The Healthcare Analytics Adoption Model: A Framework and Roadmap

Dale Sanders
Senior Vice President
Health Catalyst

Dr. David A. Burton
Executive Chairman
Health Catalyst

Denis Protti, Sc.D.
Professor Emeritus
University of Victoria

> What seems to be emerging in healthcare is a repeat of the trend of computerization and data management in other industries. Phase 1 is portrayed by systems that are designed specifically for supporting transaction-based workflow and data collection.

BACKGROUND

Over the last few years, there has been a flurry of activity around the topic of analytics (the discovery and communication of meaningful patterns in data) and even more recently, the use of "big data" (the collection of data sets so large and complex that it becomes difficult to process using on-hand database management tools or traditional data processing applications). This somewhat sudden interest in the topic can be traced back to a 2001 phone call between Dale Sanders, then serving as Director of Enterprise Data Warehousing for Intermountain Healthcare in Utah, and Pat Taylor of Blue Cross/Blue Shield of Alabama, during which time they founded the Healthcare Data Warehousing Association (HDWA) to accelerate the adoption and exploitation of analytics in healthcare. That phone call between two colleagues grew into a professional group that now includes over 300 organizational members in the U.S. and Canada.[1]

Culminating years of work in this arena and anticipating the healthcare industry's needs, Sanders published a commentary in 2012 in which he released the inaugural version of the Healthcare Analytics Adoption Model (HAAM), a proposed framework to measure the adoption and meaningful use of data warehouses and analytics in healthcare in ways similar to the well-known HIMSS Analytics EMRAM model.[2] After consultations and feedback from the industry, the second version of the HAAM is now being released.

> Healthcare is now entering Phase 3, the data analysis phase, which will be characterized by the adoption of enterprise data warehouses (EDW), now becoming synonymous with the term "Big Data."

THREE PHASES OF DATA ANALYSIS: DATA COLLECTION, DATA SHARING AND DATA ANALYTICS

What seems to be emerging in healthcare is a repeat of the trend of computerization and data management in other industries. Phase 1 of an industry's computerization is portrayed by systems that are designed specifically for supporting transaction-based workflow and data collection. In healthcare, this phase is characterized by widespread electronic medical record (EMR) adoption. In Phase 2, the need for sharing data among members of the workflow team becomes apparent. In the case of healthcare, this phase is characterized by health information exchanges (HIEs). In Phase 3 of computerization, organizations realize that the data they are collecting and sharing can be used to analyze aspects of the workflow that are reflected in the patterns of aggregated data. Healthcare is now entering Phase 3, the data analysis phase, which will be characterized by the adoption of enterprise data warehouses (EDW), now becoming synonymous with the term "Big Data."

This same three-phase evolution seen at the industry-level also applies at the micro-level within an organization. Early adopters of EMRs are thus more likely to have transitioned through these three phases, even though the healthcare industry as a whole has yet to do so. Organizations such as Intermountain Healthcare, using the HELP EMR, and the U.S. Veterans Affairs (VA) Health Care system using Vista, were also early pioneers in reaching Phase 3 of data management. Examples of integrated care models in the United States and beyond demonstrate that, when incentives are aligned and the necessary enablers are in place, the impact of leveraging big data can be very significant. The VA health system generally outperforms the private sector in following recommended processes for patient care, adhering to clinical guidelines, and achieving greater rates of evidence-based drug therapy. These achievements are largely possible because of the VA's performance-based accountability framework and disease-management practices enabled by EMRs and analytics allows them to frequently close the loop on clinical practices.

THE CHARACTERISTICS OF EACH LEVEL OF THE MODEL

The Analytics Adoption Model was designed purposely to mimic the benefits of a structured educational curriculum based on over 20 years of industry observation and lessons learned in healthcare. The curriculum is designed to ensure that organizations establish a foundational understanding of analytic technology and organizational use of analytics in step-wise fashion before attempting the more complicated topics of the upper levels. Also, each level of adoption includes progressive expansion of analytic capability in four critical dimensions:

(1) New Data Sources: Data content expands as new sources of data are added to the healthcare ecosystem.

(2) Complexity: Analytic algorithms and data binding become progressively more complex.

(3) Data Literacy: Organizational data literacy increases among employees, leading to an increasing ability to exploit data as an asset to organizational success, including new business and economic models.

(4) Data Timeliness: Timeliness of data content increases (that is, data latency decreases) which leads to a reduction in decision cycles and mean time to improvement.

> Organizations may find themselves operating quite effectively in Levels 5 or 6 but ineffectively at Levels 3 and 4.

HEALTHCARE ANALYTICS ADOPTION MODEL

Data binding grows in complexity with each level

Level 8	Personalized Medicine & Prescriptive Analytics	Tailoring patient care based on population outcomes and genetic data. Fee-for-quality rewards health maintenance.
Level 7	Clinical Risk Intervention & Predictive Analytics	Organizational processes for intervention are supported with predictive risk models. Fee-for-quality includes fixed per capita payment.
Level 6	Population Health Management & Suggestive Analytics	Tailoring patient care based upon population metrics. Fee-for-quality includes bundled per case payment.
Level 5	Waste & Care Variability Reduction	Reducing variability in care processes. Focusing on internal optimization and waste reduction.
Level 4	Automated External Reporting	Efficient, consistent production of reports and adaptability to changing requirements.
Level 3	Automated Internal Reporting	Efficient, consistent production of reports and widespread availability in the organization.
Level 2	Standardized Vocabulary & Patient Registries	Relating and organizing the core data content.
Level 1	Enterprise Data Warehouse	Collecting and integrating the core data content.
Level 0	Fragmented Point Solutions	Inefficient, inconsistent versions of the truth. Cumbersome internal and external reporting.

In addition to these trends within the model, organizations frequently operate at various stages of maturity in each level. In that regard, the model is not necessarily linear in its progression, although in an ideal state that would be the case. Organizations may find themselves operating quite effectively in Level 5 or 6 but ineffectively at Levels 3 and 4. Such was the case at Intermountain Healthcare during the early stages of their EDW development. Consequently, Intermountain adjusted its strategy and reassigned resources to address the laborious inefficiency of report production in Levels 3 and 4. Afterwards, the gains in efficiencies paid dividends in Levels 5 and above, where data architects and analysts were able to spend more time on market-differentiating analytics. Intermountain Healthcare was named the top Integrated Delivery Network in the U.S. market for seven of the eight years following this adjustment.

Level 0	Fragmented Point Solutions	Inefficient, inconsistent versions of the truth. Cumbersome internal and external reporting.

Level 0 of the Analytics Adoption Model is characterized by fragmented “point solutions” which have very focused, limited analytics capabilities, typically focused on departmental analytics such as finance, acute care nursing, pharmacy, laboratory or physician productivity. New knowledge generated by these solutions tends to be isolated to one area, which may encourage optimized sub-processes at the expense of enterprise-wide processes. The fragmented applications are neither co-located in a data warehouse nor otherwise architecturally integrated with one another. Overlapping data content leads to multiple versions of the truth. Reports tend to be labor intensive and inconsistent. There is no formal data governance function tasked with maximizing the quality and value of data in the organization.

> In Level 1, the beginnings of an enterprise data governance function are established with an initial focus upon reducing organizational and cultural barriers to data access, increasing data quality in the source systems and master data identification and management.

Point solutions in this level can satisfy the internal and external reporting that is important to Levels 3 and 4, but they are not a market differentiator and cannot scale to the more complicated analytic use cases and business models associated with the upper levels of adoption. Cumulatively, fragmented point solutions at this level also tend to require significantly more labor from data analysts and systems administrators to use and maintain than single, integrated data warehouses. The same inefficiencies of decentralization hold true for the fragmented costs of software licensing and vendor contract management.

Level 1 | Enterprise Data Warehouse

Collecting and integrating the core data content.

Level 1 is satisfied when core transaction source system data is integrated into an Enterprise Data Warehouse. At a minimum, the following data sources are co-located in a single local or hosted data warehouse: (1) HIMSS EMR Stage 3 clinical data, (2) financial data (particularly costing data), (3) materials and supplies data, and (4) patient experience data. If available, data content should also include insurance claims. A searchable metadata repository is available across the enterprise. The metadata repository provides natural language descriptions of the EDW content, describes known data quality issues and records data lineage. The metadata repository is the single most important tool for the complete democratization of data across the enterprise. The EDW data content is updated within one month of changes in the source systems.

The beginnings of an enterprise data governance function are established with an initial focus upon reducing organizational and cultural barriers to data access, increasing data quality in the source systems and master data identification and management. Data stewardship for the source data content areas in the EDW is forming under clinical and administrative ownership. Organizationally, it is best for the EDW to report to the CIO at this stage, assuming that the CIO can facilitate access to and the extraction of data from the source systems. Later, as the EDW evolves from the construction and early phases of adoption, the organizational alignment can change to another C-level executive who represents the functional use of analytics in the organization, such as the Chief Medical Officer or Chief Quality Officer.

Level 2 | Standardized Vocabulary & Patient Registries

Relating and organizing the core data content.

At this level, master vocabularies and reference data are defined and available in the EDW. These vocabularies and reference data include local master patient identity, physician identity, procedure codes, diagnosis codes, facility codes, department codes and others. Data stewardship for master data is functioning. Master vocabularies and reference data are identified and standardized across disparate source system data content in the EDW. Naming, definition and data types in the EDW data content areas are standardized according to local master reference data, enabling queries across the disparate source content areas. Patient registries based on billing codes and defined by multidisciplinary teams are available in the EDW to support basic analytics for the most prevalent and costly chronic diseases and acute care procedures in the local environment. Data governance forms around the definition and evolution of patient registries and master data management.

> The key criteria for success in Level 3 is efficiency and consistency of reports that are necessary for effective management, but alone are not enough to create differentiating value in the market.

Level 3 | Automated Internal Reporting

Efficient, consistent production of reports and widespread availability in the organization.

Level 3 is characterized by automated internal reporting where the analytic motive is focused on consistent, efficient production of reports required for: (1) executive and board level management and operation of the healthcare organization, and (2) self-service analytics for key performance indicators and interactive dashboards at the director and management level. The key criteria for success in this level is efficiency and consistency of reports that are necessary for effective management, but alone are not enough to create differentiating value in the market. Ideally, once developed and deployed, the maintenance of these reports requires little or no labor to support and are nearly entirely self-service. Also, the reports are reliable in their availability when needed, consistent and accurate, thus minimizing wasteful debate and the attractiveness of developing redundant reports that end users and analysts consider more reliable, consistent or accurate.

An analytic services user group exists that facilitates collaboration between corporate and business unit data analysts. Among other synergies, this group is organized to define consistent data definitions and calculation standards. Data governance expands to include data quality assurance and data literacy training and to guide the strategy to acquire mission-critical data elements in subsequent levels of adoption.

Level 4 | Automated External Reporting

Efficient, consistent production of reports and adaptability to changing requirements.

The analytic motive at this level is focused on consistent, efficient and agile production of reports required for external needs, such as: (1) regulatory, accreditation, compliance and other external bodies (e.g. tumor and communicable disease registries); (2) funding and payer requirements (e.g. commercial financial incentives and federal Meaningful Use payments); and (3) specialty society databases (e.g. national cardiovascular data registry). Master data management at this level requires data content in the EDW that has been conformed to current versions of industry-standard vocabularies such ICD, CPT, SNOMED, RxNorm, LOINC and others. In addition to the low-labor, low-maintenance requirement for producing reliable, accurate and consistent reports at this level, the EDW must be engineered for agility in this context, due to the constantly changing nature of external reporting requirements.

Data governance and stewardship is centralized for external reporting. Stewardship processes exist to maintain compliance with external reporting requirements and govern the process for approving and releasing the organization’s data to external bodies.

EDW data content at this level has been expanded to include text data from patient-record clinical notes and reports. EDW-based text query tools are available to support simple keyword searches within and across patient records.

"In Level 5, data governance expands to support multidisciplinary care management teams that are focused on improving the health of patient populations."

Level 5 | **Waste & Care Variability Reduction**

Reducing variability in care processes. Focusing on internal optimization and waste reduction.

At Level 5, organizations are moving away from utilitarian internal and external reporting. They have a significant opportunity to differentiate themselves in the market based on quality and cost and enabled by analytics. Data at this level is used explicitly to inform healthcare strategy and policy formulation. The analytic motive is focused on measuring adherence to clinical best practices, minimizing waste and reducing variability, using variability as an inverse proxy for quality. Data governance expands to support multidisciplinary care management teams that are focused on improving the health of patient populations. Population-based analytics are used to suggest improvements to individual patient care. Permanent multidisciplinary teams are in place to continuously monitor opportunities that will improve quality and reduce risk and cost across acute care processes, chronic diseases, patient safety scenarios and internal workflows.

The precision of registries is improved by including data from lab, pharmacy and clinical observations in the definition of the patient cohorts. The EDW content is organized into evidence-based, standardized data marts that combine clinical and cost data associated with patient registries. The data content expands to include insurance claims (if not already included) and HIE data feeds. On average, the EDW is updated within one week of source system changes.

Level 6 | **Population Health Management & Suggestive Analytics**

Tailoring patient care based upon population metrics. Fee-for-quality includes bundled per case payment.

Level 6 is characterized by organizations that have achieved a sustainable data driven culture and established a firm analytic environment for understanding clinical outcomes. The "accountable care organization" shares in the financial risk and reward that is tied to clinical outcomes. At least 50 percent of acute care cases are managed under bundled payments. Analytics are available at the point of care to support the Triple Aim of maximizing the quality of individual patient care, population management and the economics of care. EDW data content expands to include bedside devices, home monitoring data, external pharmacy data and detailed activity-based costing.

Data governance plays a major role in the accuracy of metrics supporting quality-based compensation plans for clinicians and executives. On average, the EDW is updated within one day of source-system changes. The EDW reports organizationally to a C-level executive who is accountable for balancing cost of care and quality of care.

Level 7 | **Clinical Risk Intervention & Predictive Analytics**

Organizational processes for intervention are supported with predictive risk models. Fee-for-quality includes fixed per capita payment.

Level 7 organizations are able to move into the arena of predictive analytics by expanding on their optimization of the cost per capita populations and capitated payments. Their focus expands from the management of cases to collaboration with clinician and payer partners to manage episodes of care, including predictive modeling, forecasting and risk stratification.

> iPhones, cloud computing, gene sequencing, wireless sensors, modernized clinical trials, internet connectivity, advanced diagnostics, targeted therapies and other science will enable the individualization of medicine – and force overdue radical change in how medicine is delivered, regulated and reimbursed.

The analytic motive at this level expands to address diagnosis-based, fixed-fee-per-capita reimbursement models. Focus expands from management of cases to collaboration with clinician and payer partners to manage episodes of care using predictive modeling, forecasting and risk stratification to support outreach, triage, escalation and referrals. Physicians, hospitals, employers, payers and members/patients collaborate to share risk and reward (e.g., financial reward to patients for healthy behavior).

Patients who are unable or unwilling to participate in care protocols due to constraints such as cognitive disability, economic inability, geographic limitations to care access, religious restrictions and voluntary non-participation are flagged in registries. Data content expands to include home monitoring data, long-term care facility data and protocol-specific patient reported outcomes. On average, the EDW is updated in one hour or less of source system changes.

Level 8 | Personalized Medicine & Prescriptive Analytics

Tailoring patient care based on population outcomes and genetic data. Fee-for-quality rewards health maintenance.

At Level 8, the analytic motive expands to wellness management, physical and behavioral-functional health and mass customization of precise, patient tailored care. Healthcare-delivery organizations are transformed into health-optimization organizations under direct contracts with patients and employers. Fixed-fee, per capita payment from patients and employers for health optimization is preferred over reimbursement for treatment and care delivery. Analytics expands to include natural language processing (NLP) of text, prescriptive analytics and interventional decision support. Prescriptive analytics are available at the point of care to improve patient specific outcomes based upon population outcomes.[3] Data content expands to include 7x24 biometrics data, genomic data and familial data. The EDW is updated within minutes of changes in the source systems.

At this level, healthcare organizations are completely engaged as a data-driven culture and shift from a fixation with care delivery to an obsession with risk intervention, health improvement and preventive medicine. New data content in the enterprise data warehouse is combined with not-yet-discovered algorithms that can identify relationships between genomics, family history and patient environment. Eric Topol's book, *The Creative Destruction of Medicine: How the Digital Revolution Will Create Better Health Care*, portrays how medical innovation will coalesce to change clinical practice and what the coming changes mean for today's policy debate.[4] In Dr. Topol's vision, iPhones, cloud computing, gene sequencing, wireless sensors, modernized clinical trials, internet connectivity, advanced diagnostics, targeted therapies and other science will enable the individualization of medicine – and force overdue radical change in how medicine is delivered, regulated and reimbursed. But unlike any prior time in medicine, this revolution is superimposed on a world of social networking, omnipresent smart phones with pervasive connectivity and ever-increasing bandwidth. This great convergence will usher in the creative destruction of medicine. At the same time, consumers have an unprecedented capacity to take charge – it is *their* DNA, *their* cell phone, *their* precious individual information.

The resulting analytics will be applied early in the patient's life to develop a lifelong health optimization plan. When healthcare delivery is required, the patient's treatment protocol is tailored specifically to that patient based upon the insights gained from these new data sources and algorithms. The boundaries of evidence-based medicine

> The ROI of EMR investments, let alone impactful health reform, will not be realized until the healthcare industry invests in enterprise data warehousing and commits culturally to the exploitation of analytics — that is, to become a data-driven culture, incented economically to support optimum health at the lowest cost.

are extended beyond the limited applicability of randomized clinical trials to include the quasi-experimental evidence that emerges from local and regional enterprise data warehouses. This locally derived evidence is shared with commercial clinical content providers to iteratively enhance the knowledge content from randomized clinical trials.

IN CONCLUSION

Healthcare around many parts of the world has been moving through three phases of computerization and data management simultaneously: data collection, data sharing and, now, gradually into data analytics. The data-collection phase, characterized by the urgent deployment of EMRs, will not have a significant impact on the quality or cost of healthcare. Numerous retrospective studies of EMR deployment have yet to reveal anything other than a very modest return on investment.[5] The overwhelming failure rate of health information exchanges due to unsustainable economic models is also well documented.[6] However, the investment in EMRs is fundamentally required to achieve the value that is accessible in analytics. The return on investment of EMRs, let alone impactful health reform, will not be realized until the healthcare industry invests in enterprise data warehousing and commits culturally to the exploitation of analytics — that is, to become a data-driven culture, incented economically to support optimum health at the lowest cost.

Current adoption rates of data warehousing and analytics stand at only 10 percent and just a small subset of those early adopters operate above Level 3; none operates consistently above Level 5. In informal polls conducted by Sanders during webinars on this topic, webinar participants consistently report their organization operating between Levels 2 and 3, no higher. By observing the events and tools that encouraged the adoption of EMRs, notably the EMRAM, the Healthcare Analytics Adoption Model follows suit and provides a framework for more rapid progression to analytic maturity.

Notes

1. *Healthcare Data Warehousing Association (HDWA).* 2012. Retrieved October 1, 2012. http://hdwa.org/hdwa/home/.
2. Sanders D. *A Model for Measuring Industry-Wide Adoption and Capability of Healthcare Analytics and DataWarehousing in the USA.* ElectronicHealthcare. Vol.11 No.2 2012.
3. Pestotnik, Classen, Evans, and Burke. *Implementing Antibiotic Practice Guidelines through Computer-Assisted Decision Support: Clinical and Financial Outcomes.* Ann Intern Med. 1996;124(10):884-890.
4. Topol E. *The Creative Destruction of Medicine: How the Digital Revolution Will Create Better Health Care.* Basic Books. January 2012.
5. Miliard M. *Hospital Execs Slow to Measure Technology ROI.* Healthcare IT News. October 25, 2012.
6. *Survey: Lack of Interoperability Hinders Health Data Exchange*. iHealthbeat. October 04, 2012.

ABOUT THE AUTHORS

Dale Sanders, *Senior Vice President, Health Catalyst*

Prior to his work in the healthcare industry, Dale Sanders worked for 14 years in military, national-intelligence and manufacturing sectors, specializing in analytics and decision support. In addition to his role at Health Catalyst, Sanders serves as the senior technology advisor and CIO for the National Health System in the Cayman Islands. Previously, he was CIO of Northwestern University Medical Center and regional director of Medical Informatics at Intermountain Healthcare, where he served in a number of capacities, including chief architect of Intermountain's enterprise data warehouse. Sanders is a founder of the Healthcare Data Warehousing Association. He holds Bachelor of Science degrees in Chemistry and Biology from Ft. Lewis College and is a graduate of the U.S. Air Force Information Systems Engineering program.

David A. Burton, *M.D., Executive Chairman, Health Catalyst*

A former Senior Vice President of Intermountain Healthcare where he served in a variety of executive positions for 26 years, Dr. Burton spent the last 13 years of his career co-developing Intermountain's clinical process models utilized within the EDW environment. Dr. Burton is the former founding CEO of Intermountain's managed care plans (now known as SelectHealth), which currently provide insurance coverage to approximately 500,000 members.

Denis J. Protti, *Sc.D., Professor Emeritus, University of Victoria*

Dr. Protti was the founding Director of the University of Victoria's School of Health Information Science and a former faculty member. Prior to joining the university, Dr. Protti held senior information systems executive positions in Manitoba and British Columbia hospitals. His research and areas of expertise include National Health Information Management & Technology Strategies, Electronic Health Records, Primary Care Computing and Evaluating Information Systems. Dr. Protti was also the first recipient of the Canadian Health Leadership Network's MacNaught-Taillon Award for his contributions to Canadian healthcare. In May 2009, Dr. Protti was granted an Honorary Doctor Science from City University London for his contributions to the British healthcare system. In 2012, he was the inaugural recipient of the Techna Health Innovator Award.

About Health Catalyst

Based in Salt Lake City, Health Catalyst delivers a proven, Late-Binding™ Data Warehouse platform and analytic applications that actually work in today's transforming healthcare environment. Health Catalyst data warehouse platforms aggregate and harness more than 3 trillion data **points** utilized in population health and ACO projects in support of over 22 million unique patients. Health Catalyst platform clients operate 96 hospitals and 1,095 clinics that account for over $77 billion in care delivered annually. Health Catalyst maintains a current KLAS customer satisfaction score of 90/100, received the highest vendor rating in Chilmark's 2013 Clinical Analytics Market Trends Report, and was selected as a 2013 Gartner Cool Vendor. Health Catalyst was also recognized **in 2013** as one of the best places to work by both Modern Healthcare magazine and Utah Business magazine.

Health Catalyst's platform and applications are being utilized at leading health systems including Allina Health, Indiana University Health, Memorial Hospital at Gulfport, MultiCare Health System, North Memorial Health Care, Providence Health & Services, Stanford Hospital & Clinics, and Texas Children's Hospital. Health Catalyst investors include CHV Capital (an Indiana University Health Company), HB Ventures, Kaiser Permanente Ventures, Norwest Venture Partners, Partners HealthCare, Sequoia Capital, and Sorenson Capital.

Visit healthcatalyst.com and follow us on Twitter, LinkedIn, Google+ and Facebook.

3165 East Millrock Drive, Suite 400
Salt Lake City, Utah 84121
ph. 800-309-6800