Health Information Exchanges and Big Data: Challenges and Opportunities

Hadi Kharrazi, MHI, MD, PhD

Johns Hopkins University

hkharrazi@jhsph.edu
The presentation in a nutshell

- Introduction
- Big Data and Healthcare
- HIE History, Architecture and Services
- HIE Examples (Indiana HIE, CRISP)
- HIE Big Data Exploration (translating CPGs into population health data)
- HIE and Population Health IT Framework
- JHU Center for Population Health IT (CPHIT)
Introduction

- **Hadi Kharrazi**, MHI, MD, PhD
  - Assistant Professor, Johns Hopkins School of Public Health
    Department of Health Policy and Management
  - Joint Appointment, Johns Hopkins School of Medicine
    Division of Health Sciences Informatics
  - Assistant Director
    Johns Hopkins Center for Population Health IT (CPHIT)
  - **E:** kharrazi@jhu.edu
  - **L:** https://www.linkedin.com/in/kharrazi/
  - **W:** https://jhsph.edu/cphi
  - Research interests: Population DSS, HIE & ACO Analytics
Big Data and Healthcare

Statistics:
- Data size: 2012 = 500 petabytes → 2020 = 25,000 petabytes
- Cost: 2009 = 17.6% GDP ($2.9t) → 2025 = 25.0% GDP
- Big Data Saving: ~$300b/yr

Definition:
- Big Data is a collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications.

History in Healthcare
- Last time: We pushed the data boundary in the Human Genome Project (1990)
- This time: Massive roll out of EHRs (Meaningful-Use) and Integration of various source of data (genomic, mobile, patient-generated, exchanges)
Big Data and Healthcare (cont.)

- Big Data Specs – data driven 4Vs:
  - Volume $\rightarrow$ quantity (size)
  - Variety $\rightarrow$ type (structure, standardized, ontology)
  - Veracity $\rightarrow$ quality (meaning, completeness, accuracy)
  - Velocity $\rightarrow$ time (real-time, timeliness)
Big Data and Healthcare (cont.)

Biomedical informatics methods, techniques, and theories

- Bioinformatics
- Imaging Informatics
- Clinical Informatics
- Consumer Health Informatics

Population HIT

Public Health Informatics

Basic Research

Biomedical informatics as a basic science

Molecular Research

Health Research

Applied Research
Big Data and Healthcare (cont.)

- Big Data Specs – clinical driven 5Ms:
  - Measure → size, type, quality (QMs, digitization, ...meaningful use)
  - Mapping → integration, interoperability (information exchanges, EDWs)
  - Methods → analytics (exploration, visualization, predictive, ...replacing RCTs)
  - Meanings → knowledge (EBM, CPG, ...meaningful use)
  - Matching → outcomes (triple aims, ACA)
Big Data and Healthcare (cont.)

- Johns Hopkins

---

**Epic**

- Measure -
- Mapping -
- Methods ??
- Meaning ??
- Matching ??
HIE History
HIE History (cont.) > ONC

ONC relationship with DHHS
HIE History (cont.) > Health Information Organization (HIO)

- **HIE (verb):** The electronic movement of health-related information among disparate organizations according to nationally recognized standards in an authorized and secure manner.

- **HIO (noun):** An organization that oversees and governs the exchange activities of health-related information among independent stakeholders and disparate organizations according to nationally recognized standards in an authorized and secure manner.

- An HIO can be described by many acronyms, including:
  - State Level Health Information Exchange (**SLHIE**)
  - Regional Health Information Exchange (**RHIO**)
  - Regional Health Information Network (**RHIN**)
  - Health Information Exchange Networks (**HIE[N]**)
  - Others: Integrated Delivery Systems (**IDNs**); Physician practices HIEs; Payer-led HIEs; and, Disease-specific HIEs.
HIE Architecture > Centralized

Centralized HIE

Data Repository

© Hadi Kharrazi @ JHSPH-HPM
HIE Architecture (cont.) > Federated

Federated Inconsistent HIE
HIE Architecture (cont.) > Federated

Federated Consistent HIE
HIE Architecture (cont.) > Hybrid

Hybrid HIE
HIE Architecture (cont.) > Switch

EHRs → Data Standardization Machine → PHRs

Claims → Data Standardization Machine → Registries

PACS → Data Standardization Machine → LABs

Switch HIE

Copyright Regenstrief Institute
HIE Architecture (cont.) > Patient Centric

Patient Centric HIE (e.g., PHR controlled)
HIE Architecture (cont.)

Health Information Organizations / Regional HIOs

Copyright Regenstrief Institute
Connecting RHIOs / NwHIN
HIE Architecture (cont.) > NwHIN


Nationwide Health Information Network (NwHIN)
HIE Services > Core Services

- **Presentation Services**: login, patient look-up, request patient records, view data

- **Business Application Services**: e-Prescribing, EMR, lab, radiology, eligibility checking, problem list/visit history

- **Data Management Services**: data persistence/access, value/code sets, key manage.

- **Data Storage Services**: message logs, XML Schemas, Provider/User Directory

- **Integration Services**: message translation/transport, HL7 mapping, EMR adapter

- **System Management Services**: system config, audit/logging, exception handling
HIE Services > Data Services by Constituency

- **Hospitals:**
  - Clinical messaging
  - Medication reconciliation
  - Shared EHR
  - Eligibility checking

- **Physicians:**
  - Result reporting
  - Secure document sharing
  - Shared EHR
  - Clinical decision support
  - Eligibility checking

- **Laboratory:**
  - Clinical messaging
  - Orders

- **Public Health:**
  - Needs assessment
  - Biosurveillance
  - Reportable conditions
  - Results delivery

- **Consumers:**
  - Personal Health Records

- **Researchers:**
  - De-identified longitudinal clinical data

- **Payers:**
  - Quality measure
  - Claims adjustment
  - Secure document transfer
HIE Services > Emerging Services

- Next Generation Analytics
  - Data warehouse, data analytics and business intelligence
  - Quality reporting support
  - Performance management
  - Fraud and abuse identification and prevention
  - Care gap identification
  - Care and disease management
  - Public health monitoring and analysis
  - Population monitoring and predictive profiling
The Indiana HIE (IHIE) includes (as of mid-2011):

- Federated Consistent Databases
- 22 hospital systems → ~70 hospitals
- 5 large medical groups and clinics & 5 payors
- Several free-standing labs and imaging centers
- State and local public health agencies
- 10.75 million unique patients
- 20 million registration events
- 3 billion coded results
- 38 million dictated reports
- 9 million radiology reports
- 12 million drug orders
- 577,000 EKG tracings
- 120 million radiology images
HIE Examples (cont.) > CRISP (Chesapeake Regional Info. Sys. for our Patients)

- **Focus Areas:**
  - Query Portal Growth
  - Direct Secure Messaging
  - Encounter Notification System (ENS)
  - Encounter Reporting System (ERS)
  - Health Benefits Exchange integration

<table>
<thead>
<tr>
<th>Progress Metric</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Organizations Live</strong></td>
<td></td>
</tr>
<tr>
<td>Hospitals (Total 48)</td>
<td>48</td>
</tr>
<tr>
<td>Hospital Clinical Data Feeds (Total 143 - Lab, Radiology, Clinical Docs)</td>
<td>86</td>
</tr>
<tr>
<td>National Labs</td>
<td>2</td>
</tr>
<tr>
<td>Radiology Centers (Non-Hospital)</td>
<td>5</td>
</tr>
<tr>
<td><strong>Identities and Queries</strong></td>
<td></td>
</tr>
<tr>
<td>Master Patient Index (MPI) Identities</td>
<td>~4M</td>
</tr>
<tr>
<td>Opt-Outs</td>
<td>~1500</td>
</tr>
<tr>
<td>Queries (Past 30 Days)</td>
<td>~3500</td>
</tr>
<tr>
<td><strong>Data Feeds Available</strong></td>
<td></td>
</tr>
<tr>
<td>Lab Results</td>
<td>~16M</td>
</tr>
<tr>
<td>Radiology Reports</td>
<td>~5M</td>
</tr>
<tr>
<td>Clinical Documents</td>
<td>~2M</td>
</tr>
</tbody>
</table>
Patient has diabetes and has not had an eye exam in two years. Based on guideline xxx you may want to consider asking for an eye consultation.

```php
$results = mysql_fetch();
$patient_id = $results['patient_id'];
mysql_query ('SELECT * FROM lab WHERE patient_id = $patient_id' and lab_term = 'HBA1c');
$results = mysql_fetch();
$lab_result = $results['lab_term'];
IF ($lab_result > 9){
    ECHO “Patient has a high HBA1c level”; 
}
```
HIE Big Data Exploration (cont.)

Clinical Practice Guidelines

Computerized CPG (CARE, Arden syntax)

Manual Programming for each CPG

One patient at the time

Relational DB 1

CDSS SQL 1

NQF QM eMeasures

Automated for all CPGs

All patient at the time

Relational DB 2

CDSS SQL 2

Relational DB n

CDSS SQL...
**HIE Big Data Exploration (cont.)**

- **CPG**
- **NQF**
- **CPG Consortium**
- **CARE Arden**
- **XML**
- **Lexical Integration**
- **Syntactic Mapping**
- **Semantic Integration**
- **XSLT trans.**

Generating Renewable CDSS

1. CDSS SQL 1
2. CDSS SQL 2
3. CDSS SQL... (n)

Relational DB 1
Relational DB 2
Relational DB... (n)

Pragmatic Integration
HIE Big Data Exploration (cont.)
**HIE Big Data Exploration (cont.)**

<table>
<thead>
<tr>
<th>Rules</th>
<th>Population</th>
<th>One patient*</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD4_1.sql</td>
<td>59</td>
<td>0.18</td>
</tr>
<tr>
<td>CHOLESTEROL_1.sql</td>
<td>180</td>
<td>0.11</td>
</tr>
<tr>
<td>CHOLESTEROL_2.sql</td>
<td>1</td>
<td>0.12</td>
</tr>
<tr>
<td>DIABETES_1.sql</td>
<td>541</td>
<td>0.56</td>
</tr>
<tr>
<td>DIABETES_2.sql</td>
<td>1561</td>
<td>0.47</td>
</tr>
<tr>
<td>DIABETES_3.sql</td>
<td>182</td>
<td>0.40</td>
</tr>
<tr>
<td>DIABETES_4.sql</td>
<td>121</td>
<td>0.40</td>
</tr>
<tr>
<td>DIABETES_5.sql</td>
<td>1620</td>
<td>1.51</td>
</tr>
<tr>
<td>TCL_2.sql</td>
<td>67</td>
<td>0.12</td>
</tr>
<tr>
<td>TCL_3.sql</td>
<td>69</td>
<td>0.18</td>
</tr>
<tr>
<td>TCL_4.sql</td>
<td>61</td>
<td>0.17</td>
</tr>
<tr>
<td>TETANUS_SHOT_1.sql</td>
<td>1</td>
<td>0.07</td>
</tr>
<tr>
<td>WHOLE_1.sql</td>
<td>127</td>
<td>0.16</td>
</tr>
<tr>
<td>WHOLE_2.sql</td>
<td>129</td>
<td>0.13</td>
</tr>
<tr>
<td><strong>Average (sec)</strong></td>
<td><strong>1249</strong></td>
<td><strong>0.73</strong></td>
</tr>
</tbody>
</table>

Runtime per second

* each row shows average of 20 individual runs
HIE Big Data Exploration (cont.)

Applying all applicable CARE rules for Institute-1

*Some rules are never hit for institute 1 (not shown in this diagram)*
HIE Big Data Exploration (cont.)

Institute versus Rules (Raw Numbers)  
(zero rules / zero institutes are removed)
HIE Big Data Exploration (cont.)

Institute 1: Percentage of Hit / Miss Rules

Missed applicable CARE rules for institute-1
(CDSS-generated QM fingerprint)

Manual process of deciphering the rules (hit-means-miss OR hit-means-hit-only)
The lower the better (less missed)
**HIE Big Data Exploration (cont.)**

**Institute 2: Inverse Percentage of Hit / Miss Rules**

Missed applicable CARE rules for institute-2
(CDSS-generated QM fingerprint)

*Manual process of deciphering the rules (hit-means-miss OR hit-means-hit-only)*

*The lower the better (less missed)*
HIE Big Data Exploration (cont.)

Missed rate by institute

*The effect size of the missed rules depends on the population size.*
HIE and Population HIT > Population Health Informatics

- Population Health Info.
  - Population Health Analytics
  - Population Health Data-warehouse
    - Personal Health Records
    - Insurance Data
    - Rx Data
    - Admin data repositories

CI to PubHI
PubHI to CI
CI to PubHI
PubHI to CI

Clinical Informatics
EHRs
1...n

HIE

Public Health Informatics
A
B
C
Population-based HIE

EHRs  →  MPI  →  PHRs

Claims  →  MPI  →  Registries

PACS  →  MPI  →  LABs

Population-based CDSS across databases
The Johns Hopkins Center for Population Health Information Technology
(CPHIT, or “see-fit”)

- The mission of this innovative, multi-disciplinary R&D center is to improve the health and well-being of populations by advancing the state-of-the-art of Health IT across public and private health organizations and systems.

- CPHIT focuses on the application of electronic health records (EHRs), mobile health (m-health) and other e-health and HIT tools targeted at communities and populations.

- Director: Dr. J. Weiner

- www.jhsph.edu/cphit
CPHIT (cont.)

CPHIT Organizational Linkages

External PH/IDS Orgs.

JH Health System

JHU Academic Departments and other R&D centers

JH Healthcare Solutions, LLC

Industry Foundations Government

Business Partners

Copyright Dr. J. Weiner
CPHIT (cont.)

- **CPHIT’s R&D Priorities**

  - Development and testing of measures created from EHRs and other HIT systems to quantify quality, outcome and need within target populations and communities.
  
  - Use and advancement of computing methodologies – including natural language processing (NLP) and pattern recognition tools – to improve the application of non-structured EHR data for population-based interventions.
  
  - Initiation of effective approaches for linking provider-centric EHR systems with consumer-centric internet and mobile-based e-health applications.
  
  - Development of EHR-based tools and decision support applications to identify and help manage high risk populations for preventive and/or chronic care and to coordinate care within IDS/ACO.
  
  - Strategic approaches for creating an interoperable community of EHR networks and integrating them with the current functions of public health agencies (e.g., surveillance and vital records).
  
  - Creation of legal/ethical and policy frameworks to support the development of secondary use of EHRs for public health programs and research.
CPHIT (cont.) > Research Pipeline

- Developing and Testing "e-ACGs": Using EHR / HIT Data to Enhance Johns Hopkins ACGs' Ability to Measure Risk and Health Status

- Real-time, Cross-Provider High Readmission Risk Detection and Notification System

- Evaluate the economic impact of the Shared Patient Virtual Health Record (SPVHR) in Inpatient Settings

- Prescription Drug Monitoring Program (PDMP) Evaluation

- Developing and evaluating an Inter-Provider Health Screening and Quality Measure System (IPHSQMS)

- Identifying high risk pregnancies by linking claims and clinical data using natural language processing tools

www.jhsph.edu/cphit
Thank you

Q & A

www.jhsph.edu/cphit