants in Lung Cancer Patients with KRAS Mutation
Some SMART Hotbeds

- Intermountain Healthcare
- Premier Health
- HCA
- SMART
- HP
- Duke Medicine
- Geisinger Health Plan
- OpenMRS
- Epic
- surescripts
- VECNAxG Health Solutions
- Cerner
- Lilly
- "polyglot"
- HL7 FHIR
- Department of Biomedical Informatics
A few examples of Uptake

• Argonaut
  – HL7-led implementation of SMART in 5 major EHRs
• XG Systems
  – Geisenger/Glenn Steele spinoff to implement Geisenger knowhow in SMART apps
• Intermountain and HSP
  – Implementing SMART apps
• Duke
  – Implementing SMART apps through exposed Epic services
• HCA
  – Largest health system in America implementing SMART Apps for clinical trials matching
Opportunities for Australia

• Do not follow the US into documentation doldrums
  – Automation that augments clinicians
  – Inspire young clinicians by allowing direct innovation
• Experiment with new diagnostic models
• Take risks where there is little hope.
• Accelerate empowered patients.
Australia, please lead the way!

We Only Have Our Diseases To Lose

Thank you.