BIG Data in Healthcare: Hype and Reality

By Naeem Hashmi, February 17, 2013

The Big Data bubble in Healthcare is filled with hot air - at least for now. Every one is talking about it but when you dig deep with pointed questions, very quickly you discover that it has nothing much to do with the Big Data. So why am I taking time to write this note? Well, we need to look at Big Data in context of Healthcare realities. First, we need to understand the spectrum of healthcare segments and look at each segment on feasibility of Big Data components.

I identify the following six distinct healthcare segments in context of Big Data as:

- Life Sciences
- Pharma
- Medical/Mobile Devices
- Care Provider
- Payers
- Public Health

I did not include Patients in the list as they are mostly content generators, consumers and have much to benefit from but they are on the mercy of other segments. Life Sciences and Pharma segments are already exploiting Big Data elements for OMICs research in discovering new drugs and therapies that I will discuss in later articles.

In this article, I will focus one piece of the puzzle, the realities of Big Data in the Care Providers segments.

To frame the discussion, let me describe what I mean by Big Data. Big Data consists of the following elements:

- **Content Attributes:** 4 Vs (Volume, Variety, Velocity, Veracity) - widely accepted characteristics
- **Technology Stack:** Pure Hadoop-stack and Analytics Appliances
- **Computing Model:** Extremely Reliable Parallelism

- **Query Model:** Hadoop/Map-Reduce uses search method (not complex query). Not SQL.
- **Connecting the Dots:** Understanding the associative meanings of interesting patterns hidden underneath massive content-stores.

The Healthcare Provider segment is extremely fragmented. With ongoing Healthcare reforms, there is some hope that use of EHRs may result in greater use of clinical content. Most of the CIOs are too busy in addressing ICD-10 and CMS Meaningful Use 2 mandates and none have really invested in true Big Data platform. Data management still remains on legacy platforms, often mainframes. So how /where does Big Data fit in the Care Providers segment and challenges?

Data Jail

Access to Clinical data is not easy. EHR vendors have a lock on patient clinical content within their proprietary 'Data Jails'. It is extremely difficult to access clinical content and analyze. The Volume and Velocity needed for Big Data is not accessible to providers.

I am using the term 'Data Jail' because it takes me back to the early 90's, when the 'Data Jail' term was widely used for the ERP vendors. Then, IT folks had a hard time pulling data out of ERPs for reporting and analytics. Today, ERP vendors are quite open and allow access to the data from their data-stores. The EHR vendors need to learn lessons from ERPs for openness, interoperability and scalability. Opening access to their repositories is good for their customers and to leverage Big Data elements.

Most clinical data volumes are low, static and locked in the EHR 'Data Jails' meaning that data acquisition velocity is very low. On the other hand, clinical content is extremely complex and unstructured. Content context and semantics are more important than volume and velocity, and you may find that you can live with standard relational databases or analytics appliances without Big Data elements.

Real-Time Patient Monitoring

Real-time patient monitoring applications, such as bedside heart monitors, OR monitors for anesthesia, Dialysis patient monitors, etc., are being piloted in care facilities. These applications require a Big Data platform that can quickly consume real-time events and signal possible life threatening or other adverse events.

Moreover, the increasing use of mobile personal health monitoring devices that stream data to the care providers also require Big Data elements to capture and predict possible health issues and alert the care takers and patient for necessary actions. Real-time patient monitoring applications are good candidates for Big Data driven analytics.

ACO Level Analytics Hub

ACO level data volumes are expected to be quite high and Big Data will play a key role for developing analytics. Again, the challenge will be how to pull detailed clinical content from EHRs and harmonize the content in an ACO hub with billing and operational data, ensuring content quality.

Lack of data exchange standards are a problem. Even when vendors conform to a patient data exchange standard, such as Continuity of Care Document (CCD), the actual implementation varies significantly. For example, CCD implementation between two hypothetical vendors can results in completely different patient clinical content. The reason - one vendor packages a latitudinal CCD while the other encounter-specific CCD. Both conform to the CCD specifications but the payload is different. For the ACO Big Data repository, you need a high performance harmonization engine to ensure content consistency before you are able to embark on meaningful Analytics.

Valuing Big Data

After going through the big pain of a Big Data implementation, the end goal is to bring valuable insight back in the clinical processes and embed within clinical workflows. Most people think that simply displaying charts and graphs on a stand-alone dashboard will do the job.

In my experience, however, clinicians will not leave their workflow to access a stand-alone dashboard, no matter how insightful. Physicians and clinicians do not have time to hop between many screens while caring for patients. Instead, we need to embed analytical insights within clinical work-flows so the provider has the right alert at the right time for the right patient. , The limitations of EHR interoperability make embedding non-EHR events within clinical workflows and work-lists difficult, and sometimes impossible.

Emerging Big Data Opportunity - Comparative Effectiveness Research

Humans are social animals. We socialize and chat about our daily life stories, including health issues, families, care givers, coping behaviors, and medications. Such social communication is a gold mine to Pharma and Clinical Outcome researchers to understand the effectiveness of specific therapy. This new field is called Comparative Effectiveness Research (CER).

The CER is similar to Randomized Control Test but that CER provides a lot more insight on clinical outcome when certain therapies are put into real practice. Through use of the Big Data

analytics, Comparative Effectiveness Research will significantly improve patient care through personalized therapies.

Emerging Providers of Big Data Clinical Analytics

Last year, I served as a judge for Massachusetts' first Health Datapalooza competition. Several Big Data driven clinical applications were part of this competition. I had good discussions with a few vendors. Though their solutions are unique and valuable, none have thought through how their solution will fit into the clinical setting, client EHR workflows, or leverage the client's clinical platform. These Big Data analytics providers are mostly start-ups with innovative care solutions.

When I scan the Big Data marketplace, it reminds me of 'dot com' era of the early 90's. Back then, many small innovative groups were developing fascinating tools, just like today in the Healthcare world. Unfortunately, few will survive. My suggestion to Big Data innovators is to remember that if they want to succeed in the clinical marketplace, their solution must be interoperable with EHRs and ACO business/clinical applications.

Closing thoughts

Use of Big Data in the clinical practice is in the early stage of its evolution. Most CIOs may talk about Big Data, but to best of my knowledge, there are no pure Big Data implementations in hospital or clinical practices. Only a handful of large commercial Clinics networks have gone beyond the pilot stage and have fully implemented Big Data platforms.

Admittedly, building a Big Data platform and maintaining a team of skilled data scientists is not easy. It is also very expensive. Currently, Big Data pilots in healthcare settings are only found in University Hospitals because they are able to tap into academic research funding and a pool of [cheap] graduate students who are ready to investigate unexplored content. Similarly, companies developing healthcare analytics that use social data for public health and patient engagement applications are mostly small start-ups. Today, we are just taking baby steps, and we have a long road ahead to attain Big Data's true value for clinical settings.

Author
Naeem Hashmi
Information Frameworks
http://infoframeworks.com
nhashmi@infoframeworks.com