



# Mine Big Data To Advance Clinical Decision Support

Only through innovative analytical techniques will we be able to truly leverage the healthcare data collected and improve the way we deliver care.

By Barry Chaiken, MD, FHIMSS; CMO, DocsNetwork

Although much effort focuses on the improvement of clinical workflows, an opportunity exists to transform healthcare delivery by implementing evidence-based clinical decision support at the point of care. Clinical content delivered effectively within new, efficient clinical workflows can aid in directing patients toward evidence-based therapeutic plans that produce desired clinical and financial outcomes. While informaticists work on developing these clinical workflows, the lack of clinical knowledge limits the ability of organizations to leverage HIT in order to personalize therapeutic care plans.



the author

Barry Chaiken  
MD, FHIMSS; CMO of  
DocsNetwork,  
[www.docsnetwork.com](http://www.docsnetwork.com)

The digital age is the age of big data, where every piece of technology captures data that can be available for later use. The McKinsey Global Institute (MGI) describes data generated in this way as digital “exhaust data,” or data created as a by-product of other activities. Wikipedia defines

big data as: *a term applied to data sets whose size is beyond the ability of commonly used software tools to capture, manage, and process the data within a tolerable elapsed time. Big data sizes are a constantly moving target currently ranging from a few dozen terabytes to many petabytes of data in a single data set.*

## BIG DATA MEANS BIG VALUE IN HEALTHCARE

The rapid expansion in the use of EMRs and digitally driven technology — MRI scanners, body sensors, automated lab tests — brings the era of big data to healthcare. MGI estimates that big data presents a \$300 billion potential annual value to the U.S. healthcare system. The five broad areas to deliver that value are:

- 1) clinical operations,
- 2) payment/pricing,
- 3) R&D,
- 4) new business models, and
- 5) public health.

Sub-areas include comparative effectiveness research (CER), clinical decision support, remote patient monitoring, health economics, and personalized medicine.

The four large data sources for healthcare include clinical (e.g. EMR, images), pharmaceutical (e.g. clinical trials), administrative (e.g. utilization, claims), and consumer (e.g. home monitoring, retail purchases). Each of these data sources has its own pros and cons.

The large investment in EMRs and the increased use of digitally connected medical devices drive the rapid expansion of available clinical data. As these technologies evolve, the data collected becomes more expansive and granular, yet poorly utilized.

Pharmaceutical and medical device companies collect clinical trial data to substantiate the safety and efficacy of their products. Although the clinical trial group represents a subset of the real target population, limited analytics aided by expert opinion provide the basis for a suboptimal product review process.

Both payors and providers utilize administrative data to monitor their business practices. Lacking in clinical meaning, analysis often leads to poor decisions based upon erroneous results gleaned from analysis of incomplete data sources.

As the availability of information from consumers grows with their use of technology, retail entities utilize the data to assist in the management of their businesses. The expanded deployment of remote patient monitoring devices and the collection of data points through social media and consumer monitoring programs (e.g. pharmacy purchases with an affinity card) now offer additional data sets that were unavailable only a short time ago.

These data sources present a valuable area for analysis by researchers striving to find ways to improve care delivery while lowering costs. CER, supported by data mining, allows organizations to identify affordable therapies that enhance patient care. With the implementation of HIT, data



## Trend #5: Clinical Decision Support

warehouses contain petabytes of searchable clinical, outcome, genomic, and financial data across multiple patient populations. Bringing this data together using sophisticated knowledge, analytic tools, and domain-specific interfaces allows researchers to discover relationships among multiple variables

Organizations that properly collect, analyze, and use big data will achieve a significant competitive advantage over those organizations that fail to recognize the opportunity big data presents.

gleaned from previously unconnected databases.

In turn, this new clinical knowledge enables clinicians to personalize treatment for patients

based upon their genetic background by linking it to descriptive patient data and outcomes. Personalized medicine transcends analysis of a population-based cohort by placing the patient within a sub-population that better reflects the expected outcome from a prescribed treatment. Embedding this personalized knowledge within an EMR's clinical decision support module facilitates the delivery of these evidence-based best practices at the point of care.

### SEMANTIC WEB LINKS DISPARATE HEALTHCARE DATA

New analytic tools such as Semantic Web 3.0 — linked data — offer ways for machines to analyze these data sets, leveraging approaches that would be impossible using standard relational databases and statistical methodologies. These new tools permit researchers to work around the barriers presented by data sets' nonconformance to standards for data collection or storage.

Similar to the use of metadata, semantic web techniques allow the assignment of descriptors to each data point, providing a context and meaning to the data. This allows machines applying powerful statistical techniques to analyze the disparate data sets in ways not available to humans alone due to the data sets' size and complexity.

Effective use of semantic web technology in medical research requires the indexing of the available clinical and nonclinical data sets. These data sets

include clinical data taken from EMRs; patient genomic data; existing genomic pharmaceutical databases; curated disease-specific, peer-reviewed research; and nonclinical data taken from claims systems, consumer behavior records, and monitoring of social media. Experts can utilize semantic web technology to query multiple large data sets to explore comparative effectiveness hypotheses.

For the entire history of medical research, investigators posed hypotheses and tested them to see what therapies proved effective. Advances in clinical knowledge grew from frequent comparison of different therapies, with clinicians shifting to those that offered the best results. Comparative effectiveness analysis forms the basis of all medical research. The availability of semantic web technology and newly constructed clinical data sets presents researchers with an extraordinary opportunity to rapidly explore clinical relationships within subpopulations of patients using data formerly unavailable. These results then form the basis for evidence-based protocols specifically targeted at those subpopulations

### BIG DATA AS HEALTHCARE'S COMPETITIVE ADVANTAGE

The knowledge obtained from big data offers additional benefits to healthcare. CER delivers medical knowledge that can be applied using clinical decision-support tools deployed at the point of care. The analysis of subpopulations allows for the delivery of personalized medicine that accounts for genetic variation between and among ethnic groups.

Big data applied to health economics and outcomes research facilitates the development of performance-based pricing plans that reward quality outcomes rather than incentivizing utilization. Accountable Care Organizations will derive great value from using big data.

The uses of big data are numerous and far-reaching. Only through innovative analytical techniques will we be able to truly leverage the healthcare data collected and improve the way we deliver care. Organizations that properly collect, analyze, and utilize big data will achieve a significant competitive advantage over those organizations that fail to recognize the opportunity big data presents. □