Health Analytics: Delivering Evidence-Based Insights and Decision Support

leidos
HEALTH

Data Science Solutions for Healthcare

Interoperable electronic health records (EHRs) can help healthcare organizations improve chronic disease management, increase operating efficiencies, transform their finances, and improve patient outcomes.

However, EHR implementations are in various stages of maturity across the country, and their benefits have not been fully realized. Healthcare decision-makers must make meaningful use of the data collected, available, and accessible in EHRs.

From Raw Data to Decision-Making

By optimizing the use of data accessible in EHRs, we can uncover hidden relationships and identify trends and patterns in this diverse and complex information to improve chronic disease management, increase operating efficiencies, and transform healthcare organizations' finances. Emerging capabilities in data science can identify useful, actionable — and perhaps unanticipated — recommendations to improve healthcare delivery from large volumes of diverse and complex data.

Data Science Framework and Process

Leidos Health has developed, implemented, refined and applied data science techniques to real-world problems in healthcare, national security, intelligence, commerce, and other fields. Success in these distinct domains relies on a framework of practice that is distinct from the general concept of "big data" in that it recognizes the tight coupling among three skill areas:

- Strong domain expertise and knowledge of the available data resources
- Analytical skills and techniques, based on mathematics, statistics, and machine learning
- Computer science and programming skills

This data science framework comprises four fundamental stages – *data collection, data curation, data analysis and data explanation*, each of which is underpinned with a firm understanding of the business or clinical problem and healthcare delivery environment. It is not uncommon to proceed through each stage to establish a basic data science workflow, then revisit earlier stages to explore new information and to accommodate shifts in the problem definition, environment resources, new data sources, additional analyses.

Data Collection – starts from the moment raw data are gathered from their originating sources to the time of their placement (possibly in a transformed state) in repositories appropriate for storage, retrieval, and analysis. It also includes the planning and systems required for ongoing management and security of the collected data.

Data Curation – perhaps the most important step in determining whether data will produce measurable business value. This stage includes the traditional extraction, transformation, and load (ETL) methods of both commodity and custom-developed big data management systems.

- **Improves and standardizes the quality of the raw data;** e.g., cleans data corrupted during collection and data values outside meaningful or permissible ranges, as well as reformatting and conversion of data values to standard units and formats. It may also include transformations such as conversion of audio files to text transcripts. These standardized data may be merged, aligned and linked, or aggregated into new records specific to an activity, process, or organization.

- **Enriches the dataset with the addition of new "features" generated from collected data.** These features may be as simple as summary statistics (e.g., mean and standard deviation), or as complex as a map generated from medication data of pharmacy locations that visualizes this one factor of patient compliance.
- **Generates new representations of data**, such as a network graph of clinical protocols and medications administered among patients created from clinician order entry records. Alternate data representations enable privacy protection, especially when an analytics project is not allowed the same access to data as the process that originated the data.

Data Analysis – includes identification, selection, and execution of algorithmic and computational models and methods. Exemplar methods of data analysis include: information retrieval (traditional searches or more complex Boolean queries of complex data relationships), classification, clustering, pattern recognition and matching, anomaly detection, predictive modeling, and multi-relational network analysis.

Successful data analysis solutions do not take a "kitchen sink" approach of applying analysis methods indiscriminately and hoping for good outcomes. Instead, they are driven by the organization's need for insightful and actionable intelligence and focus on producing relevant and accurate evidence to answer the general question "What is going on in my business today?" e.g.,

- **Prescribed question:** "How is my administrative budget tracking against plan?"
- **Performance question:** What are the factors driving readmission rates?

Data Explanation – provides decision-makers with the intelligence they need to take action. This can happen only if leaders have timely access to clear, accurate, and relevant data analytic conclusions in a form that aligns with their missions, needs, and expectations. Communicating analytic conclusions may employ any of the following strategies: data summarization, synthesis of analytic findings, narrative creation, forensic analysis, and visualization.

An example would be the application of forensic analysis, such as longitudinal analysis of a drug's efficacy in clinical trials to understand and explain how the drug might affect readmission rates, with the goal of informing a hospital's choice to incorporate it into their standard of care.

Conclusion

Healthcare organizations are transitioning to EHRs to improve collaboration and information sharing, reduce inefficiency and administrative effort, and ultimately improve patient care. Data science is a deep, flexible, customizable next level of opportunity for organizations using EHRs. Healthcare decision-makers can choose from among the depth, breadth, and complexity of the data sharing, analytics, and explanation tools they deploy based on their missions, budgets, and patient needs.

As healthcare organizations grow, adapt, and serve changing populations, the data science mix of methods can be tailored to keep delivering benefits. At every point, data science can be harnessed to derive useful, actionable information from the complexity of data in EHRs.

leidos
HEALTH

USE CASE 1:

Public Health Influenza Data Sharing

Homeland Security Presidential Directive 21[1] states the necessity of establishing a national biosurveillance system for human health that integrates state, regional, and community-level capabilities. To better support early detection and rapid response to major infectious disease outbreaks or pandemics, the Centers for Disease Control (CDC) wanted to expand the real-time exchange of clinical health information among health information exchanges (HIEs) and local, state, and federal public health organizations.

CDC awarded Leidos a contract to design and develop an integrated HIE "to connect public health with HIEs to improve public health's real-time understanding of communities' population health and healthcare facility status."[2]

Leidos led a team of bioinformaticists, systems engineers, HIE operators, and subject matter experts in state and local public health requirements to assemble the Northwest Public Health Information Exchange (NW-PHIE) to help achieve this goal.

A key challenge was establishing a real-time, multidirectional information flow across federal, state, and local jurisdictions. Until this project, biosurveillance platforms such as RODS[3], ESSENCE[4], and BioSense[5] collected de-identified biosurveillance information, but did not share or visualize these data across platforms. No syndromic surveillance data feeds were being sent to public health from hospitals in central and eastern Washington State.

The starting points for NW-PHIE's syndromic surveillance requirements were the American Health Informatics Community (AHIC) Biosurveillance Use Case and the MBDS. NW-PHIE augmented the AHIC requirements with additional requirements from local and state public health agencies.

The Leidos team developed a data format specification that described how the highly diverse clinical data would be transmitted to public health. We followed best practices and industry standards, such as those from HL7.

Leidos next mapped the data fields in the HL7 messages to the clinical care information provided by the various data sources to meet the data output specifications. Through these alignments, information for a single patient is split into many discrete biosurveillance HL7 messages triggered by events at a healthcare facility, such as registration, lab order, lab result, etc.

To make clinical data useful for population health purposes, several analyses were needed. First, individual messages are reassembled into a longitudinal, comprehensive view of a visit encounter. Next, the encounters are classified according to surveillance criteria, counted over a fixed time interval, and paired with an appropriate denominator, such as total visit volume or catchment population. From these data, absolute counts and rates of illness within the patient population are obtained.

1. http://fas.org/programs/bio/resource/documents/hspd-21.pdf
2. M. Trebatoski, J. Davies, D. Revere, D. Dobbs, "Methods for Leveraging a Health Information Exchange for Public Health: Lessons Learned from the NW-PHIE Experience," Journal of Public Health Informatics, ISSN 1947-2579 * http://ojphi.org * Vol.2, No. 2, 2010.
3. Real-time Outbreak and Disease Surveillance Laboratories (RODS), https://www.rods.pitt.edu/site/
4. The ESSENCE II Disease Surveillance Test Bed for National Capital Area, Joseph Lombardo, http://techdigest.jhuapl.edu/TD/td2404/Lombardo.pdf.
5. http://www.cdc.gov/BioSense

Explanation

To assist public health departments in evaluating the clinical data, we developed tools and algorithms to help visualize data and support the development of appropriate public health responses. The resulting analysis is fed back to clinicians via health alerts, investigation findings, and case criteria updates.

Now, thanks to NW-PHIE, 16 hospitals (representing close to a million patient encounters a year) are sending a full set of syndromic surveillance data to public health. These feeds proved to be vitally important in helping manage the 2009 H1N1 outbreak and are available to help manage future disease outbreaks.

leidos HEALTH

USE CASE 2:

Analysis of Congestive Heart Failure Readmission Rates

Patients with a diagnosis of congestive heart failure (CHF) pose a challenge to health delivery systems. According to the Centers for Medicaid & Medicare Services (CMS), the rate of readmission within 30 days of discharge for CHF patients ranges from 18.2% to 33.8%[6], compared to the average readmission rate of approximately 24.8% for all patients in the United States.

CMS views these variations as an opportunity to decrease the rate of CHF readmission. Beginning in October 2012, CMS has penalized payments to health systems that exceed the average rates of readmission. Other diagnosis categories, such as hypertension, also trigger penalties from CMS when readmission rates exceed prescribed norms.

To address this business challenge, Leidos conducted a research and development effort to study whether our information-based health approach to clinical challenges could be used to create a model that could calculate an accurate CHF readmission probability based on a broad set of clinical and non-clinical factors.

A multidisciplinary team including experts in commercial and federal healthcare, health informaticists, physicians, scientists, statisticians, and mathematicians conducted initial research into causes of CHF readmission. They identified three types of factors that may affect readmission of CHF patients: patient-specific factors (clinical and demographic), hospital-specific factors (practices and continuity), and socioeconomic factors (support and access).

Our team used these SME findings to develop our pilot prediction model for CHF readmission rates, in the form of a Bayesian inference network (INET) model with multiple nodes and relationships (links). Each node depicted on a network graph represented a grouping of data that affects CHF readmission (directly or indirectly), such as whether a patient has extensive comorbid conditions or a social support system after discharge. The model, which includes multi-path relationships (chains of inference), then generates predictive values by statistically weighting and combining the complementary or competing chains of inference.

Note that this use case focuses on analysis; requirements for data collection and curation would need to be considered based on a specific implementation of this proof-of-concept experiment.

However, in preparation for a decision on technology investment into data collection and curation, the team also investigated whether electronic patient data were available to populate the model's parameters directly. Figure 1 illustrates the availability of EHR data in our statistical analysis of more than 5,000 CHF patient encounters in a 10-hospital system between 2010 and 2011.

6. CMS Hospital Compare Database, January 26, 2012, http://www.hospitalcompare.hhs.gov/hospital-search.aspx

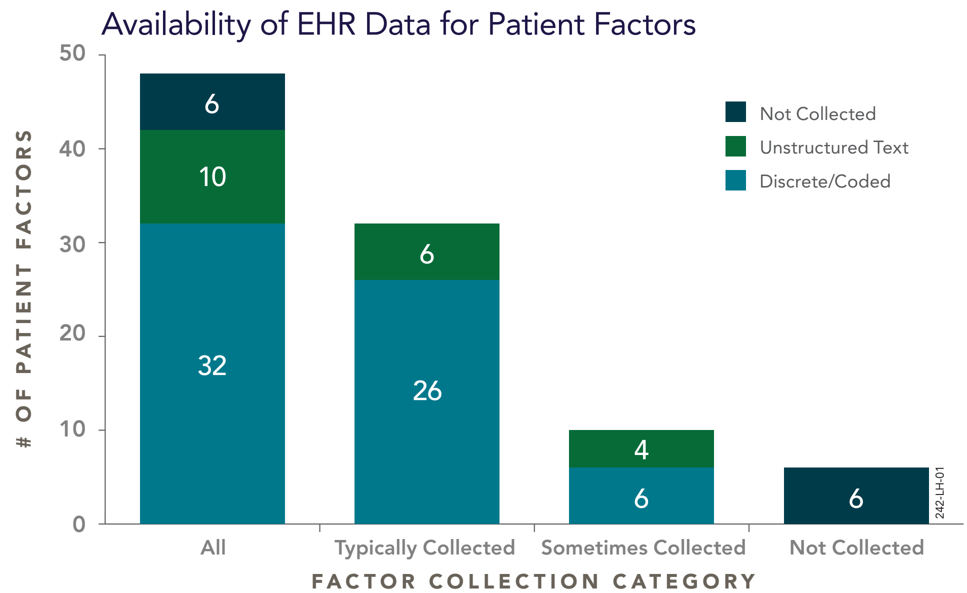

Figure 1. Availability of EHR Data for Patient Factors

Explanation

Having constructed the network and populated it with data, the team investigated a set of outcomes for three use cases identified by the team's clinical SMEs, using historical "hard" CHF data and subjective, but expert, judgments.

The INET model's findings — including outcome likelihood (chance the patient would readmit) and "top N" indicators that, if addressed correctly, would reduce the readmission — were presented to a focus group consisting of physicians, executive leaders, informaticists, and statistical analysts from a 10-hospital university health system and a large, inner-city safety net hospital.

The focus group members reviewed the outcomes against their own CHF readmission experience and concluded that the INET model would be a useful decision-support tool that provides insight for hospital clinical and administrative staff members responsible for identifying CHF patients at risk prior to discharge.

leidos HEALTH

ABOUT LEIDOS HEALTH

Leidos Health helps healthcare organizations achieve their goals of meeting regulatory requirements, improving quality of care, reducing costs and enhancing the patient experience. Our services include implementation and optimization of EHRs for all major vendors, as well as solutions for critical initiatives such as IT strategy, revenue cycle, clinical optimization, Meaningful Use, and ICD-10, technology infrastructure, and cybersecurity. Leidos Health (a subsidiary of Leidos, formerly known as SAIC, and including the businesses formerly known as maxIT Healthcare and Vitalize Consulting Solutions) is a new company with unique capabilities and a 25-year legacy of success.

For more information, visit **leidoshealth.com**.

FOR MORE INFORMATION

877.652.4099 / contact.us@leidoshealth.com
705 E. Main Street / Westfield, IN 46074
Visit us at **leidoshealth.com**

leidos
HEALTH

242-LH:VF3