

A Rapid-Learning Health System

Using HIT for Medical Research

Lynn Etheredge

March 17, 2009

Rapid Breakthrough Science Using IT

- A new international model for science
 - Large computerized research databases
 - Data-sharing via WWW
 - Networks of researchers, collaborative science
- Key developments
 - WWW created at CERN (1992): high energy physics, astronomy, global warming, earthquakes, etc.
 - Human Genome Project – spectacularly successful US-NIH model of *in silico*, collaborative research (2000)
- Cyberinfrastructure = future basis of science, technology, economic growth (NSF)

Applying IT to Medical Research

- **HIT research leadership**
 - Electronic health record (EHR) databases: Kaiser-Permanente (8 m); VA (8m); Geisinger (3m)
 - Research networks: HMO Research Network (15 HMOs, 11 m enrollees), Cancer Research Network (12 HMOs, 10 m enrollees), Vaccine Safety Datalink
 - Major institutions (e.g. Mayo) & companies (e.g. IBM)
- **Policy interest in HIT**
 - HHS national HIT office, HIT/EHR strategy, data and standards development
 - *Health Affairs* (rapid-learning health system), IOM (learning healthcare system) (2007)
- **Payoffs of *in silico* research**
 - Vioxx side effects (Kaiser)
 - Genome-disease associations (2007)

Recent Developments

- **New networks & grids**
 - FDA Sentinel network (100 M patients)
 - Cardiovascular research network
 - NCI's CaBIG & BIG Health Consortium - a WWW for medical research
- **New databases & capabilities**
 - Kaiser-RWJF: Biobank (200,000→500,000 patients, w/ clinical, genomic & environmental data), world's largest biobank
 - HHS: Genome-Wide Association Study database (dbGaP, w/ Framingham as core (9,000 patients))
- **New predictive models**
 - Archimedes-RWJF: ARChES project for very large-scale, physiology-based simulation models for clinical decision-making

Recent Developments

- **New policy consensus: HIT for**
 - Efficiency
 - Quality problems
 - Effectiveness research
 - Fiscal problems
- **New financing (2009) !!**
 - \$19 billion for HIT - EHRs
 - \$10 billion for NIH
 - \$1.1 billion for comparative effectiveness research

Rapid Learning Potential

- From today's *data-poor environment* -- data scarcity & expense of acquiring is a critical limit -- to a future *data-rich environment*, w/potential for near real-time learning from experience of millions of patients
- Research centers, pharma, biomed companies
 - If 100 institutions contribute 100 patient records /year to a computerized databank = 10,000 records
 - Research centers' net gain: 9900 research records for 100 records = 99:1, for each institution
 - Reduces duplication of research; saves \$
- NIH
 - Beyond laboratories with “postage stamp” databases, non-standard & non-interoperable
 - Beyond temporary networks (200+/yr)
 - Petabyte databases & petaflop computing power ! Peta = 10^{15}

Rapid Learning Potential

- **Researcher's productivity/discoveries— our most important resource/outcome**
 - Research in days vs years; far more research results – per year & per researcher. “Research at the speed of thought” (or faster!)
- **Research creativity, inspiration and breakthroughs**
 - Allows inexpensive & extensive data exploration
 - Allows researchers to quickly generate and test new hypotheses, from molecules to populations
 - Allows researchers to use relevant data from many sources
 - Allows researchers to study groups excluded from clinical trials (patients with multiple conditions, children, women), patient subgroups, heterogeneity of response, genetic variations, low incidence and orphan diseases, special needs populations
 - Allows many more individuals, with many more ideas and perspectives, to engage in research

Rapid Learning Potential

- Potential of HIT for effectiveness research
 - “Is treatment A or treatment B better for a patient group C?” is an *applied research* question
 - With the appropriate data, comparative effectiveness questions can be answered quickly & efficiently
 - *In today’s data-poor environment, the appropriate data are usually scarce, expensive and difficult to obtain*
- Examples of new databases/patient registries & networks under discussion/development
 - Pediatric research network; patients with multiple chronic conditions (Medicare/Medicaid dual eligibles); new technologies; medical device safety; orthopedics; women (breast cancer); autism; mental health; survivors of pancreatic & lung cancer; sickle cell; alcohol and drug abuse; high risk pregnancies, Medicare cancer registry

Government's Role

- The new international model for breakthrough science requires government support
 - Cyberinfrastructure, e.g. BIG Health Consortium, NSF high-performance computing, interoperability, data standards
 - Research funding, e.g. \$1.1 B for effectiveness research
 - National data policy for collaborative research
 - Publicly funded data must be “open science”, e.g. NSF/CODATA initiative for global, open-access scientific data
 - National research databases and learning networks must be developed and supported
 - National data policies are a key missing element in US (and international) support for breakthrough science. Government policies and incentives are critical for the “science of the commons”

What's Possible: Example #1

- **A rapid-learning system for new medical technologies**
 - New technologies are a priority: least known about benefits and risks, driver of future costs & benefits, relatively easy to organize research
 - At point of FDA approval/Medicare's first payment, HHS (w/ private sector, physicians & patients) would define a national research plan, data to be collected, research registries to be created. Public and private sector collaborate on coverage with evidence development
 - After three years, the accumulated database would be analyzed for comparative effectiveness and safety

What's Possible: Example #2

- National research databases for government-funded clinical studies
 - HHS does not yet have a national policy requiring public availability (w/ standards, interoperability) of publicly funded data. *US should build on highly-successful Human Genome Model*
 - The evidence base for medical care could include clinically rich, carefully collected (de-identified) data *from millions of patients* in computerized databases. (63,000 trials registered in ClinicalTrials.gov)
 - Clinical medicine needs to develop a robust learning model, a la physics & other sciences, that uses:
 - Experiments (RCTs) +
 - Observational data (research databases) → that lead to
 - Predictive models – for personalized medicine

What's Possible: Example #3

- **A national Medicare cancer registry**
 - A uniquely valuable research database on cancer
 - Includes clinically rich data on 750,000 new Medicare patients diagnosed with cancer annually
 - For Research
 - To advance cancer science and care as quickly as possible.
 - To personalize cancer care
 - For Quality
 - To implement a national cancer quality reporting system, starting with breast cancer and colorectal cancer (National Quality Forum)

Summary

- Some parts of the US health system have been leaders in using IT to advance science
- New HIT investments (EHRs, NCI's BIG Health Consortium, effectiveness research) make possible extraordinary advances in medical research. National data policy must be addressed
- A rapid learning health system may provide lessons to advance science, technology, and economic growth.