Data in Motion – Big Data Analytics in Healthcare

A White Paper
# Table of Contents

1. Introduction
2. Big Data in Healthcare
3. The IT Challenge
   4. Stakeholders' Dream and Reality
4. Big Data Analytics Case Study
5. Why Information Builders and EMC
   6. Integration Solutions
   7. Intelligence Solutions
   8. Integrity Solutions
6. EMC Solutions for Healthcare Analytics
7. Conclusion
Introduction

There are a number of questions – and lots of marketing hype – surrounding big data. Some healthcare providers argue that the use of big data is the key to achieving sustainable competitive advantage and optimizing clinical outcomes. Others wonder if it is a buzzword with more risk than reward. All will agree that new approaches to analytics are needed to fully harness the available data points across the patient journey and transform it into actionable insight that will truly enhance care quality.

Some providers believe they already have solutions in place, but those solutions often consist of just an Excel spreadsheet and a database. For healthcare providers and payers alike, more powerful and robust tools are needed to successfully leverage big data and analytics to improve patient care delivery, accelerate accountable care business models, reduce costs, and drive organizational change throughout the value chain.

Data – especially when generated in the healthcare setting – is never at rest. As patients see physicians and receive diagnostic tests, information about those visits and procedures flows between systems, providers, insurance companies, and healthcare networks. Other data may also be required, including details about enrollments, provider credentialing, or fee schedules, as well as information collected from medical collaborators and other external sources, such as electronic health records, medical images, and care management documentation. This data in motion drives the need for the continual use of analytics to improve services, reduce costs, and enable informed patient care and decision-making.

This paper explores the real-world challenges healthcare providers face when considering big data and analytics. We will present a successful case study, as well as a comprehensive strategy for implementing big data analytics within an organization.
What is big data, how does it impact today’s healthcare firms, and what can they do to most effectively leverage it?

First, big data can be characterized by three “V”s:
- **Velocity**: the speed with which data is generated
- **Variety**: the range of elements contributing to data generation
- **Volume**: the amount of data collected from each source

Some people cite a fourth “V”: veracity, or the ability to qualify and quantify the data collected from each source, and leverage it for analytics.

Big data needs a place to reside. It needs powerful hardware, backed by solid software, so the variety and volume of information, which often comes from a vast array of sources (personal medical records, radiology images, clinical trial data, FDA submissions, human genetics and population information, etc.), can be efficiently processed, and analytical workload complexity and agility can be supported. A high percentage of big data is often described as multi-structured to distinguish it from the structured operational data used to populate a data warehouse. In most organizations, multi-structured data is growing at a considerably faster rate than structured data.

The challenge lies in unifying these disconnected data sets and fully leveraging the potential of existing information. Technology is available to enable this integration, as well the generation, capture, and analysis of the new healthcare data streams, such as genomics, sensor readings, and population and disease statistics.

Just as the volume and variety of data has changed, so too has the velocity at which it is generated and the speed with which it must be retrieved, analyzed, and compared to facilitate better decision-making. Advances in technology – particularly storage and analytics platforms – provide the ability to create and leverage a data warehouse to cost-efficiently bring all that information together and make it readily available for reporting and analysis.

A common mistake is to presume that all data within a warehouse is clean. Data quality issues are of particular concern in healthcare because they can impact informed care-related decision-making. The quality of healthcare data, especially unstructured data, is highly variable as it comes from a multitude of sources – written prescriptions, doctor’s notes, inaccurate charting, etc.

Veracity in healthcare data is imperative. Are the patient, hospital, payer, or reimbursement codes correct? Are diagnoses, treatments, prescriptions, procedures, or outcomes properly captured? Variety, volume, and velocity all impact veracity by fueling the cycle for continuous improvement and analytics. The highest value in healthcare data will be realized with enhanced data quality and representation of relevant measures.
The collection, standardization, integration, and consumption of this rapidly growing structured and unstructured data – with regard to all four Vs – defines big data analytics. A software solution that integrates, normalizes, and validates clinical data from across the continuum of care can not only improve outcomes and reduce costs, but also help identify disease trends, coordinate rules-driven patient registries, and drive performance management strategies – a key ingredient for meeting Accountable Care Organization (ACO) requirements.
Gaining an enterprise-wide view of healthcare data helps organizations provide fact-based answers to questions ranging from which therapeutic approaches work best and which patients are at highest risk for readmission or infection, to how physicians are performing in relation to quality and cost. The challenge is getting that data into a format that allows clinicians to make decisions in a faster and more coordinated fashion.

While traditional analytics focuses on the use of data warehouses, big data calls for immediate synchronization. Big data analytics is more iterative and predictive, anticipating what may happen in the future so providers can take a more proactive approach to patient care.

However, eliminating data silos and bringing together fragmented information in real time doesn’t have to be so difficult. Imagine the value of identifying approaches to use leading indicators, instead of trailing ones, so that interventions could be made more rapidly – and with more predictive power behind them.

The move to big data analytics for healthcare is an opportunity to explore and seek competitive advantage through greater insight. Like any investment, it comes with risks and considerations. A solid business case must be made. Whether or not more data will likely yield valuable insights must be determined. In the end, it’s about using data more effectively.

The best way to approach big data analytics, and support the extreme processing involved, is to deploy optimized hardware and software solutions for processing different types of big data workloads. These solutions can then be combined with the existing enterprise data warehouse to create an integrated information supply chain.

Big data analytics platforms – like the EMC Greenplum™ Unified Analytics Platform (UAP) – are a major advance in solving these challenges. Solutions like these offer an integrated system built from the ground up to enable comprehensive installation and management of the environment, while boosting price and performance for analytical workloads.

Once an organization has a sense of its data, it must determine what the most critical business and clinical decisions are and whether or not the required data and analytics are available to make them. How many of these decisions would be further supported with better analytics? Is the organization structured in a way to capitalize on better decision-making – especially with more data being created every day?

With powerful big data analytics tools, healthcare organizations have the storage and computation needed to ingest and digest information from across their data-rich enterprises. Through techniques such as machine learning and statistical modeling, organizations can learn from the details in a more comprehensive manner – capturing and analyzing all the data. Whether the focus is on patient care delivery or clinical research for disease management, these systems will help successfully drive any analytics strategy forward.
For example, researchers can leverage the same scalable and well-understood techniques used in weather forecasting models, applying them to huge epidemiology data sets to more accurately predict viral outbreaks. Hospitals can evaluate the propensity for fraud or bed churn, and readily achieve critical performance improvement goals by overcoming information integrity issues, enabling compliance with government regulations, and facilitating the flow of accurate information across the organization. Doctors can have more accurate pictures of the genetic and environmental causes of disease, and discover more effective treatments as a result.

**Benefits to Stakeholders**

Every stakeholder in the healthcare value chain can benefit from big data analytics stemming from accurate data:

**Providers** can be granted real-time, single-view access to patient, clinical, and other relevant data to support improved decision-making and facilitate effective, efficient, and error-free care. They can also ensure accurate, on-time payment that promptly reimburses them for their time and care.

**Hospitals** can use big data analytics as a means of gaining more control over quality initiatives, with an eye toward reducing treatment costs and cutting readmissions. They can chart clinical performance against national standards and benchmark against other hospitals; examine the effects of length of stay, complications, and readmissions; or expand from looking at individual patients to looking at populations of patients with problems, such as congestive heart failure or diabetes. They can also establish detailed clinical histories for patients to ensure that diagnoses and treatments are based on the best available patient information.

**Patients** will benefit from more timely coordination of care and the flow of error-free information between themselves, caregivers and providers, and payers.

**Researchers** can leverage predictive modeling, statistical tools, and algorithms to improve the design and outcome of experiments, and to provide a better understanding of how to develop treatments that address unmet needs, while navigating regulatory and marketing processes.

**Pharmaceutical companies** can better understand the causes of diseases, find more targeted drug candidates, and design more successful clinical trials to avoid late failures and market safer and more effective pharmaceuticals. Once in the market, they can leverage accurate formulary and reimbursement information to customize their marketing efforts and conduct less costly post-marketing surveillance.

**Medical device companies**, many of which have been collecting data from hospital and home devices for safety monitoring and adverse event prediction, can leverage home monitoring devices, tablets, and the Internet to improve care and partner with hospitals to minimize patient stay length.
Payers can move from fee-for-service to pay-for-performance, and can use big data to help stratify population risk and guide them to more sustainable business models such as expanding into wellness management, data analytics, and API platforms.

Government agencies can reduce costs, enforce regulations, and maximize the social value of data.
A community-based, not-for-profit healthcare provider based in Illinois has implemented software and hardware from Information Builders and EMC to move its analytics strategy forward. Together, these solutions provide a powerful, comprehensive environment for data integration, master data management (MDM), business analytics, and performance management. As a result, it can readily achieve critical performance improvement goals by overcoming information integrity issues, enabling compliance with government regulations, and facilitating the flow of accurate information across the organization.

Working in conjunction with EMC, Information Builders helped this healthcare provider create an enterprise master patient index (EMPI), a comprehensive enterprise data warehouse that provides a single view of the patient, complete with full auditing capabilities. The warehouse contains insight about the various activities that take place as each patient moves through the care continuum, which is collected from multiple diverse sources across the business, including Cerner, Lawson, and Allscripts. This environment is extended with Information Builders’ WebFOCUS business intelligence (BI) and analytics platform, which provides advanced reporting and analysis features that allow clinicians, administrators, and other stakeholders to quickly and easily generate compliance reports that monitor critical metrics for performance management purposes.

The provider will take advantage of data integration components, data quality software, and capabilities leveraging the EMC® Greenplum® Data Computing Appliance (DCA) and UAP to help gain an enterprise-wide view of its patient information. This is particularly important because, for many healthcare providers, the disparate nature of data across labs, clinical departments, hospital systems, and even clinical applications, such as electronic health records, can hinder efficiency. To improve quality of care and lower costs, they must use best practices for integrating assets.
The promise of big data in healthcare is revolutionary. Use of big data will ease the transition to authentic data-driven healthcare, allowing healthcare professionals to improve the standard of care based on millions of cases, define needs for subpopulations, and identify and intervene for population groups at risk for poor outcomes. To date, few healthcare professionals would claim that the promise of big data has been fulfilled.

Big data and analytics can create immediate value for an organization. Analytics can improve an organization’s performance by making everything more transparent and measurable, while exposing variability as well as potential issues and opportunities. This transparency, through operational dashboards or performance management, fosters better decision-making across an organization and enhances patient care delivery at a lower cost and improved efficiency, while also satisfying benchmark and regulatory issues.

A successful business and clinical analytics strategy requires a singular mechanism for gathering, consolidating, preparing, and storing data from any source, in any location. Only with the right comprehensive platform – one that can enable integration of all information assets, dramatically improve the integrity of the data contained in those assets, and transform that data into powerful intelligence for wide-scale use – can companies ensure that their enterprise information is truly analytics-ready.

Information Builders offers a wide array of powerful, proven intelligence, integration, and integrity solutions that allow organizations to rapidly and economically build data warehouses that are optimized for reporting and business analytics. These solutions reduce costs and accelerate return on investment by tapping into the power of real-time information, and making it readily available to any user, anywhere, at any time.

Integration Solutions
iWay data integration solutions from Information Builders allow for direct access to all of your data, so you can design your architecture to address the unique information needs of all your users. iWay provides:

■ End-to-end integration of a wide variety of sources, more than 300 in all, including cloud-based information, social systems, and big data
■ Support for real-time and batch integration
■ Flexible extract, transform, and load (ETL) and message-based styles of integration

Intelligence Solutions
The WebFOCUS BI platform empowers everyone – from internal employees to customers and partners – to make smarter, more confident decisions by leveraging powerful, yet intuitive analytic tools from any browser or mobile device.
Business Intelligence
From financial and HR reporting to compliance reporting and dashboards, WebFOCUS makes vital information valuable to everyone, at every level of the enterprise, on any Internet-connected device. With the growing importance of digital channels, customer-facing analytics solutions can help to enrich relationships with patients and partners, too.

Business Analytics
Cutting-edge business analytics, including predictive analytics, visualization, location intelligence, and enterprise search, allow even non-technical users to manipulate data in an almost unlimited number of ways to uncover critical patterns and trends that would be impossible to identify otherwise.

Performance Management
An out-of-the-box performance management framework empowers organizations to monitor, track, and drive performance by communicating goals, measuring execution, and assessing enterprise risk against performance strategies and objectives.

Integrity Solutions
iWay provides comprehensive, real-time management of any information from any source. Regardless of where data resides, whether it’s in a structured or unstructured format, iWay can seamlessly cleanse and enrich it, ensuring a consistent and accurate view of data from every interaction point. This is particularly important in analytics-driven environments, where the accuracy and integrity of the information used as the basis for mission-critical decision-making must be ensured.

Data Quality Management
Real-time data quality management and profiling capabilities help to optimize the consistency, completeness, and correctness of data across all information assets.

Master Data Management
Leading-edge MDM technology consolidates millions of records, and makes unified and validated master data instantly available to a wide range of enterprise applications and systems.

Omni-Patient™
Omni-Patient delivers a unique combination of an enterprise master patient index (EMPI) with pre-packaged models designed specifically for the healthcare industry. It empowers healthcare firms to overcome the challenges associated with fragmented information architectures by facilitating the validation of patient identification across enterprise systems and enabling the creation of a single golden record for each patient, provider, physician, payer, workforce, and facility.

Data Governance
Data governance tools allow data stewards and other stakeholders to manage and control information at every stage and in every system.
EMC Solutions for Healthcare Analytics

The EMC Healthcare Analytics Appliance is a complete, purpose-built solution that combines the industry-leading speed and scalability of the EMC Greenplum analytics platform with a robust ecosystem of application partners. It provides a complete solution in one device, with the flexibility to integrate with existing tools. Whether an organization is just beginning to strengthen its BI capabilities, or is implementing accountable care business models that require broad analytics, EMC can help build the supporting IT infrastructure, on-premise or hosted as an ITaaS, with the highest levels of performance, availability, virtualization, security, and integration.

The solution leverages EMC Greenplum’s UAP, which combines co-processing of structured and unstructured data with a productivity engine and social network that enables collaboration among data scientists. The UAP fuses three Greenplum products into a unified offering to support healthcare provider analytics use cases: the EMC Greenplum database for structured data; the enterprise Hadoop offering EMC Greenplum HD for analysis and processing of unstructured data; and EMC Greenplum Chorus™, which acts as a productivity engine and social network for data science teams.

EMC Greenplum’s DCA offers the power of a massively parallel processing (MPP) architecture, while delivering the fastest data-loading – without the complexity and constraints of proprietary hardware. It is a unified big data analytics appliance, a modular solution for structured data, unstructured data, and Greenplum partner applications such as BI and ETL.
Conclusion

As information environments become larger and more complex, healthcare firms must find better, smarter ways to harness and tap into their big data, in order to optimize clinical, financial, and administrative operations. To remain competitive and compliant, they need a comprehensive platform for collecting, storing, and enhancing big data – and making it readily available for use by all stakeholders in the healthcare value chain.

Big data analytics platforms, such as the EMC Unified Analytics Platform, are designed to help healthcare firms overcome the obstacles associated with managing and maintaining big data. When combined with BI, integration, and integrity solutions from Information Builders, healthcare organizations are empowered to capture and analyze all their data – no matter where it resides – whether it’s in structured or unstructured format, and use it to improve operations, contain costs, meet stringent regulatory demands, and most importantly, maximize the quality of patient care.