

Curating Knowledge in the Digital Age:

Mining Massive Data Sets for Healthcare Insights





In the recent article *Healthcare's Age of Liquid Data*, we explored why the healthcare industry must move aggressively to make healthcare data extremely connected. We identified three steps to make this happen — the industry must aggregate data, curate data and use the data to engage consumers in a personalized manner. Each step is examined in this series of follow-up articles.

Healthcare is on the brink of transformation. Payers, health systems and other providers that embrace robust data and analytics and extreme interoperability will dramatically enhance their capacity to engage patients, support clinicians, enhance efficiencies, improve revenues, expand services and individualize the prediction, prevention, diagnosis and treatment of illness and disease. The technology for managing and utilizing healthcare data is already widely available and improving quickly. Now, new rulings by the Center for Medicare and Medicaid Services (CMS) and the Office of the National Coordinator for Health Information Technology (ONC) have expanded patient data access to providers, payers, third-party vendors and patients themselves. This will turbocharge data sharing, care coordination and consumerism while leading to an explosion of new products and services.

To compete effectively in this age of liquid data, traditional healthcare organizations will need to step up their technological capabilities and adopt a new mindset. This article focuses on the importance of curating massive data sets to generate actionable data insights.

Data Curation Process



STAGE 1 Grouping Data into Meaningful Units of Analysis



STAGE 2 Determining Current/ Future Care Utilization and Financial Risk



STAGE 3 Targeting Opportunities to Improve Care Quality and Organizational Performance

DATA AND VALUE

Value-based care is a data-driven enterprise. As healthcare organizations expand their payment and care delivery models to embrace value through population health strategies, they increasingly rely on data to optimize organizational performance, engage members/patients and enhance care quality and outcomes.

That starts with aggregating and standardizing massive data sets from multiple sources. Many healthcare organizations, however, lack the data curation capabilities to derive powerful and actionable insights from that data. By framing data curation as a three-stage process, they can blueprint the workflow, capabilities and technologies necessary to mine data insights consistently and effectively.

The three-stage process consists of the following components.

Stage 1: Grouping Data into Meaningful Units of Analysis

Stage 2: Determining Current/Future Care Utilization and Financial Risk

Stage 3: Targeting Opportunities to Improve Care Quality and Organizational Performance

Let's explore each of these stages of the curation process.







Stage 1: Grouping Data into Meaningful Units of Analysis

Health systems can collect multi-terabytes of clinical, claims, environmental and social data on individual consumers. These massive, complex and highly detailed data sets are not useful until grouped into meaningful units of analysis.

Consider a patient with diabetes, Jane Smith, on a typical doctor visit. In her consultation, Jane will likely undergo diagnostic testing and receive a diagnosis and care plan that may require therapy and prescription drugs.

These activities generate data from physicians, labs, pharmacists and therapists as well as payment claims. In addition, these data sources combine with Smith's historic data in her electronic medical record (EHR) as well as with disease-specific, payment and personal data related to her condition and the broader population.

Segmenting and grouping that data appropriately gives analysts the ability to dig deeper and ask the right questions: How has she responded to specific treatments? What does a year in the life of her chronic condition look like? How does she compare to thousands of other patients with a similar condition? Each look requires a specific analysis. To assist in these grouping efforts, healthcare organizations deploy these three things:

Analytic Tools: Analytic tools enable the organization to establish the level of disease burden by assessing individual disease markers. This informs the predictive models and business intelligence tools that analysts apply to produce actionable insights.

Augmented Intelligence: Data aggregation and standardization are replete with errors. The claims system may not recognize the pharmacy or procedure code. A diagnosis might be missing. Treatment codes may be incompatible with the condition or diagnosis.

Augmented intelligence improves data quality. Clinical teams review application-generated feedback to identify which codes have created errors, then adjust the application to increase accuracy for both treatment and billing.

Natural Language Processing: Natural language processing is an algorithmic technology that translates spoken data entries into the medical records, a solution that is increasingly relevant with the rise in computing power and reliance on voice interaction with computers. Advances in NLP also enable analysts to extract data from large bodies of unstructured text and written documents, scans and images.





Stage 2: Determining Current/Future Care Utilization and Financial Risk

Grouping tools sort and mark all data related to Jane Smith's diabetes. Analysts combine this disease-specific information with information related to comorbidities to assess Jane's total disease burden. Jane is not alone, however. There are likely many other patients with similar health status and care needs.

Effective value-based care and population health requires understanding current care utilization and costs, and developing the ability to predict and manage future care utilizations and costs. What resources do patients like Jane actually require? What's the likelihood that those resource needs will increase exponentially next year?

An understanding of baseline and projected needs is the foundation for lowering costs and improving outcomes within risk-based contracts. This type of analysis can be generated through two things.

Business Intelligence: Business intelligence tools convert meaningful units of analysis into actionable insights. Assessing disease markers identifies the prevalence of disease and calculates associated care utilization and costs.

Predictive Analytics: Based on business intelligence reports, predictive models determine which patients/members will have the heaviest future disease burden, and who would benefit most from active care management and/or disease management interventions.





Stage 3: Targeting Opportunities to Improve Care Quality and Performance

With a solid understanding of the projected disease burden, treatments and aligned costs for Jane and her population cohort, healthcare organizations can determine the appropriate targeted intervention that will improve care quality and outcomes while reducing costs. The overall goal is to optimize care delivery, resource utilization and patient/employee experience. Analytic tools, machine learning and bots aid care professionals in delivering these targeted interventions. **Analytic Tools:** An assortment of analytic tools help identify gaps in care. Potential gaps include:

- Care inconsistent with clinical guidelines
- Patients' non-adherence to treatment and medication plans
- Lab reports indicating poorly controlled or managed disease
- Inappropriately monitored treatments



Gap analysis can determine which patients and treatments will likely incur unnecessary care and resources. These patients may benefit from enhanced engagement, direct interventions and/or enrollment in disease management programs. Not all gaps are equal. It is important to address care gaps where targeted interventions offer the highest potential for improved care outcomes.

With these insights, healthcare organizations can make strategic decisions regarding the treatment of specific populations that absorb disproportionate care resources. Depending on overall care utilization and financial risk predictions, the organization might focus more directly on populations with specific chronic diseases or on individuals who are frequent emergency department visitors, for example. Organizations can also use gap analysis analytic tools to enhance clinician performance. If physicians drive disproportionate utilization and costs, are there specific approaches that can be improved to drive better outcomes, processes and results?

Machine Learning: Machine learning tools can help healthcare organizations improve clinical and business processes, drive efficiencies and optimize performance. However, these tools require careful oversight since they can also potentially reinforce inherent errors and biases.

Bots: As organizations move increasingly to automated systems for triaging and directing routine care, bots help process data through machine learning software programs and direct members/patients to the most appropriate administrative or care delivery resources.

CONCLUSION: VALUE-BASED CARE REQUIRES CONTINUOUS LEARNING

To transition to value-based care delivery and population health management, healthcare organizations must engage with data in new and powerful ways.

Success at risk-based contracting requires constant improvements in outcomes, costs and performance. Those capabilities can only be achieved by leveraging massive data sets for actionable knowledge and insight. The curation process is an essential step, enabling the organization to continuously monitor and apply data in ways that inform and enhance ongoing analysis. The right technology solutions and tools can support healthcare professionals in making the decisions that improve processes, care outcomes and consumer experience. In this age of liquid data, human-machine collaboration offers a path to better care and more robust financial performance.

AUTHORS



David Johnson is the CEO of 4sight Health, a thought leadership and advisory company working at the intersection of strategy, economics, innovation and capital formation. Dave wakes up every morning trying to fix America's broken healthcare system. Prior to founding 4sight Health in 2014, Dave had a long and successful career in healthcare investment banking. He is a graduate of Colgate University and earned a Masters in Public Policy from Harvard Kennedy School. Employing his knowledge and experience in health policy, economics, statistics, behavioral finance, disruptive innovation, organizational change and complexity theory, Dave writes and speaks on pro-market healthcare reform. His first book *Market vs. Medicine: America's Epic Fight for Better, Affordable Healthcare*, and his second book, *The Customer Revolution in Healthcare: Delivering Kinder, Smarter, Affordable Care for All* (McGraw-Hill 2019), are available for purchase on www.4sighthealth.com.

4<u>sightHEALTH</u>



Dr. Thomas Lynn serves as Vice President and Chief Medical Officer at Conifer Health Solutions. In this role, Tom leads the clinical and financial analytics team responsible for measuring physician quality, provider performance and network costs. He provides perspective on innovating care delivery models and helping physicians shift to fee-for-value models. Tom began his career practicing family medicine and directing an emergency department in the U.S. Air Force and serving as an associate clinical professor with Duke University. He holds an MD from Georgetown University School of Medicine and a Master of Science in Biomedical Engineering - Clinical Informatics from Duke University. He has a BA from Amherst College where he studied Mathematics, Chemistry, and Neuroscience.

