Predictive Modeling Integrated in Healthcare, Toward the Learning Health System

Jonathan Silverstein, MD, MS, FACS, FACMI
Chief Medical Informatics Officer
Kanter Health Foundation
U.S. healthcare challenges in a slide?

- People are dying of preventable causes.
- Cost is out of control.
- Quality can’t be measured.
- Variability is local and widespread.
- New technology is exponentiating.
- Decision-making is maximally distributed.
- Data is not available routinely for learning.
Informatics and Predictive Analytics has fundamentally changed our world!

- Credit card fraud detection
- Target can predict when you are pregnant
- Traffic monitoring and prediction
- SPAM filters
- Smart meters and utility grids
- Weather forecasting
- Netflix and Amazon predict your tastes
- Mobile phones become interactive assistants
- Science’s third pillar of data and computation!

• WHAT ABOUT BIOMEDICINE?!
Universe of Biomedical Data

- LIFESTYLE
- ENVIRONMENT
- CLINICAL
- GENOMIC
Health Spending as a Share of GDP
United States, 1962 to 2022, Selected Years

Recent Detail

<table>
<thead>
<tr>
<th>Year</th>
<th>Spending</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>16.4%</td>
</tr>
<tr>
<td>2009</td>
<td>17.4%</td>
</tr>
<tr>
<td>2010</td>
<td>17.4%</td>
</tr>
<tr>
<td>2011</td>
<td>17.3%</td>
</tr>
<tr>
<td>2012</td>
<td>17.2%</td>
</tr>
</tbody>
</table>

Notes: Health spending refers to National Health Expenditures. Projections (P) include the impact of the Affordable Care Act.
Provider, Patient & Payor Faced With Bewildering Choices: The Current Practice of “Qualitative” Medicine

Has tumor spread?
Surgery or Chemotherapy?
What molecular subtype?
What stage?
What dose?
Pre-operative Chemotherapy?
What schedule?
In combination with which drugs?

From Patrick Soon-Shiong, MD
Patients with same diagnosis → Non-toxic responders

Non-responders, toxic responders

Misdiagnosed
The Future of Medicine is:

• Evidence based (data driven)
• Practice based (generation of data)
• Targeted and precise (ER+ Breast Cancer vs. Triple Negative Breast Cancer)
  – Personalization to individual mutations
  – Genomics for biomarkers and gene therapy
• A Learning Health System
  – “…gets the right care to people when they need it and then captures the results for improvement…” – Institute of Medicine
Trends in EHR adoption show increasing use of advanced functionality.

Figure 3: Percent of non-federal acute care hospitals with adoption of EHR systems by level of functionality: 2008-2014.

NOTES: Definitions of Basic EHR and Comprehensive EHR systems are reported in Table A1.
*Significantly different from previous year (p < 0.05).
What’s in a Health System Enterprise Data Warehouse?

- Extreme granularity of data capture - Terabytes of transactions of who did what to whom for millions of patients
- Demographics
- Medical history
- Orders
- Meds
- Procedures (e.g. details about implants)
- Test results (including molecular markers?)
- Vital signs, detailed characteristics (e.g. pain)
- Text of all notes
- Perhaps charges, Perhaps imaging, …
Digital Health Funding Snapshot Year Over Year

2015 saw a significant shift towards validated innovation. Today’s capital is supporting companies with proven leadership, strategy, business models and demonstrated outcomes. Although total funding is down from last year, there is more market impact per invested dollar than ever.

Source: StartUp Health Insights | startuphealth.com/insights Note: Report based on public data on seed, venture, corporate venture and private equity funding only. © 2015 StartUp Health LLC
A “Fundamental Theorem” of Biomedical Informatics

Data Collection, Modeling, Integration, and Learning Opportunities

• Collection of Discrete Clinical Data
  – System collects data in context-specific EHR workflows for research and quality improvement.

• Integration of EMR/EDW & External Systems
  – System exchanges data and extends workflow.

• Surveillance Algorithms
  – System monitors and recommends specific intervention for patients who meet known criteria.

• Predictive Algorithms
  – System recommends specific intervention for patients by analyzing thousands of similar patients’ data.
Statement of Problem

Clinical documentation is a rich source of information on interactions between the health system and individual patients.

Question: How can we capture this information **Consistently and Completely** for analysis—especially the interesting parts of progress notes?

Answer: Tools Balance Expressivity and Workflow
Three Different Approaches

Free Text
- Manual Chart Review
- Web Form

Abstract Data
- Parse and Abstract Data

Enter Data
- Structured Tools

Generate
- Database

Natural Language Processing
Discretization

the translation of continuous information into discrete data

“Patient reports dyspnea”
“X complains of Shortness of breath”
“Pt denies SOB”

“Examination revealed rales”
“PE result: Crackles”
“negative for lung noises”

Dyspnea: Yes/No
Rales: Yes/No

Analogies:
• analog signals vs. discrete waveforms
• film vs. JPG
Quality improvement and practice-based research in neurology using the electronic medical record

Demetrius M. Maraganore, MD
Roberta Frigerio, MD
Nazia Kazmi, MS
Steven L. Meyers, MD
Meredith Sefa, MS
Shaun A. Walters, MS
Jonathan C. Silverstein, MD, MS

Abstract
We describe quality improvement and practice-based research using the electronic medical record (EMR) in a community health system-based department of neurology. Our care transformation initiative targets 10 neurologic disorders (brain tumors, epilepsy, migraine, memory disorders, mild traumatic brain injury, multiple sclerosis, neuropathy, Parkinson disease, restless legs syndrome, and stroke) and brain health (risk assessments and interventions to prevent Alzheimer disease and related disorders in targeted populations). Our informatics methods include building and implementing structured clinical documentation support tools in the EMR; electronic data capture; enrollment, data quality, and descriptive reports; quality improvement projects; clinical decision support tools; subgroup-based adaptive assignments and pragmatic trials; and DNA biobanking. We are sharing EMR tools and deidentified data with other departments toward the creation of a Neurology Practice-Based Research Network. We discuss practical points to assist other clinical practices to make quality improvements and practice-based research in neurology using the EMR a reality. 

Genomic Health Initiative (efficient research)

Enrollment/consenting at pre-existing medical contact points
- Fully electronic through patient portal
- At routine blood draws

Multi-purpose consent
- Bio-sample collection
- Research
- Options for return of clinical data, & patient re-contacting

Biological specimens
- Consenting patients are ‘flagged’ for bulk ordering
- At next routine blood test, we collect the specimen

Processing of blood specimens
- Robotized methods
- Transfer of DNA aliquot to genomics lab

Integration with NS EHRs warehouse
- Fully identified protected linkage inside the health system

Genomics
- Results not part of EPIC
- Sharing of de-identified samples/data
- Back to patient is defined in consent
- Alignment with NS business plans
Health Monitors for Chronic Disease by Gait Analysis with Mobile Phones

Fig. 2. Prediction accuracy of Global Initiative for Chronic Obstructive Lung Disease (GOLD) status for different walk periods. For the continuous model, the spatiotemporal gait parameters are calculated over discrete periods of time. Thirty seconds of walking is selected as the optimal walk period to balance prediction accuracy with practicality.
What’s Going Around
Syndromic Surveillance Across the NorthShore Population

Influenza | Strep Throat | Pertussis | Pediatric Asthma | Gastroenteritis

Influenza-Like Illness as of November 06, 2013
Three-Year Lookback Period

Colors represent the proportion of all patients seen in the Medical Group who had the syndrome of interest. Areas with insufficient data for accurate estimates are not colored.

Below BLUE line = LOW syndromic activity. Above RED line = HIGH syndromic activity.

Age Specific Instances of syndrome. Grey represents 95% confidence intervals.

Sample Maps | Change Location | How WGA Works | Feedback

How What’s Going Around Works.

‘What’s Going Around’ (WGA) demonstrates a new way in which an Electronic Medical Record system can benefit providers and patients. We believe WGA is the first platform in the world to translate EMR-based data into up-to-date local epidemiological information, and provide this information to clinicians at the point-of-care, in a patient-specific manner.

Click here for full description.
Eliminating Undiagnosed HTN 2010-2012

Mike Rakotz, Chad Konchak, Ari Robicsek (and a cast of thousands)
Eliminating Undiagnosed HTN 2010-2012

Patients Newly Diagnosed with Hypertension by Primary Care

Research Pilot Study:
• 112 Physicians
• 24 Practices

MG Quality Program
• 211 Physicians
• 40 Practices

# of New Diagnosed Patients

# of New AOBP Readings

Mekhala Padwardan, Mike Rakotz, Chad Konchak, Ari Robicsek (and a cast of thousands)
"We seek the development of a learning health system that is designed to generate and apply the best evidence for the collaborative healthcare choices of each patient and provider; to drive the process of discovery as a natural outgrowth of patient care; and to ensure innovation, quality, safety, and value in health care."
Your Life, Your Health
Sharing your digital health data could save your life

JOSEPH H. KANTER
Kanter Health Foundation

• MISSION
  • To mobilize diverse organizations and people to collaboratively advance human health

• VISION
  • Every decision affecting health is informed by knowledge of *what works best*
Health systems--at any level of scale--become learning health systems when they can continuously study and improve themselves.
Checklist View: A Health System That Can Learn

✓ Every consensus patient’s characteristics and experience are available to learn from

✓ Best practice knowledge is immediately available to support decisions

✓ Improvement is continuous through ongoing study

✓ This happens routinely, economically and almost invisibly

✓ All of this is part of the culture

Charles Friedman
A Nationwide Learning Health System: A System of Health Learners Across Our Nation

All-Inclusive  Large Scale  Decentralized  Reciprocal

Governance
Engagement
Data Aggregation
Analysis
Dissemination

Healthcare Delivery Networks

Research Institutes

Tech Industry

Universities

Government/Public Health
Micro View: How Learning Happens
“Virtuous Cycles” of Study and Change

1. Collect Data
2. Assemble Data
3. Analyze Data
4. Interpret Results
5. Deliver Tailored Message
6. Take Action to Change Practice

Technology and policy for making knowledge persistent and sharable

Mechanisms for tailoring messages to decision makers
Mechanisms for capturing changed practice

Policies governing access to data
Mechanisms for communities of interest to form

Charles Friedman
Learning Systems Can Exist at Any Level of Scale

One organization

Many organizations

Several countries

A region or country

The world

Charles Friedman
Many “Use Cases”…
A Learning System Routinely Enables:

• **Pursuit of Best and Safer Care at Lower Cost:** Communities of interest discover what interventions are most cost-effective and are supported in implementing them.

• **Enhanced Public Health:** During an epidemic, new cases are reported directly from EHRs, the spread of the disease is predicted, and clinicians are alerted.

• **Consumer Empowerment:** Patients facing difficult medical decisions discover the experiences of other patients like them.

• **Enhanced Research:** Data of many types (all –omics) can be rapidly aggregated and analyzed.
105 Endorsements of the LHS Core Values*

(As of 6/7/2016)

*To be included on the www.LearningHealth.org website.
<table>
<thead>
<tr>
<th>Core Values of the Learning Health System</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Person Focused</td>
</tr>
<tr>
<td>2) Privacy</td>
</tr>
<tr>
<td>3) Inclusiveness</td>
</tr>
<tr>
<td>4) Transparency</td>
</tr>
<tr>
<td>5) Accessibility</td>
</tr>
<tr>
<td>6) Adaptability</td>
</tr>
<tr>
<td>7) Governance</td>
</tr>
<tr>
<td>8) Leadership</td>
</tr>
<tr>
<td>9) Scientific Integrity</td>
</tr>
<tr>
<td>10) Value</td>
</tr>
</tbody>
</table>

https://lillypad.lilly.com/entry.php?e=8284
A New International Journal

Learning Health Systems

A new open access journal dedicated to improved individual and population health

All articles are freely available to read, download and share

www.lhsjournal.com
The mission of CAPriCORN is to provide a data and informatics infrastructure to support collaboration among Chicago area scientists in conducting patient-centered outcomes research.
...to provide software applications and appropriate governance to facilitate the sharing of EHR data among healthcare organizations for research purposes
Collaborative data sets for specific studies

Specialty-focused data sets for ongoing research (e.g., consortium data sharing)

De-identified queries across entire community
LHS.network cohort

LHS.network is building a cohort of millions of patients’ full electronic medical records with the objective of answering the following question:

Among patients with my diagnosis and similar to me, what treatments did they choose and what were the outcomes?

Achieving our objective will proceed along two parallel projects:

- Assembling base set of millions of full electronic medical records from dead people (decedents)
  - **Key mechanism**: non-human subjects research can progress rapidly
- Building a living cohort of altruists who will both contribute their data and benefit directly
  - **Key mechanism**: Touching millions of people will be achieved by an efficient scalable consent
Research informatics platform aggregating decedent subject data from health systems and transforming it for use in generating hypotheses and piloting informatics methods.

The HITECH Act modified HIPAA to enable research access to decedent records under all circumstances with no method for patients or their representatives to opt-out.

PHI of Decedents {45 CFR 164.512(i)(1)(iii)}

§ 164.512 Uses and disclosures for which consent, an authorization, or opportunity to agree or object is not required.

(iii) Research on decedent's information. The covered entity obtains from the researcher:

- (A) Representation that the use or disclosure is sought is solely for research on the protected health information of decedents;
- (B) Documentation, at the request of the covered entity, of the death of such individuals; and
- (C) Representation that the protected health information for which use or disclosure is sought is necessary for the research purposes.
why do we need patient identifiers?

- Precise verification of records for any findings (e.g. Chart Review) and source verification is impossible with de-identified records without re-identification.
- We need to know the context of time and location in order to optimize the similarity metric between patients.
- Project goals are totally dependent on a similarity metric among people, so it is essential to de-duplicate all records before analysis.
- EMRs do not account for behavior and other social determinants of disease, so if we can determine with precise identification that one person lives in the same house as another, then we can assert with high confidence that many of the same social determinants of disease are similar without knowing what they are precisely.
Two Questions to Consider*

1. What can the LHS do for me?
2. What can I do for the LHS?

*In your role as a healthcare professional, as a patient/caregiver, as a citizen, and in other roles.