Computer Science Challenges from Medicine

Peter Szolovits
MIT Computer Science and Artificial Intelligence Lab
Prof. of Electrical Engineering & Computer Science
Prof. of Health Sciences & Technology

CRA Snowbird
July 14, 2008

http://medg.csail.mit.edu/

US Health Care is Broken

- IOM: 48-98,000 “unnecessary” deaths/year
- 45M uninsured
  - Emergency Room as primary care
- Poor communication among providers
  - Repeat tests, incoherent care (no continuity), delays
- Spending ~17% of GDP, and growing
  - GM cars contain more health care than steel
  - BTW, education spending ~8-9% !!!
- Poor IT deployment and use
  - Most IT adoption for “low-hanging fruit”, e.g., billing
  - Low investment levels
  - Major systems tend to “melt down” (e.g., Kaiser, NHS)

NAS/NRC/CSTB Study

(in progress; comments mine, not committee's!)

- Challenges in CS ∩ IT ∩ healthcare
  - Site-visit based study, led by Bill Stead (Vanderbilt)
- Fragmented data from heterogeneous systems
- Documentation of what *has been* done, not mediation of what *is being* done
- UI's look like paper predecessors
- Very rare decision support/evidence based advice
- Unclear, *ad hoc*, complex processes
  - Not recorded, not analyzed
- Frequent interruptions
- Speed is paramount for users

Points of Leverage

- Policy
  - Insurance
  - Incentives
- Technology
  - Improved collection, handling & use of data
  - Communication and workflow
  - Decision Support
  - Privacy and Confidentiality
- Transformational Opportunities
  - Patient involvement & control
  - Research integrated with care
  - Healthcare as a *system*
    - (IOM+NAE Report *Building a Better Delivery System: A New Engineering/Health Care Partnership*, 2005)

Data: Examples of the Good

- MIMIC II: 30,000 ICU patients @ BIDMC
  - Signals (~4000), numerics, notes, labs, pharma, HIS
- Harvard Crimson
  - Save all blood samples, available for studies
- Gene Expression Omnibus (GEO)
  - All “raw” data from NIH-supported genomic experiments
  - Available for data re-use

Data: The Bad

- Poor interoperability
- How to fix?
  - Standards
    - HL7 CDA, CCR, ASN12, DICOM, LOINC, ICD, SNOMED…
    - Office of National Coordinator for Healthcare
      - AHIC, HITSP, CCHIT, HISPC, …
  - “Semantic Web”--loosely coupled declarative data

Data: The Opportunity

- Improved acquisition methods
  - *Intelligent Listening*--new modalities such as speech
  - *Aware examining room*--gestures, seeing & interpreting actions
  - *Walking ICU*--wearable real-time instrumentation
- Lifelong, patient-controlled records
  - E.g., indivohealth.org, MS HealthVault, Google Health

Decision Support

- Models of disease and of healthcare
  - "Expert systems"--rules or patterns
  - Statistical predictive models
  - Machine learning/data mining (*neo-statistics*)
  - Qualitative "causal" models
  - Differential equation models of pathophysiology
- Integration with workflow
  - E.g., CPOE
  - Built-in follow-up actions with each action
- Support patients, not just providers

Patient Control

- Who cares most about your health?
- Who is “on the spot” for all events & decisions?
- Who knows your preferences best?
- Who is willing to work without payment?

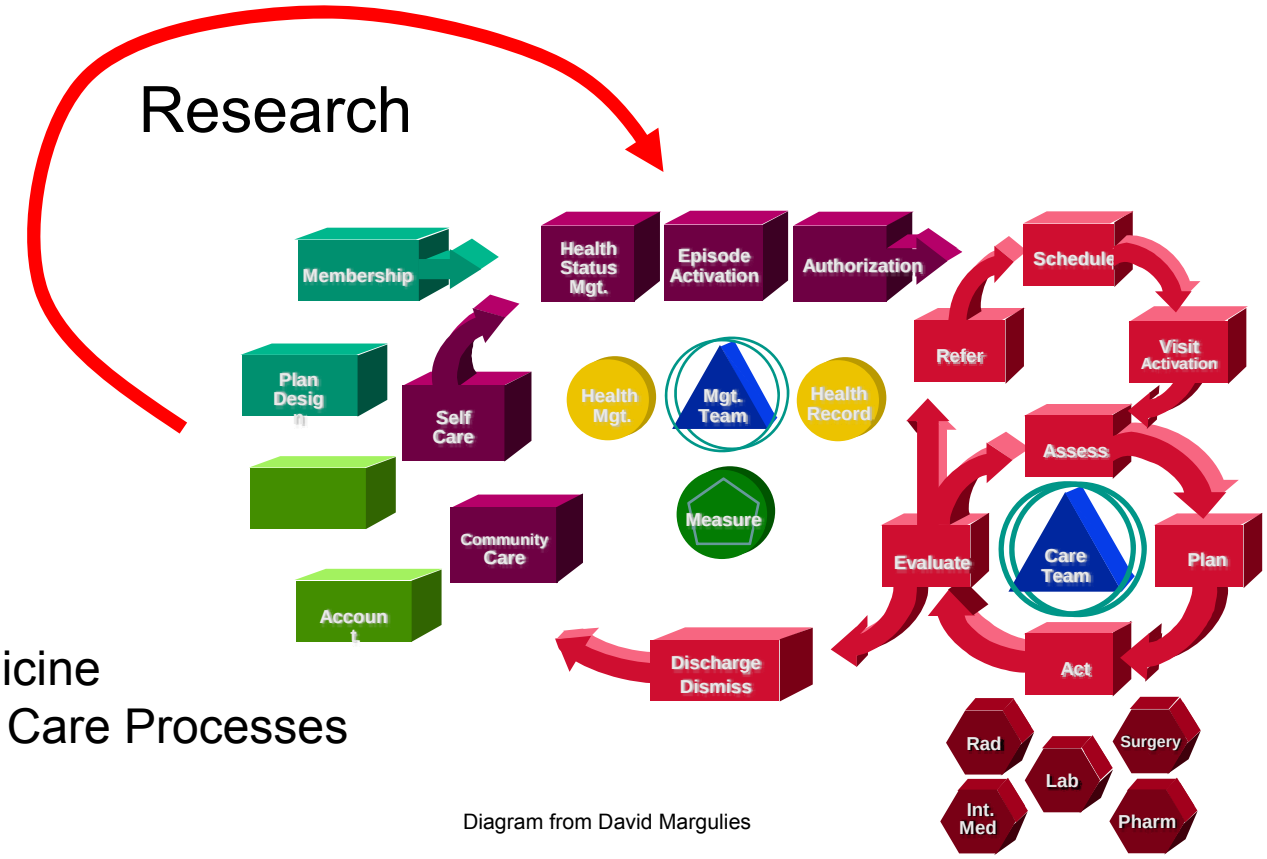
So, why not put ***you*** in charge of your continuity of care?

Desired Functionality

(from 1994 *Guardian Angel* proposal)

- **Patient-owned life-long individual record:** all medical conditions, care, preferences, …; allows individual to collect data on own medically-relevant experiences
- **Personal interface** to health-care information systems: hospital, lab, clinic, billing, …
- **Individualized medical encyclopedia**: explains results and plans to patient
- **Communication interface** with care team
- Permit unobtrusive **continuous monitoring** of relevant health-related activities and conditions
- **Decision support** for the patient and caretakers

Integrating Research with Care



Biomedicine
Clinical Care Processes

Diagram from David Margulies

I2b2: Integrating Information from Bench to Bedside

- Phenotype = Genotype + Environment
- We're getting very good at measuring G
- P is represented by clinical history
- E.g., Scott Weiss' asthma study
  - Use Partners Health Care RPDR (Research Patient Data Repository) to select especially poorly-responding asthma patients
  - Collect genomic data
  - Find predictive relationships

Privacy and Confidentiality

- Improving trust
  - Transparency
  - Patient control of access and dissemination
- Cryptographic framework using digital signatures
  - Allows separation of possession from authenticity
  - Practical problem: authenticating patients, providers
- Separating individuality from identity
- De-identification
  - Tabular data: *k*-anonymity, geographic fuzz
  - Text: NLP models for finding PHI