Data Driven Medicine: Data, Not Programs
Sharing, Not Hoarding Individuals, Not Populations

CAS Interdisciplinary Innovation Forums
BIG DATA FORUM
June 5, 2012

Joanne S. Luciano, PhD
Research Associate Professor
Tetherless World Constellation
Rensselaer Polytechnic Institute
Times have changed

Keep the information about the data with the data, then it can be (re)used

- The end of the blockbuster era
- The right treatment to the right patient at the right time
- Lowering cost of pharmaceutical development
- Improved patient response to treatment
- Aging population
- Technology (Web and Mobile)
- Shifts in thinking

Photo: http://www.flickr.com/photos/sepblog/4014143391/
Data Driven Medicine:

Data, Not Programs
Sharing, Not Hoarding
Individuals, Not Populations
Data, Not Programs

Feet?
Years?
December?
Noon?
Dozen?

Data, Not Programs

Data Dictionaries:
Without a data dictionary, a database management system [or any program] cannot access data from the database.”

Duh!

Shifting the Burden from the User to the Provider

Balancing resources for developing Online access

- Web Services
- Dynamic pages on-the-fly from databases
- Static HTML pages

Increasing level of resources and skills to create

Increasing complexity of data structures

Increasing level of resources required to maintain
Data-Information-Knowledge Ecosystem

Producers

Data
Creation
Gathering

Information
Presentation
Organization

Context

Consumers

Knowledge
Integration
Conversation

Experience
Data Driven Medicine:

Data, Not Programs

Sharing, Not Hoarding

Individuals, Not Populations
Sharing, Not Hoarding

Now get it out there, where it can do some good.

Yes, you can share data and you can know you are referring to the same thing if it has the same URI.

Photo: http://www.flickr.com/photos/sepblog/4014143391/
The Semantic Web is the New Global Web of Data

It is about standards for publishing, sharing and querying knowledge drawn from diverse sources.

It makes possible the answering of sophisticated questions using background knowledge.
Establishing Communities of Interest/Practice

- BioPathways Consortium
- BioPAX
- W3C Semantic Web for Health Care and Life Sciences (HCLSIG)
Biological Pathways of the Cell

BioPAX

What’s a pathway?
Depends on who you ask!

Metabolic Pathways
Molecular Interaction Networks
Signaling Pathways
Gene Regulation

BioPAX Level 1
BioPAX Level 2
BioPAX Level 3
BioPAX Level 4
BioPAX

Biological PAthway eXchange

An abstract data model for biological pathway integration

Initiative arose from the community
BioPAX Biochemical Reaction

OWL (schema)

Instances (Individuals) (data)

phosphoglucose isomerase 5.3.1.9

β-D-glucose-6-phosphate → D-fructose-6-phosphate
BioPAX Ontology

parts

how the parts are known to interact

a set of interactions

Level 1 v1.0 (July 7th, 2004)
Before BioPAX

With BioPAX

>150 DBs and tools

Common “computable semantic” enables scientific discovery
Data Driven Medicine:

Data, Not Programs
Sharing, Not Hoarding
Individuals, Not Populations
Order and Time
Symptoms Improve
is Different for CBT and DMI
Individuals, Not Populations

Social Web
Blogs
Patients Like Me
Cure Together

http://curetogether.com/depression/ig/treatment-effectiveness-vs-popularity

Photo: http://www.flickr.com/photos/sepblog/4014143391/
A 2D barcode that helps predict a patient's personal response to medical therapy

A Medicine Safety Code captures data about a patient’s 400 most important pharmacogenomic markers and allergies as a 2D barcode. It can help to predict how medications are metabolized and tolerated by the patient. Medical doctors can use this information to make the delivery of medications safer and more effective.

Quickly retrieve pharmacogenomic markers of patients when needed

No central storage of data is necessary, giving patients full control over their personal health information.

http://safety-code.org/

Photo: http://www.flickr.com/photos/sepblog/4014143391/
Personalized Medicine

Components

• Understand disease heterogeneity
• Comprehend disease progression
• Determine genetic and environmental contributors
• Create treatments against relevant targets
  – drugs against relevant targets (molecular structures)
  – Yoga against stress
  – Exercise against obesity
  – Elimination against food intolerance or allergy
• Develop markers to predict response
• Identify concrete endpoints to measure response
We need to shorten the time

tighten the loop between research and practice;

15 years + is too long,

way too long
“While it took 2,300 years after the first report of angina for the condition to be commonly taught in medical curricula, modern discoveries are being disseminated at an increasingly rapid pace. Focusing on the last 150 years, the trend still appears to be linear, approaching the axis around 2025.”

Now **You** Try It!

Challenge: What can you do (combine) with your data? What will you find?