Big Data: Past, present, & future perspectives from an integrated delivery system—Group Health

AHCJ conference, Santa Clara, CA

Eric B. Larson, MD, MPH
Executive Director
Group Health Research Institute (GHRI)
Vice President for Research
Group Health Cooperative
Seattle

April 24, 2015
Key points

• What do we mean by Big Data?

• Big Data: *not* new. Larger-scale efforts and value: *new*.

• Most patients support sharing their data with some concerns.

• Big Data alone can’t improve care or health

• Big Data: prone to misleading claims and reports.
What do we mean by Big Data?

• Broad (many people) &/or deep (many measurements per person)

• Structured (and unstructured) information from various sources:
  • Patients
  • Electronic medical records (EMRs)
  • Medical charts (natural language processing)
  • Administrative claims
  • Tests & results
  • Self-reported & self-collected data
Big Data in integrated care: Group Health

• Group Health, a pioneering nonprofit HMO, has had Big Data on a stable population since mid-1970s, when it computerized prescription refills, lab tests, & diagnoses.

• Advantages: Representative population, ready access to vast amounts of health care & health data with “complete” capture (HMO).

• Group Health’s Big Data: attractive to outside researchers. (GHRI established in 1983.)

• Data, research, & advances = Reputation for innovation & aspirational concept of learning health care system.
Our vision for Big Data: Learning health care system

In a learning health care system, research influences practice and practice influences research.

EVALUATE
Collect data and analyze results to show what works and what doesn’t.

IMPLEMENT
Apply plan in pilot and control settings.

ADJUST
Use evidence to influence continual improvement.

DESIGN
Design care and evaluation based on evidence generated here and elsewhere.

DISSEMINATE
Share results to improve care for everyone.

INTERNAL AND EXTERNAL SCAN
Identify problems and potentially innovative solutions.

Figure from Group Health Research Institute, in Greene SM, et al. Annals of Internal Medicine, 2012.
Cautionary tale:  
Big Data can mislead

An early lesson:


• Babies with defects from planned pregnancies—with no exposure to spermicides (1986). Definitive study disproved original finding (1987).

• Outside researchers knew only “data.”

• Familiarity with data & local practice can be critical to avoid mistakes.
Local & global: Big Data @ Group Health

Broader data networks developed, e.g.:

NCI’s Cancer Research Network & CDC’s Vaccine Safety Datalink:

• Integrated systems share data

• Solution: a “federated” data model, not centralized model.

• Virtual Data Warehouse (VDW): Data stay local, and programs written to extract “de-identified” data.
Similar current activities

- HMO Research Network (HMORN)
  - 20 health systems combine information on diverse population of 15 million people for statistical power to answer pressing questions
Similar current activities, cont.

• National Patient-Centered Clinical Research Network (PCORnet): Patient-Centered Outcomes Research Institute (PCORI)

• FDA’s Sentinel Initiative
  • Rapid-response electronic safety-surveillance system monitors drugs, devices, and vaccines in 120 million people

• Electronic Medical Records and Genomics (eMERGE) Network:
  • National Human Genome Research Institute (NHGRI) project with 10 institutions
  • Links genomic data from existing biobanks to EMR data to discover genetic causes of disease
  • Uses VDW approach but with standardized download of limited data to public Big Data repository: database of Genotypes and Phenotypes (dbGaP)
The importance of trust

- **Current Examples:** FDA’s Sentinel & NIH’s Health Care Systems Research Collaboratory using rapidly conducted pragmatic clinical trials

- **Problem:** Overly burdensome “research” oversight

- **Solution:** “Building Trust in the Power of Big Data: Research to Serve the Public Good”

“Routinely collected data provide great potential for extracting useful knowledge to achieve the triple aim in health care.”

—2013 IOM Clinical Effectiveness Research Innovation Collaborative (CERIC) report
Patients’ perspectives on Big Data

• Most engaged patients are **enthusiastic** about sharing their data—but have some **concerns**.

• **Trust** matters.

• **Local** trust, relationships, & engagement help.

• One prominent concern expressed: **Profit making**
How to build trust in using routine health care records data

• Local control (VDW) & patient engagement

• Restricted virtual and physical access: “need-to-know principle”

• Culture and policies that respect patient privacy

• Commitment that research serves the public good

“Consider how the consent process could foster respectful engagement rather than merely mitigate risk.”

—Trinidad et al., Science 2011, etc.
Efficient use of available data will speed discovery & translation

Use of electronic health records (EHRs)
- EHRs allow efficient and cost-effective, recruitment, data collection, & participant communication, monitoring, & follow-up.

Randomize treatment alternatives based on normal health care operations
- This sometimes mean randomizing at clinic or provider level ("cluster randomization").
Synchronize your watches

• Increased availability of routine health care data speeds research up.

• Studies that once required decades of data collection can now be accomplished in just months.

• Goal: *quick*—but not *dirty*. 
But Big Data will not automatically improve health & health care

Big Data must be transformed into usable and actionable information that can be used to improve health—& health care’s quality and safety.

- Clinicians need 24/7 decision support based on high-quality, generalizable science from representative populations.

- Patients & general public will also increasingly have that decision support directly at hand.

- Process will involve shared decision making & support behavior changes.
How will we distinguish responsible decision support providers from vast amount of bogus material in cyberspace today?
Misleading report: Alzheimer’s: even small rise in risk creates fear

• January 2015 paper went around world in hours: United Kingdom front-page news, sensationalizing findings, unleashing panic.

• National Health Service created factsheet to correct misimpressions created by sensationalistic press coverage.

• No dangerous anticholinergics in Benadryl in UK.
Association of Health Care Journalists: Help!

• Big Data & Internet have tremendous persuasive power.
• Too much self-promoting advice & products—from private & public (including academic) sectors
• Overdiagnosis, overtreatment, & medicalization of everyday life = huge threat.
• Honest brokers are needed, so quick Big Data don’t lead to quick & dirty hype.

“Improving public understanding of health and health care”

——AHCJ tagline is ideal for this task
Thank you.

Eric B. Larson, MD, MPH
larson.e@ghc.org
206.287.2988

grouphealthresearch.org
@GHResearch