Data Makes the Difference: The Future State of Data Collection

James E. Tcheng, MD
Professor of Medicine
Professor of Informatics
Duke University Health System

2019 Nov 11
Disclosures (none relevant to presentation)

• Advisor
  – GE Healthcare, Philips Healthcare, International Guidelines Central, Lumedx, Medstreaming

• Board Membership
  – Regenstrief Foundation, ABIM CV Specialty Board, HL7 Advisory Board, Livmor, Alliance for the Implementation of Practice Guidelines (chair)
Future State of Data Collection

• Are we there yet?
• Interoperability of what?
• The need for granular clinical data
• The 4 tenets of data capture
• The Duke journey
• Structured reporting – and you!
The View from the President’s Office

- 2004 - President Bush establishes a 10 year goal to develop the electronic health record (EHR)
- 2009 - President Obama signs ARRA, pushes EHR adoption through incentives, targets full implementation by 2016
Magical Thinking

... a disorder of thought content ... the false belief that one's thoughts, actions, or words will cause a specific consequence in some way that defies commonly understood laws of causality

_Wikipedia_
10 Years & $36+ Billion Dollars Later ... 

**Are We There Yet?**

<table>
<thead>
<tr>
<th>Envisioned</th>
<th>Reality</th>
</tr>
</thead>
<tbody>
<tr>
<td>EHR “Meaningful Use”</td>
<td>EHR meaningless burden</td>
</tr>
<tr>
<td>Usability and productivity</td>
<td>Death by a thousand clicks</td>
</tr>
<tr>
<td>Patient engagement</td>
<td>AVS drivel</td>
</tr>
<tr>
<td>Effective clinical care</td>
<td>CDS trivial pursuit</td>
</tr>
<tr>
<td>Population health</td>
<td>Resource consumption focus</td>
</tr>
<tr>
<td>Bending healthcare cost curve</td>
<td>Cost control and penalties</td>
</tr>
<tr>
<td>Better provider work life</td>
<td>NOT!</td>
</tr>
<tr>
<td>Torrent of real-world data</td>
<td>Puddles of document exchange</td>
</tr>
<tr>
<td>Big (clinical) data analytics</td>
<td>Transactional (admin) data</td>
</tr>
<tr>
<td>Leveraged RCTs via registries</td>
<td>Electronic bridge to nowhere</td>
</tr>
</tbody>
</table>
We Have Robust Evidence & Guidelines--Why Registries?

Evidence → Guidelines → Clinical Practice
“Science tells us what we can do; Guidelines what we should do; Registries what we are actually doing.”
Data Demand: Multiple Masters

- Registries
- Health system
- Payers
- Patients
- Federal, State regs.
- FDA
- Research
- Machine learning, AI ...
- Oh yes ... clinicians

... who are time-challenged, short-staffed, overloaded with information and have increasing expectations placed upon them.
ARRA HITECH
HIT Standards Committee

Clinical Operations Workgroup Report

Jamie Ferguson, Chair
Kaiser Permanente

John Halamka, Co-chair
Harvard Medical School (& HITSP)

20 August 2009
Clinical Operations is recommending standards for interoperability between entities, not within an entity.

Recommended standards should not apply to internal data capture, storage or uses – only to external representation and data exchange between entities.

Content should be able to be represented in the specified vocabularies and exchanged in the specified standards at the boundary between entities, regardless of how it is managed internally.

Many methods may potentially be used to achieve interoperability standards, e.g., mapping, external services, or native data capture.
Boundary-Based Interoperability

<table>
<thead>
<tr>
<th>Focus on recording clinical content</th>
</tr>
</thead>
</table>

- **SNOMED Clinical Terms (SNOMED CT)**
  - International Health Terminology Standards Development Organization (IHTSDO)

- **Logical Observation Identifiers, Names and Codes (LOINC)**
  - Regenstrief Institute for Healthcare

- **RxNorm**
  - National Library of Medicine

- **International Classification of Diseases – Clinical Modification (ICD-9/10-CM)**
  - World Health Organization
  - National Center for Health Statistics

- **Current Procedural Therapy (CPT)**
  - American Medical Association

Focus on reimbursement
Search Term: myocardial infarction
Returns 308 matches in 2.33 seconds
Term defined by pathologic, anatomic relationships
No clinical definition

SNOMED-CT
Problems with Boundary-Based Interoperability

• Duke participates in ~20 CV registries
  – ETL, ETL, ETL, ETL, ETL
• (Lack of) vocabulary specificity
  – E.g., ICD-10, SNOMED-CT
• (Lack of) clinical vocabulary
  – EHR (text-based) documentation lacks discipline to capture information per se, as well as information as data
How Registries Solve the Data Capture Problem

Standardized NCDR data elements and processes

The CathPCI Registry uses standardized data elements and definitions for:

- Patient demographics for diagnostic coronary angiography and percutaneous coronary intervention (PCI) procedures
- Patient history/risk factors, cath lab visit indications and coronary lesion information
- Provider and facility characteristics
- PCI Indications, lesion information, intracoronary device utilization and intra/post-procedure events
- 30-day and 1-year follow-up information on patients who had PCI

The registry supports a variety of data entry and submission options including certified third-party vendors and secure web-based entry. Data collection options

Swivel Chair Interoperability  

Wes Rishel

Clinical Systems  \[\rightarrow\]  Registry Data Entry
What is Artificial Intelligence?

The theory and development of computer systems able to perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.
Machine Learning Predictive Modeling Based on Labeled Data

End goal: predict $y$ from $x$

$x$, data/features for a subject

$y$, associated label 0/1
Extended Logistic Regression

\[ \sigma(\zeta_i) \]

Probability of particular outcome

\[ \sigma(z_{i1}) \quad \sigma(z_{i2}) \quad \ldots \quad \sigma(z_{iK}) \]

Probability of \( K \) latent processes

May be viewed as logistic regression on \( K \) latent features, rather than directly on the \( M \) components of raw data.
Well-Formed Clinical Data Today

- Transactions (things you can bill for)
  - Events (e.g., clinic visits, medication administration, chest x-ray, surgical procedure)
  - Orders (CPOE)
  - Prescriptions (a type of order)
  - ICD-10 diagnoses (EHR MU)

- Automated data (machine output)
  - Laboratory results
  - Devices (smart watch, ICU monitoring)

- But granular clinical data ... is MIA
Not So Well-Formed Data ... for Clinicians, Quadruple Aim, Registries, Industry, Regulatory, Science, ...

• Clinical documents – text!
• Clinical information
  – Medical history, risk factors, social history, family history, physical exam, test interpretations, events, complications, assessment, plan
• Qualified / summative assessments
  – E.g., rationale for no treatment
• Patient-reported information
• ...
The Importance of Data Quality

\[ \int_{i=0}^{\infty} \text{garbage} = \text{garbage} \]

D Fridsma, 2018
@PaulLomax: The most unbelievable aspect of the Star Trek universe is that every ship they meet has compatible video conferencing facilities ...
The Four Tenets of Data Capture

• Capture high quality “good” data once, use many times
• Collect data at the point of care using a team-based, high usability system
• Use the computer (not humans) to abstract and compile views of the data
• Reduce clinician cognitive burden
THE Foundational Issue

Tower of Babel

Pieter Bruegel the Elder and Pieter Bruegel the Younger, 1563
Steps to Native Data Interoperability

Clinical concepts as data elements

Professional societies
Academic consortia
FDA

Data elements as database specifications

Informatics modeling
Regulation (ONC, ASC X12)
HIT vendors

Capture of data per db specs integrated into workflow

HIT vendors
Healthcare entities
Professional societies
Exchange, Use, and Reuse of Data Requires Shared Data Definitions (including semantics)

Native Data Capture → Native Data Interoperability
“Dammit, Jim, I’m a Doctor, Not a Computer!”
Concurrent Data Acquisition Facilitated Abstraction

Quality of Data

Lower

Higher

Quarters

Days

Time from Data Collection to Use

Data Abstraction Models

Facilitated Abstraction

Traditional “Swivel Chair” Abstraction

Concurrent Data Acquisition
What is **Structured Reporting**?

- Specific data captured by the person closest to that data, integrated into clinical workflow (e.g. MA, tech, RN, pt)
- Informatics formalisms: universal, well-defined common data elements; data model that parallels (i.e., is representational of) clinical care model
- Data is compiled by the computer to produce most of the content in a report; MD focuses on assessment of data quality, cognitive interpretation
- Output: the *structured report*
- ROI: ↑ data quality /quantity, ↓ redundancy / repetition, time to final reports, FTE requirements → augmented knowledge, financial gains
Duke Heart Center - Dataflow End State

Build infrastructure

Use the data

Near Real Time Clean Up

HIT / EHR (POC Form)  Discrete Data (CDEs)  Structured Documentation  DQR Credible Data  Analysis, Measures  Benchmark Registries

Heart Data Mart

Research

Active Quality Improvement Cycle
How Is Structured Reporting Done?

- Engineered, best-practice workflows
- Just in time, context specific, high usability, point of care data capture via forms
- Lots of business rules
- Optimized IT form factors
- Computer is a compiler

In other words ...

- Command of who does what when, where, and how
Episode of Care: Invasive & Interventional Cardiac Cath

Clinician evaluation
1. History of present illness (presentation)
2. Past medical history (risk factors)
3. Past clinical events, procedures
4. Social history, family history, ROS, PE

Functional testing
Diagnostic cath
Dx cath-specific info: operator, dx cath clinical status, results / findings, recommendations

Common data (dx + intervention)
Combined info: start/stop time, start vitals, access site, contrast, radiation, IABP timing, compassionate use

Interventional cath
Interv cath-specific info: operator, interv clinical status, meds STEMI details, lesions, interventions, outcomes, complications

In-hospital follow-up
Long-term follow-up

Data Tables
Demographic
Episode_Of_Care
History_Past (med, surg, risk factors)
Current_Common(sxs, meds)
Current_STEMI
Current_Not_STEMI
Proc_Cardiac_Cath (shared dx + PCI)
Proc_Cardiac_PCI (PCI only)
Coronary_Anatomy
[Proc_Meds]
[Proc_Labs]
InHospital_Events
Discharge

ELECTIVE
ACS/STEMI

Pre-Procedure
Diagnostic
Intervention
Follow-up
[new sx$s, events, staged cath (~5% in-hospital)]

outpatient
inpatient
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Information Sources</strong></td>
<td>History &amp; Physical Other documents Laboratories</td>
<td>Existing clinical data History &amp; Physical Other documents Laboratories</td>
<td>History &amp; Physical Other documents Laboratories Consents</td>
<td>Pre-procedure evaluation packet Hemodynamics Catheterization images</td>
<td>Hemodynamics Catheterization images Measurements Calculations</td>
</tr>
<tr>
<td><strong>Information Captured as Digital Data</strong></td>
<td>Patient identifiers Demographics Diagnosis Laboratories Insurance</td>
<td>Patient identifiers Demographics History Physical Exam Previous studies Laboratories Diagnosis</td>
<td>Patient identifiers Height, weight, vital signs Medications</td>
<td>Patient identifiers Procedures Hemodynamics Findings Measurements Medications Inventory</td>
<td>Patient identifiers Cath results Interpretation Tree diagram</td>
</tr>
<tr>
<td><strong>Actors</strong></td>
<td>Physician requestor Scheduling hub / Communications Center</td>
<td>Advanced Practice practitioners Physician operator</td>
<td>Outpatient / inpatient nurses</td>
<td>Physician operator Cath lab nurses Cath lab technologists</td>
<td>Physician operator</td>
</tr>
<tr>
<td><strong>Information Systems</strong></td>
<td>Registration system Scheduling app Electronic Health Record</td>
<td>Electronic Health Record Procedure Reporting system</td>
<td>Electronic Health Record</td>
<td>Radiography Modality Hemodynamic Monitoring system Procedure Documentation / Reporting system</td>
<td>Procedure reporting system</td>
</tr>
<tr>
<td><strong>Form Factor (for Actors)</strong></td>
<td>Desktop workstation</td>
<td>Mobile tablet</td>
<td>Bedside workstation</td>
<td>Multiple workstations: Radiography Modality Hemodynamic Monitoring Procedure Documentation</td>
<td>Desktop workstation</td>
</tr>
<tr>
<td><strong>Data Output</strong></td>
<td>Schedule – to scheduling app Orders – to Electronic Health Record (EHR) system</td>
<td>Clinical data – to procedure reporting system (history section) Patient status – to scheduling system → electronic schedule Orders – to EHR</td>
<td>Nursing documentation – to EHR Patient status – to scheduling system → electronic schedule</td>
<td>DICOM Modality Worklist to Modality, Hemodynamic, and Procedure Documentation systems → procedure log report; and data for procedure report (procedure section) [See also IHE CATH, CRC profiles]</td>
<td>Procedure results – to procedure reporting system (results section) → structured procedure report</td>
</tr>
</tbody>
</table>
Near Real Time Clean Up

Concurrent Data Acquisition Model and Data Flow Into Registries

Clinical data collected during case

Medical History event

- Type I diabetes
- Type II diabetes

Cath/PVI/EP/CT Surgery report

- Type I Diabetes

Registry data mapping rules

If Type I = Yes then Diabetes = Yes

Registry submission

Diabetes = Yes

Unidirectional carry forward (CF) with CF button trigger N+1

Unidirectional CF with automated CF, plus “Lock/Unlock” data, plus ad hoc CF button trigger

Frequent / real time, accurate, missingness reports, real time metrics, real time risk calculation

Registry Nurse

- Data quality
- Performance improvement
What Does SR Fix?

• MINIMAL CHART ABSTRACTION
• Single source of data (trust and verify)
• Reusable data – “collect data once, use many times”
• Explicitly prompts for presence / absence of data – not just charting by exception
• MD emphasis on findings, results, interpretation, recommendations – not “art”
• ↑ workflow efficiencies, ↓ FTEs
• ↑ Clinical data, data quality, completeness
What is Needed for **Structured Reporting**?

- Unified clinical vocabulary $\rightarrow$ interoperable data standards
  
  -- clinical use to CVIS to EHR to registry
What is Needed for **Structured Reporting**?

- New MD professionalism standards
  -- conversion from dictation to information model
What is Needed for Ubiquitous Structured Reporting?

1. MD, staff, professional society transformation
   -- conversion from dictation to information model

2. Government, payer, health systems transformation
   -- shift emphasis from payment to data

3. Informatics: common data elements (CDE) →
   controlled vocabularies; common data model (CDM);
   data interoperability (HL7, IHE, etc.)

4. Clinical industrial (process) engineering to describe,
   model, implement best-practice workflows
   -- who does what when, where, and how
   -- implementation science, change management

5. IT platform, solution set
Key Duke CVIS RFP Requirements

- Information stored as data elements
  - Limited text (e.g., history, recommendations)
- Unfettered access to data (SQL database)
- High usability POC data capture
  - Development (programming) environment
- Vendor maintenance of data model
  - Data submission to registries
- Interoperability ready (e.g., UDI)
- Structured reporting, structured reports
A Little Behavioral Economics ...

Human frailties - and the need for “choice architecture”:

• Unrealistic optimism
  - If interoperability were that easy ...
• Loss aversion
  - Inertia favors stasis
• Status quo bias
  - “Easy Button” default option
• Framing effects
  • How to convince (“sell”)

Nudge
RICHARD H. THALER
WINNER OF THE NOBEL PRIZE IN ECONOMICS
and CASS R. SUNSTEIN
NEW YORK TIMES BESTSELLER
Improving Decisions About Health, Wealth, and Happiness
Concurrent Data Capture: Key Concepts

- Capture data once, use many times
- Directed data capture, relevant (pertinent) charting, prompted charting by exception
- Distributed data capture, integrated into workflow
- Team-based documentation
- Data persistence, data liquidity
- Data compilation into views (reports)
- On framework of semantic interoperability
- => Structured reporting
Three Tiered “Data In” Framework

<table>
<thead>
<tr>
<th>Iteration Time</th>
<th>System</th>
<th>Utility</th>
<th>Data Store</th>
</tr>
</thead>
<tbody>
<tr>
<td>III Months</td>
<td>Epic</td>
<td>3 stars</td>
<td>+ SQL Server</td>
</tr>
<tr>
<td>II Weeks</td>
<td>LUMEDX</td>
<td>5 stars</td>
<td>Cognos</td>
</tr>
<tr>
<td>I Days</td>
<td>REDCap</td>
<td>3 stars</td>
<td>MySQL</td>
</tr>
</tbody>
</table>

Production level slow rate of change

Structured Reporting
Medium Cycle Learning

API Data Integration

LUMEDX Data Warehouse

Rapid Cycle Learning

LUMEDX front end is Microsoft Screen build toolkit

Copyright Duke Heart ©
Duke “Data Out” Framework

Data Store

- Duke Health Data Warehouse
- LUMEDX Data Warehouse

Duke Heart

Data Lake

Exadata

LUMEDX

Self Service Reports
(SSRS/SAP BO)

Visualizations
(Tableau)

In-App Analytics
(LUMEDX + Epic)

Scheduled Direct Emails
(SSRS/SAP BO)

Self Service Datasets

SQL Server ODBC (Excel/SAS)

LUMEDX back end is really just SQL Server plus all its tools
Thank You!

james.tcheng@duke.edu

Visit the Project website:
https://dcric.org/registry-data-standards