IT ALL STARTS WITH A DATA WAREHOUSE

Why the Healthcare Data Warehouse Is Becoming the Critical Foundational Platform for Analytics Success in the Upcoming Healthcare Transformation Environment

Health Catalyst
ignite outcomes improvement™
It All Starts with a Data Warehouse

Why the Healthcare Data Warehouse Is Becoming the Critical Foundational Platform for Analytics Success in the Upcoming Healthcare Transformation Environment
INTRODUCTION

One day in 1928, Alexander Fleming, a Scottish biologist, pharmacologist, and botanist, neglected to clean his workstation before going on vacation. When he returned, Fleming noticed a strange fungus on some of his cultures. Curiously, bacteria seemed not to thrive near those cultures. The rest of the story is history, and thus, quite by accident, did penicillin become the first — and still one of the most widely used — antibiotics.

Accidents happen in technology, too. And we’re glad they do. Back in the 1980s, a Wal-Mart computer operator one day grew tired of retrieving archival tapes for historical sales data. So he secretly “borrowed” excess storage space on a company server, where he downloaded and stored data from the most requested tapes. This giant data stash couldn’t stay secret for long, and it didn’t. When Wal-Mart managers found it, they quickly realized the enormous value of timely and widespread access to data. Thus was born the Wal-Mart data warehouse (although the roots of data warehousing date to the 1960s). Soon, every transaction in 6,000 Wal-Mart stores was available for analysis in the data warehouse within seven minutes. This treasure trove of data supported Wal-Mart’s strategic planning and astonishing growth.

Wal-Mart’s story shows that good things do indeed happen by accident. The company’s story also underscores the power of analytic technology coupled with a data-driven culture.

While Wal-Mart, other retailers, and companies in most verticals have adopted data warehouse solutions, healthcare organizations have failed to do so broadly or comprehensively. This is surprising, considering the volume and variety of data residing in the systems of hospitals and group practices. And it’s unfortunate, considering how very much is at stake: the health of Americans, the cost and quality of care delivery, and the productivity of our country.

Letting go of old beliefs and making way for new ones is never easy. I can’t help but go back to my days studying physics and philosophy and recall the evolution of the way we perceive the universe. For centuries, Western man accepted Ptolemy’s claim that the earth was the center of the universe. Then, 1,500 years later, Copernicus came along and said, “You know what? My observations indicate that the sun is actually the center.” Accepting this paradigm shift was difficult for most, but it proved to be right.

The same kind of revolution needs to happen in the healthcare industry. In the very near future, we must shift away from the very patient-specific, billing-centric
electronic medical record (EMR) as the center of the universe to an analytic-centric environment based on an enterprise data warehouse (EDW).

Once, all things data-related revolved around the EMR, and it held prime position at the center of a healthcare organization’s universe. Today, organizations must understand that the EMR is simply one source of collected data supporting a particular workload. The EMR generally has basic integration within a local area network and may be connected to the external environment through a health information exchange. But this is insufficient for the sophisticated data analytics the industry needs today.

If you’re going to achieve high performance analytics, the EMR alone won’t cut it. You need an enterprise data warehouse (EDW). In fact, there is no viable alternative to an EDW if you want to successfully use analytics to improve the cost and quality of care. By incorporating the EMR’s and other systems’ data into an EDW, you create an infrastructure that enables health systems, accountable care organizations (ACOs), physician groups, and others to predict and manage patient care and improve cost and quality.

Our goal in this e-book is to provide the business and clinical leaders of healthcare organizations with all they need to know to leverage healthcare enterprise data warehouse solutions to dramatically improve their own clinical and operational performance, as well as U.S. healthcare generally.

We, the authors of this e-book, are members of the leadership team of Health Catalyst®, a healthcare information technology company that brings decades of experience at leading health systems including Intermountain Healthcare, Northwestern, and PeaceHealth. Health Catalyst has applied the principles learned at these health systems into a commercial-grade, stress-tested data platform that scales vertically and horizontally to meet the needs of everyone from small, rural hospitals to large, multi-state health systems with dozens of hospitals. It performs effectively in academic medical centers and in specialty hospitals and supports inpatient, ambulatory, and population management initiatives.

Healthcare in the United States is undergoing dramatic, unprecedented change — shifting from fee-for-service to fee-for-value without the historical investment in analytics technology so common in other industries and so essential to success. Hospitals and medical practices need enterprise data warehouse, or EDW, technology to help answer mission-critical questions and repeatedly and reliably deliver insights about clinical and operational performance.

Dale Sanders
Sr. Vice President
Health Catalyst
Arming yourself with knowledge about the platforms, options, vendors, and unique needs of the healthcare industry will help you ensure the data warehouse you select leads directly to success.

What a Data Warehouse Does for Healthcare

Organizing, cataloging, and structuring information for the benefit of the user — it’s what a data warehouse does so well. It allows teams to turn to the data warehouse to make queries, conduct research, and analyze trends. For healthcare, a data warehouse is essential to making the most informed business and clinical decisions.
Essentially, the utility of an enterprise data warehouse (EDW) that saves a copy of data collected from internal source systems as well as external registries is limitless. Reports generated and the addition of graphical formats, eye-catching charts, and comparisons allow teams across the healthcare system to see and prioritize areas with the greatest opportunities for improving outcomes. Everyone from clinicians to analysts benefits from an enterprise data warehouse. Highlighting specific measures for improvement, an EDW helps an organization critically evaluate not only care processes for chronic diseases but procedure-specific clinical process as well.

But not all EDWs are the same, and this book aims to explain some of the differences and how these can result in varying degrees of effectiveness. As a baseline, however, the following characteristics should comprise any well-designed EDW:

- **Dependent on Multiple Source Systems.** A data warehouse is populated by at least two source systems, also called transaction or production systems. Examples include EHRs, billing systems, registration systems, and scheduling systems. In large enterprises, it is not unusual for a data warehouse to contain data from as many as 50 different source systems, internal and external.

- **Cross-Organizational Analysis.** Data warehouses are designed specifically to enable data analysis across business and clinical processes — that is, the ability to analyze and link data across multiple source systems that support various business processes, particularly the full continuum of care for a patient, from birth to death. For example, a data warehouse could enable the analysis of data from an EHR coded in SNOMED and data from a billing system coded in ICD by aggregating the key elements required for the analysis from each system, regardless of the terminology used.

- **Trends, Metrics, and Reports.** A data warehouse helps identify trends and previously unknown relationships in business processes. The data output is characterized by metrics and reports. In large enterprises (15,000 employees and more), it is not unusual for a data warehouse to produce hundreds of reports and process tens of thousands of queries per month.

- **Large.** Data warehouses in today’s information intensive environments may contain billions of records constituting dozens and even hundreds of terabytes of data.

- **Historical.** A data warehouse stores many years of data, typically at least five- and sometimes as much as 30-years’ worth.

Beyond these key, common components, there are structural differences and functional nuances that will turn the EDW your organization chooses
into an indispensible tool. Arming yourself up front with knowledge about the platforms, options, vendors, and unique needs of the healthcare industry will help you ensure the EDW you select leads directly to success.
WHY AN EDW IS A FOUNDATIONAL PLATFORM FOR FUTURE ANALYTIC SUCCESS

The aggregation and integration of clinical and financial data are among the greatest virtues of an EDW.

Visualize for a moment the reporting rush. Finance, quality, human resources, and clinical departments scramble to compile data for their reports. Once each department completes its task, the group must combine its reports so management can review progress on organizational goals. Management wants to see the big picture and draw valid conclusions about quality, satisfaction, and cost performance across the organization.
While this can be a stressful, time-consuming process, a healthcare EDW can ease management reporting and improve its efficiency in three ways: by enabling a more efficient, scalable process; ensuring consistent data that everyone can trust; and enabling meaningful, targeted quality improvement. Here’s how.

1. Enabling a More Efficient, Scalable Reporting Process

Typically, hospital or group-practice executives meet to determine the categories of healthcare data they need in order to track progress toward strategic goals. You may already have a process in place for getting financial data. But now, with new value-based purchasing pressures requiring clinical and financial data, you’re suddenly tasked with getting more data than ever before. Questions may arise, including:

- Where do I start?
- Who do I approach for the data?
- Where is the data stored?
- How do I get at the data after I’ve found it?
- How do I compile and make sense of the data?

Locating the right people with the right data — whether it’s a single person who updates an Excel document or a team overseeing a database — can be a time-consuming, manual process. You’ll spend a lot of time setting up this process to gather and compile data to keep the executives’ dashboard up to date.

A healthcare EDW streamlines and scales this process. It integrates disparate data from a wide variety of sources, including billing, financial, patient satisfaction, and clinical sources. You can access it in the same place every month. And with the tools the healthcare EDW delivers, you can spend your time analyzing and interpreting the data, running visualizations and reports, and gaining an understanding of the best ways to achieve quality and cost goals.

The process enabled by a healthcare EDW is highly preferable to the tired, inefficient approach of chasing down data and sources month after month. It also provides the healthcare analytics leadership needs to make the most informed decisions for continued success.

2. Ensuring Consistent Data Everyone Can Trust

Maybe you’ve been in this situation before: at a monthly meeting concerning the organization’s finances, someone begins presenting when you suddenly realize that your data contradicts the data on his or her slides. Why does the
data show a different trend on net income? Why does one set of data show that length of stay (LOS) is going down when another one shows just the opposite?

When people throughout an organization access information in many different ways and from many sources, discrepancies become more likely to occur. The question then becomes which data to trust.

A healthcare EDW establishes a single source of truth and enables healthcare analytics. When data definitions and tools are consistent, as in a healthcare EDW, everyone can rely on the accuracy of the information used to drive critical decisions. An EDW also serves as a foundation for developing and maintaining a data governance program. With such a program, data owners and experts can identify data issues within the organization, resolve them, determine who needs to use the data, and define the best access path to the data.

The aggregation and integration of clinical and financial data are among the greatest virtues of an EDW. Consider this real-life example: a hospital targeted length of stay following appendectomies as a key opportunity for quality and financial improvement. Its frontline team looked at the data in its EDW and discovered clinicians were prescribing a wide range of antibiotics after appendectomies. Based on the outcomes data for each antibiotic, the team decided on a protocol for the entire facility and encouraged clinicians to prescribe one particular antibiotic following surgeries. The antibiotic was expensive, so clinicians ordinarily might have suggested a lower-cost, alternative medication. Since the clinical and financial data had been linked in the data warehouse and made available on dashboards, clinicians could see that while pharmacy costs rose, a parallel drop in length of stay more than offset the costs. When clinicians saw the encouraging data, they requested further refinements to the system. They asked the team questions such as, “Are we giving patients the antibiotic at the correct time after surgery?” The clinicians had engaged, becoming willing and proactive participants in the process.

3. Enabling Meaningful, Targeted Quality Improvement

On an ongoing basis, multi-disciplinary teams from across clinical, technical, financial, quality, and performance-excellence departments can meet to evaluate a hospital or group practice’s quality measures and use the EDW to identify opportunities for improvement. The organization then can develop and deploy highly targeted, specific interventions to promote those improvements in care, whether it’s reducing early term deliveries, lowering the rate of septicemia, or eliminating unnecessary X-rays.

Data-driven quality improvement is what implementing an EDW is all about. An EDW can drive consistent insights, better collaboration, and more streamlined processes across a healthcare organization.
Without a way of organizing all sources — clinical, financial, patient satisfaction, and administrative data — into a single source of truth, a healthcare organization is unable to harness the real power of its data.

If not an EDW, then what? That’s a good question. While there are alternatives to an EDW, nothing else creates a healthcare analytics foundation that’s as solid as an EDW. Consider the following approaches, each of which has great promise but fails to live up to the reality of a true EDW:
The Report Factory

The report factory approach uses an analytics platform alone and assumes that if you build it, people will come. When a healthcare institution creates or buys an analytics platform, one of two things happens: users come, or they don’t.

When they do come, the first indication is a backlog of report requests to the IT department. Why? Data can be addictive, and once users get a taste, they want more and more data until a dependency is formed. It’s not uncommon to see queues of report requests into the thousands and growing. Soon, clinicians and department heads decide the analytics platform or the IT shop is too slow, and they hire their own dedicated analysts or architects.

And sometimes the systems simply weren’t created to provide the information users really need. This results in an even bigger problem — the users never come at all and the effort it took to develop the analytics solution goes to waste.

Flavor of the Month

In order to avoid becoming a report factory, organizations sometimes opt for a different, more measured approach when developing their analytics platform. This, however, can result in a project-by-project or flavor-of-the-month approach to analytics.

Here’s how it happens: the hospital or medical practice may tackle projects based on some sort of prioritization — possibly responding to squeaky wheels or management’s pet projects. Initially, the organization may experience employee enthusiasm, improved care, and reports that really do help. Inevitably though, clinicians and staff grow dispirited. It becomes difficult to keep more than a handful of projects. Early gains are quickly lost.

Point Solutions, a.k.a., One-Solution Wonders

Yet another route an organization may take when developing their analytics platform is the deployment of one or more solution-specific apps. Often called “best-of-breed” or “point solutions,” these apps focus on a single goal and a single slicing of the data. Ministry Health Care’s CIO, Will Weider, defines them as extract-ware because they receive “an extract of your data with the promise of allowing you to have the ability to analyze and monitor your data.”

One problem with this approach is something called sub-optimization. While the organization may be able to optimize the specific area of focus, single-focused point solutions offer little insight into likely impacts both up and downstream.

Another problem is the technology spaghetti bowl. When a hospital or group practice has only a few point solutions in its dish, a small IT shop can provide adequate support. But with additional point solutions (consider them noodles, if you will) an IT department finds it all but impossible to unravel the spaghetti.
Eventually, one of the senior IT employees may own this mess, effectively holding the organization and improvement efforts hostage. Worse yet, if this IT professional resigns, a vacuum is left behind.

Point solutions typically result in multiple contracts, multi-cost dependencies, and multiple interfaces. These, in turn, lead to complexity, confusion, and organizational chaos that impede improvement and continued success.

**The EHR Enigma**

While this is clearly a necessary step towards data-driven care delivery, an EHR system alone is insufficient to enable an enterprise-wide, consistent view of data from multiple sources. The conversion of clinical data from paper to an electronic format is a necessary step; it allows for the use of data to improve care. However, without a way of organizing all sources — clinical, financial, patient satisfaction, and administrative data — into a single source of truth, a healthcare organization is unable to harness the analytic power of the data. For actionable clinical, financial, and operational insights, an EDW is necessary. Only an EDW enables near real-time capture, aggregation, and analysis of data from the EHR and other internal and external systems that reside in silos.

**The Partial Data Warehouse**

Most healthcare data warehouses currently in use cannot deliver a complete, integrated view across the enterprise. Many health systems have a partial data warehouse, and this is a good start. But without an enterprise-wide view that incorporates all major sources of data within the health system, there is no single platform and source of truth that can scale to current and future needs. It’s rare to find hospitals or medical practices that have built a true enterprise-wide healthcare EDW because it’s very difficult to do without a flexible, agile architecture. Only when clinical data is married with financial, administrative, and patient satisfaction data can breakthrough analytics, the kind that lead to significant care improvement and efficiencies, be realized.
In the healthcare industry where the data environment is much more complicated than a sales receipt and the analytic use cases are constantly changing, early-binding data models can be disastrous to agility and initial time-to-value.

The Concept of Data Binding

Data can be bound to business rules that are implemented as algorithms, calculations, and inferences acting upon that data. Examples of binding data to business rules in healthcare include:
Calculating length of stay (LOS)

Attributing a primary care provider to a particular patient with a chronic disease

Calculating revenue (or expense) allocation and projections to a department or physician

Data definitions of general disease states for patient registries

Defining patient exclusion criteria for disease/population management

Defining patient admission, discharge, and transfer rules

Data can also be bound to vocabulary terms, for both local and industry standards. Examples of vocabulary binding include:

- Patient identifier
- Provider identifier
- Location of service
- Gender
- Diagnosis code
- Procedure code

Knowing when and how tightly to bind data to rules and vocabularies is critical to the agility and success — or failure — of a data warehouse. In healthcare, the risks of binding data too tightly to rules or vocabularies are particularly high because of the volatility of change in the industry. Business rules and vocabulary standards in healthcare are among the most complex in any industry, and they undergo almost constant change.

**Lessons from Software Engineering**

The idea of late binding in data warehousing borrows from the lessons learned in the early years of software engineering. In those early years, very large software programs characterized software development — it was common to program hundreds of thousands of lines of code in a single module, supporting numerous and widely different business functions. The code for these varied business functions was tightly bound (also known as coupled) together all at once, at compile time. It was a time-consuming process to write and troubleshoot these large programs. If one piece of the program failed at compile time, the entire program failed. It was all-or-nothing programming. Also, if the program required changes or modifications because of new business rules and requirements after compile time, the entire
program had to be modified, re-compiled, and placed back in service, often with significant downtime. Agility suffered enormously.

Object Oriented Programming and Late Binding

In the 1980s, software engineering practices changed significantly, moving away from large, tightly coupled, early binding programs. Alan Kay from the Universities of Colorado and Utah and Xerox/PARC introduced the concept of late binding and object-oriented programming. This new approach was based upon two radically new concepts:

1. Writing code in smaller modules or objects that were modeled after processes and services in the real world that the software was designed to support; and

2. Binding these software objects at run time, not compile time, and only when those objects were needed to support the services they reflect.

Alan Kay’s new concepts for software engineering sat underutilized and largely unknown, confined to PARC and academic circles until Steve Jobs founded NeXT. Jobs was not a programmer, but he instinctively understood the elegance of Kay’s concepts. Object-oriented, late-binding software engineering became the standard practice at NeXT and paved the way to commercial, large-scale adoption of Kay’s philosophies. Steve Jobs receives due credit for his innovation and leadership at Apple, but by making object-oriented, late-binding software a new commercial norm at NeXT, he paved the way for the entire software revolution in Silicon Valley. The agility, scalability, and performance of platforms such as Amazon, Google, Facebook, and Salesforce were enabled by this new approach to software engineering.

Data Engineering and Late Binding

After witnessing and reflecting upon the failure of several multimillion-dollar data warehousing projects in the U.S. military, Dale Sanders, now Senior Vice President for Strategy at Health Catalyst, saw the same patterns in data engineering as those in software engineering prior to object oriented programming. Early and tight binding of data to rules, models, and vocabulary led to unnecessary complexity, delaying time-to-value. This approach also created a fragile and inflexible data warehouse infrastructure, which lacked the ability to adapt to rapidly changing analytic use cases or new data content.

In the late 1990s, while Sanders was employed by TRW Inc., he was sponsored by the Pentagon to study advanced decision support in nuclear warfare operations — a project called the Strategic Execution Decision Aid (SEDA). He turned to the healthcare industry for what he expected to be role-model examples of computer-aided analytics to drive better decisions.
in time-critical, life-critical situations but instead found almost no examples, with the notable exception of a scattered few at Intermountain Healthcare in Salt Lake City. Intermountain clearly possessed the culture and willingness to fully leverage data for improving care, but the industry at large was many years behind. Anticipating the eventual demand for analytics in the industry, Sanders made a career transition from the military, national intelligence, and manufacturing sectors into healthcare.

Sanders’s late-binding data engineering concept is now fundamental to Health Catalyst’s data warehouse platform. The Late-Binding™ Data Warehouse enables time-to-value that is measured in days and weeks, not months and years, and has proven many times more scalable and adaptable to new analytic use cases and data content than methodologies utilizing early binding, tightly coupled data models and vocabulary management.

**Data Binding Points in an Enterprise Data Warehouse**

There are six points in a data warehouse at which data can be bound to rules and vocabularies. As the data flows from left to right in the diagram below, points 1 and 2 are appropriate for binding to rules and vocabularies that exhibit low volatility; that is, those rules and vocabularies that change infrequently, such as patient identifiers and provider identifiers. Late binding — at points 4 and 5 — is appropriate for rules and vocabulary that are likely to change on a regular basis, or for which no standard rule or vocabulary exists. For example, binding in the visualization layer is appropriate for the what-if scenario analysis associated with modeling different reimbursement models or defining disease states. Once the exploratory what-if phase is complete, the new models and definitions can be locked down and bound in points 3, 4, or 5.

It is a best practice to retain a record of the bindings in the data warehouse. This record will allow analysts to quickly run models on rules and vocabulary (e.g., ICD-9 to ICD-10) that change over time, which is helpful for forecasting and predictive analytics. At Health Catalyst, we recommend embedding the history of vocabulary and business rule binding into the data structures of the data warehouse so they become a self-contained configuration control library that can easily be used to retrace analytic history when necessary.
Knowing what to bind and when in the flow of data in a data warehouse requires more than technical skills. Data engineers and architects who work in a Late-Binding™ Data Warehouse environment must possess a strategic understanding of the short and long-term evolution of the entire industry. They must appreciate the historic volatility of vocabulary and business rules as well as an ability to predict the velocity and the specifics of volatility in the future. Healthcare is undergoing changes to business rules and vocabulary at an unprecedented rate. Data warehouses must be designed to keep pace with the market, and the Late-Binding™ architecture has a proven track record of agility and adaptability to new rules, vocabularies, and data content that other designs have not matched.

**Data Modeling and Data Binding**

Relational data models are inherently binding — they bind data to business rules and relationships. When developing transaction-based applications for capturing data, a data model is an important aspect of the application design and data integrity strategy. When designing a data warehouse, data models can inhibit adaptability to new analytic use cases. The following are the current options for modeling data in a data warehouse, listed in order of progression, from early to late binding.

1. **Early-binding: Inmon, Kimball, and I2B2.** The Inmon, Kimball, and I2B2 approaches to data modeling are inherently early binding. They require all source system data to be mapped into predefined data models, a process called conformance and normalization. The terms imply exactly what is required — data that was modeled and captured in disparate source transaction systems must conform to a new data model in the data warehouse. While at first this approach might appear reasonable, in practice, it leads to major problems when applied to the healthcare industry.

   Here’s why: in analytic environments where data content, use cases, data rules, and vocabulary change infrequently — for example, the retail industry, where the data model is largely reflected in the simplicity of a transaction receipt — the Kimball and Inmon approaches are adequate. In the healthcare industry where the data environment is much more complicated than a sales receipt and the analytic use cases are constantly changing, the consequences of these early-binding data models can be disastrous to agility and initial time-to-value. The process of mapping and conforming data to early-binding models in a healthcare delivery data warehouse typically takes 18 to 24 months or longer. When new data sources are added to the data warehouse, as occurs in mergers, acquisitions, and ACO partnerships, this lengthy time-to-value is repeated again and again. Likewise, as the complexity of analytic use cases inevitably matures in an organization, the early-binding data
model must be modified and the source system data must be conformed and mapped again. Early-binding data models cannot keep pace with the changes in the analytic environment, and the data warehouse subsequently fails to deliver its initial appeal.

2. **No-binding: Hadoop, MapReduce, PIG, and SoSQL.** MapFile structure association, popularized first by IBM mainframes some 60 years ago, is reappearing in the form of Hadoop, MapReduce, PIG, and NoSQL. Data warehouses based upon this technology exhibit the lowest degree of data binding and coupling. In fact, since there is no data model in this methodology, there is no data binding until the binding is declared between the files of data through associative programming. These data warehouses are very adept at quickly loading data into the warehouse, but the benefit of time gained in loading is more than negated by the complexity of the programming required to minimally bind and declare associations between the files and data.

3. **Late-Binding™:** The Late-Binding™ Data Warehouse is a balance between the extremes of early-binding found in Inmon and Kimball and the no-binding environment of Hadoop. The Late-Binding™ Data Warehouse emphasizes the following fundamental principles related to data modeling:

› The key to success for data warehouses is relating data, not modeling data. When in doubt, model less, not more.

› Minimize the use of new conformed data models in the data warehouse by instead leveraging the data models used in the source systems.

› Apply data models to subsets of data — in data marts — when binding formerly disparate data in new contexts to support new analytic use cases.

› Approximately 20 core data elements are fundamental to almost all analytic use cases in the healthcare industry. Early binding to these core data elements is a best practice. Bind to other terms and vocabularies later and only when required by analytic use cases.

The core data elements are shown below, illustrating how those data elements leverage the data models of the source systems to act as a data bus for the Health Catalyst's Late-Binding™ Data Warehouse platform. This approach allows queries across disparate source system content in the data warehouse in exactly the same fashion as the theoretical benefits of an enterprise data model but does not require development of, and conformance to, an enterprise data model. The diagram also illustrates how the Health Catalyst analytics platform can easily feed data to non-Health Catalyst applications.
Late Binding to Other Vocabulary Terms and Rules

As organizations progress in analytic maturity and sophistication, the need to bind to new and more complex vocabularies and rules will follow. By focusing first on the core data elements and then binding to additional rules and vocabulary when a clear analytic use case requires it, data engineers can deliver rapid time-to-value initially as well as later, when adaptability to new analytic use cases arise. Some of the additional vocabularies that are typically appropriate for later binding in healthcare include LOINC, RxNorm, SNOMED, and HCPCS. Business and clinical rules about data are even more complex and volatile.

The Healthcare Analytics Adoption Model, below, illustrates the relationship between progressively higher levels of analytic capability and the need to bind to more complex rules and vocabularies. The important concept to know is: it’s best not to bind data to rules or vocabularies until the analytic use case requires it. Too often, data warehouse projects in healthcare attempt to bind data to rules and vocabularies in anticipation of functioning at Level 8 of this model, even when the organization is still operating at Level 0. It takes years to progress to Level 8, and during that time, rules and vocabularies in healthcare will undoubtedly change. As the old saying goes, don’t drive beyond your headlights. In healthcare, it’s a dangerous waste of resources and time to bind to rules and vocabularies.
that are far beyond the current analytic use cases of the organization. For more information about the healthcare analytics adoption model, visit www.healthcatalyst.com/healthcare-analytics-adoption-model.

Summary of Principles in the Health Catalyst Late-Binding™ Data Warehouse

Below is a summary of the principles that underlie the Health Catalyst approach to analytics. These principles have enabled data warehouses in the military, manufacturing, and healthcare industries to operate and adapt for over 20 years and have unparalleled track records for proven results.

1. Minimize remodeling data in the data warehouse until the analytic use case requires it. Leverage the natural data models of the source systems by reflecting much of the same data modeling in the data warehouse.

2. Delay binding to rules and vocabulary as long as possible until a clear analytic use case requires it.

3. Use earlier binding only when it’s appropriate — for business rules or vocabularies that change infrequently or that the organization wants to lock down for consistent analytics.

4. Bind late in the visualization layer, which is particularly effective for what-if scenario analysis.

5. Retain a record of the changes to vocabulary and rule bindings in the data models of the data warehouse. This will provide a self-contained configuration control history that can be invaluable for conducting retrospective analysis that feeds forecasting and predictive analytics.
A simple concept such as taking vital signs has a multitude of conditional data points associated with it. It’s too complex for a basic star schema to manage effectively.

Most discussions about data warehouse approaches in healthcare naturally lead to the question, “Why not just use a star schema or other early-binding approach?” After all, the star schema is one of the most commonly deployed data warehouse methodologies. Its simple, straightforward design is considered a best practice for a wide variety of industries, including manufacturing, retail, telecommunications, and financial services.
One of the most prominent differentiators of Health Catalyst is our approach to the healthcare enterprise data warehouse (EDW). The foundation of our technology is a Late-Binding™ Data Warehouse platform designed specifically for the analytics needs of healthcare providers.

So the short answer to the star-schema question revolves around the number of variables in healthcare data, the flexibility required to mine that data, and the expense/time to value of creating and delivering the EDW.

**Advantages of a Star Schema Data Warehouse in Other Industries**

A star schema works well when you are doing counts or aggregation of counts by various (but fixed) dimensions of data. For example, if you want to know how many pairs of a particular brand and style of blue women’s shoes in size 7 there are in the Northeast region in Q3, a star schema works well. It’s easy to provision a small data mart quickly to support the needs of different subject areas.

Another positive to a star schema is that, since analytics have been used in manufacturing, retail, and financial services for so many years, a wide range of tools have been optimized to leverage that model — to the extent that you almost feel forced to go that route. So why didn’t Health Catalyst follow that model?

The answer is simple: healthcare data is too complex and has too many variables to make a star schema (or any early-binding model) practical. Compare the shoe sales example to the data surrounding the vital signs of a patient. When you take blood pressure, you need to understand more than the reading. You also need to know:

- When the blood pressure was taken (date and time)
- What type of device was used
- Whether the patient was standing, sitting, or lying down
- Whether the reading was taken from the left arm, right arm, or some other location on the body

As a result, a simple concept such as taking vital signs has a multitude of conditional data points associated with it, making it too complex for a basic star schema to manage effectively with its one-to-many construct.

In a healthcare setting, you’re really dealing with many-to-many relationships. You typically will have multiple patients who have multiple encounters with multiple caregivers, multiple diagnoses, multiple procedures, and multiple results — and you have to resolve all of those relationships at the database level in order to draw useful analytics from it. To make that happen with a star
schema you can try using a helper table, but that’s a Band-Aid approach. It doesn’t get to the root of the issue.

**Star Schema vs. Late-Binding™ in Healthcare**

The better option is a Late-Binding™ Data Warehouse built on top of a relational database structure. Unlike a star schema, which requires significant transformation to move the data out of the source application and make it useful for analytics, a Late-Binding™ Data Warehouse allows you to get data out of the transactional system at the most detailed, lowest level of granularity with minimal transformation.

Much of the concern around when to bind the data is tied to accuracy, a critical element in any analytics program, but especially in healthcare where quality of life, and maybe even life itself, is on the line. The more transformation required to get the data from the source system to the data warehouse/analytical environment, the greater the chance of introducing errors into the data or covering up issues in the data-capture process. More transformation also means: it’s more expensive and time-consuming to map and change the data to get it into a usable format, which translates into very protracted project timelines.

With a star schema, you have to make binding decisions before you load the data into the target analytical repository. This often occurs before you really understand the binding rules you’ll need. In addition, you have to transform all the data at once because it looks nothing like the applications that were used to capture the data, as opposed to building the data warehouse incrementally. The more data you have the longer it takes. If you discover after a few weeks or months that your binding rules have to change, you need to re-map and transform the data again, slowing or even bringing your analytics efforts to a halt until the process is finished. So, not only is it more expensive and time-consuming up-front, but the costs mount over time because of these cascading changes.

In a Late-Binding™ approach, you can move the data from the transactional (source) system to the data warehouse with minimal transformation. The data is available and ready to use, but it hasn’t been committed to any particular relationships, so you have considerable flexibility in terms of what you can do with it. You can then transform and bind it only when and if you have an actual need for it — similar to the “just-in-time” approach in manufacturing.

For example, suppose you want to measure readmission rates for heart failure patients. You can build a data mart to support that particular set of analyses by binding and transforming only the data sets you need. Put another way, rather than binding the contents of an entire library, you can just transform and bind the biography section within it.
With a Late-Binding™ approach, you spend far less time up front, since you’re merely moving data to source marts to make it ready rather than actually binding it. Over time your costs are reduced because you’re only putting effort into binding the data you actually need. When you need to move the data from the source mart to the individual subject area data marts, you’re doing so within the boundaries of the EDW over which you have control, saving additional time.

Something else a Late-Binding™ approach brings is flexibility. Because it is built on top of a relational database, you can use many different commercially available tools to access it. Since you’re binding as late as possible, you also have greater flexibility if the data source or the rules change. The downstream changes are minimal and easily managed.

The net result is a faster time to value and a lower total cost of ownership.

**How a Late-Binding™ Approach Works with Population Health Management**

Here’s why the Late-Binding™ approach is so critical to healthcare organizations. Suppose you want to perform population health management for a cohort of patients who have diabetes and then build a diabetes patient registry or data mart. For these patients, you want to know their hemoglobin A1c (current and historical), their lipid profile (LDL, HDL, and triglyceride), foot and eye exam history, the claims they’ve submitted, their medication history, their BMI, and the history of their office visits.

With a star schema/early-binding model — built for counting and aggregation rules related to a fixed set of facts — it would be difficult to support all the variables and the parent/child relationships for that population. The more you want to know about how these facts relate to other facts, and how those facts relate to a third set of facts and the attributes around each of those facts, etc., the more complex everything becomes.

With a Late-Binding™ approach, you have the flexibility to include, exclude, and change the rules regarding these relationships quickly and easily, whenever it’s needed.

**Healthcare Data Warehouse: The Right Construct**

A star schema or other early-binding data warehouse makes sense and works well in many industries. Healthcare, however, isn’t one of them.

A Late-Binding™ data warehouse offers the flexibility to mine the vast number of variables and relationships in healthcare data effectively and leave room for the inevitable future changes. It saves time, lowers costs, and, most importantly, delivers the results that will help you mine your data in a way that improves patient outcomes.
Virtual data warehouses have been around for some time and, on paper, they sound great. But they can’t meet the analytics demands of today’s healthcare industry.

At Health Catalyst, we believe an enterprise data warehouse is the only viable solution for health systems and physician groups looking to use analytics to drive sustainable quality and cost improvements across an entire organization. It’s so integral to success that Level 1 of the Healthcare Analytics Adoption Model is implementing an EDW.
Because of its importance, a number of business intelligence (BI) tools and visualization solutions are being marketed as virtual data warehouses that offer quick analysis and flexible visualizations in user-friendly packaging. Promises of virtual data warehouses have been around for some time and, on paper, they sound great. The premise is that these systems forego the steps of building and loading data into an EDW by extracting data directly from the source transactional systems. Users query the BI tool, which quickly extracts the information from finance systems, EHRs, lab systems and other sources. Then, at the last-minute, the tool mashes the information together in the visualization layer to provide the user with the answers he or she was looking for.

But the laws of physics — CPUs, memory, and disc reading — mean these visualization tools can’t yet live up to their promises or replace an EDW. Here’s why.

**What Healthcare Business Intelligence Tools Do Very Well**

The core strength of front-end BI tools is visualizing data and exposing it to end users. First and foremost, these are reporting tools that capture a snapshot of information at a particular point in time and provide access to that information via easily digestible charts, graphs, and similar visualizations. BI tools may also offer a certain level of drilldown to the data itself. Additionally, because many BI vendors offer a cloud-based option, these tools make the visualizations available — securely — whenever users want to access them.

**Why You Still Need a Healthcare Enterprise Data Warehouse**

However, no matter how slick or accessible a visualization is, its effectiveness is limited if it’s not based on a centralized, robust data foundation. Unlike EDWs, which send all of the data down a single pipeline and put it into an EDW built specifically for analytics, BI tools pull information they need directly from their sources. While this may sound efficient initially, at closer inspection, it can lead to a number of shortcomings.

**Shortcoming #1: BI Tools Don’t Optimize Healthcare Data.** Anyone who has undertaken a data-warehousing project knows that aggregating data from source systems uncovers a slew of data-quality issues. In fact, optimizing data and exposing data-quality issues represents a significant chunk of the effort in the initial stages of an EDW project.

Say, for example, your query includes gestational age. But gestational age may be stored in 12 different locations within the same EHR and stored in a number of different formats. With an EDW, information from all of these source systems is copied directly into a repository so you can set up an agreed-upon and consistent definition for gestational age...
that everyone in the organization can use. This approach also exposes you to problems with the data so you can isolate them and work to correct the data at its source.

BI tools don’t make it quite so simple. Your query may pull details from three separate transactional systems, each of which includes gestational age. You, therefore, have to define which gestational age you’d like to use for the purpose of this query. And now you run the risk of the definition you chose being different from another department’s choice. The end result can be similar queries that aren’t based on the same data points, which leads to reporting discrepancies.

Additionally, with data being pulled directly from multiple source systems, isolating problems with the data and subsequently fixing the problem can be a challenge. The ideal is to have everyone working with the same set of definitions, regardless of query. That way all reports and analyses are made using consistent data.

Shortcoming #2: BI Tools Can’t Handle Large Amounts of Healthcare Data. BI tools rely on other systems to store the data, so every time a BI tool wants to access data, it goes straight to the transactional system that is the source of that data. But conducting analysis in a transaction system leads to its own problems: these systems are built to store transaction details, one person at a time. When a BI tool tries to conduct analysis in that same environment, it can result in a huge drain on the system, which is felt by everyone, particularly the people on the front lines trying to use the transactional system.

Consider this: one patient encounter can generate hundreds of rows of data associated with billing systems, EHRs, and a variety of other sources. Now, say you have a query that addresses thousands of patient visits. With a BI tool, your query will be accessing millions and billions of rows of data in various source systems. For a small, independent hospital, this may be possible. But running this type of query directly against the transaction systems in a medium-sized hospital or even a health system with two small hospitals doesn’t scale. The sheer volume of information a BI tool is sorting through makes the process inefficient. Reports are slow to generate, and, worse, people who need to input information into the transactional systems are also faced with delays and inefficiencies. Multiple queries can generate more problems: each query goes straight to the sources, which results in a web of redundant data streams — another drain on the system.

An EDW, on the other hand, is optimized for analytics. It is designed to facilitate the analyses across entire populations of patients or events. Apps are created that call on the information in the data warehouse rather than going directly to the transactional system. Queries on the
EDW are run against the information stored there, which eliminates redundant feeds and ensures the efficiency of transactional systems isn’t adversely affected.

Shortcoming #3: BI Tools Don’t Work Well with Healthcare Data at Different Levels of Granularity. Many BI tools cannot handle data at different levels of granularity. In data warehousing, granularity refers to the level of detail stored in a database and how that level relates to other data. For example, one database list might store patients; another list might store individual patient encounters, which is a finer grain. The lists have a one-to-many relationship (since one patient can have many encounters).

Aggregating and digesting one-to-many and many-to-many relationships in the data requires a sophisticated system, especially when millions of rows of data are involved. How well your data warehouse handles granularity is very important for analysis and system performance. While some BI tools are adept at displaying data with different grains in intuitive and easy-to-use ways, other BI tools have a very difficult time — or simply can’t display data with varying levels of granularity at all.

Shortcoming #4: BI Tools Can’t Optimize Healthcare Data for Multiple User Types. How many departments and people could benefit from data if it were readily available to them? Data architects, executives, nursing managers, clinicians, the person responsible for regulatory reporting … the list goes on and on. And each person is looking for different insight from the same data.

Applying logic against the data so it is understandable at multiple levels for different audiences is something BI tools simply cannot do. BI tools focus on providing answers to specified sets of questions; an EDW can allow those questions to be customized by the audience. With a true EDW, you can build subject-specific data marts to answer specific questions — questions that originate with the audience and are customized and standardized to meet the audience’s unique needs. In the process of creating these data marts, you engage the different audiences, which helps them understand the data — and even their own processes — better, too.

Shortcoming #5: BI Tools Don’t Provide for Modularity, Understandability, and Code Reuse. A BI tool limits the ability to reuse code and logic. One data architect might maintain the code/logic for six months, and then another might need to learn it and further maintain it for the next six months. On the other hand, a well-designed data warehouse includes sets of logic that work well in stand-alone sets (e.g. two or three files of code/logic that, combined, show a population
with hospital-acquired pneumonia and specific metrics for that population, such as readmission rate).

Since the logic code is stand-alone (and not bound to the particular BI tool), they can be more easily and inexpensively reused elsewhere in the organization at a relatively low incremental cost, without having to keep the logic in distributed desktop files. The data warehouse not only stores a central repository of data, but it also stores centralized logic. There are two further benefits to this approach: understandability of the code and transfer of knowledge. Data architects need to understand why the code was written a certain way to understand how to maintain a particular data mart or to reuse the logic in another data mart. Centralized logic and code support this goal.

In short, data-driven healthcare transformation requires an EDW. While BI tools remove the steps of building and loading data into a data warehouse, this shortcut comes at a very big price. BI tools cannot provide the consistency, efficiency, or meet the analytics demands of today’s healthcare industry.
How can you ensure your healthcare data warehouse project succeeds?

Data warehouses are the answer for healthcare. But they, too, can come with risks. It’s no secret that the failure rate of data warehouses across all industries is high — Gartner once estimated as many as 50 percent of data warehouse projects would have only limited acceptance or fail entirely. When a project fails, a lot of money may be spent on something that’s never used or never even gets launched.
How can you ensure your healthcare data warehouse project succeeds? Going into a project knowing the potential pitfalls is critical. Pay careful attention to these six missteps many failed projects have in common.

**Six Common Missteps That Can Cause a Healthcare Data Warehouse to Fail**

1. **A Solid Business Imperative Is Missing.**

   The build-it-and-they-will-come attitude to data warehouse creation is sure to sink a project. Success requires a clear understanding of your financial and clinical needs and how you expect a data warehouse to address them — defined with input from a team of business and clinical staff who will be using the information every day to make decisions.

2. **Executive Sponsorship and Engagement Is Weak or Non-Existent.**

   It’s not enough for executives to green light a project and tell IT to let them know when it’s done. A governing body comprised of senior leaders from various operating areas of the organization including clinical, finance, quality, strategy, and administration should steer the project from day one through deployment. The governing body should be fully engaged in approving the budget, technical team, vendor partner, and priorities.

3. **Front-Line Healthcare Information Users Are Not Involved from Start to Finish.**

   An organization can have a strong business case and executive sponsorship but still fail if it doesn’t involve clinicians and other users on the front lines of care when devising its data warehouse strategy. Who better understands information needs and clinical processes than the frontline information consumer? Who will actually put the insights from the data into action? Multidisciplinary, frontline teams — comprised of clinicians, technologists, analysts, and quality personnel — should be involved throughout the design and deployment process to ensure the solutions meet their tactical and strategic requirements.

4. **Boil-the-Ocean Syndrome Takes Over.**

   Trying to be all things to all people spells disaster. Value can’t be delivered quickly if you attempt to “boil the ocean” and identify every possible piece of data just in case it might be needed at some time in the future. Take an iterative approach instead. Choose one clinical area at a time to focus on and gradually improve quality and cost.
5. The Ideal Trumps Reality.

Too often organizations fail to consider what information they are currently capable of capturing through their financial, EHR, and other systems. They take a bottom-up approach to analytics planning, starting from scratch to come up with every possible data need the organization might ever encounter but with no consideration for what will be involved in obtaining the data. If they don’t balance the ideal with the reality, meaning the data that’s on hand, the organization can put itself on the road to failure. Adopting a pragmatic approach of working with the data that’s already available to them is more effective and feasible.

6. Worrying about Getting Governance “Perfect” Immobilizes the Project.

Governance has two dimensions: priority governance, which is the process of deciding what data should be made available first; and data governance, which establishes policies to ensure consistent data quality and definitions. When it comes to governance, too often “perfect” is the enemy of progress — attempting to have all decisions wrapped up before starting can immobilize a project. It’s more efficient to put governance further down the development path. Once you have data to govern and react to, the data governance and prioritization process will accelerate.
A phased approach enables you to evaluate improvements in financial performance and care quality before investing further in the enterprise data warehouse platform.

You’ve decided to implement an EDW and you know what to watch out for. So what can you do to ensure everything runs smoothly when the time comes to implement the data warehouse?

CEOs and CIOs rightly press for the lowest-risk, most economical plan to invest in a healthcare EDW. The best plan is to structure the investment in manageable, bite-sized chunks. Each chunk of investment should be associated with a tangible benefit or ROI. An initial investment could be tied to a projected tangible benefit or ROI within a finite period, perhaps six to nine
months. If the first project demonstrates a return after that time period, move forward with the next investment decision.

Several veterans of the healthcare data warehouse space have advocated this approach. CIOs and CEOs should be pressing vendors to show a return within a finite period and scope levels of investment so they can see tangible progress before investing further in the solution.

**Investing in Healthcare Analytics in Four Stages**

Ideally, investment in a data warehousing initiative should occur in four phases, which include implementing both the enterprise data warehouse platform itself and analytic applications that run on the platform. Each stage delivers a measurable, tangible benefit or ROI within a finite time period.

**Phase 1: Implementing the Enterprise Data Warehouse Platform in Healthcare**

The first phase implements the enterprise data warehouse platform itself. By aggregating data from disparate sources into a single data warehouse platform, organizations establish a foundation on which to build all future analytics initiatives. This stage is implemented in three to nine months; it can be completed within a single budget cycle.

The ROI for this stage is admittedly less tangible than the others, but there are clear, concrete benefits. After this phase, the data warehouse is up and running and users enjoy a new level of visibility into what is actually happening clinically, financially, and operationally. This highlights the role of the data warehouse platform in informing an organization’s focus, in realizing measurable care improvement, and in significantly and systematically improving efficiency.

Importantly, the cost of this first phase is less than a quarter of what will likely be invested in the total solution. Significantly, adopting a phased approach reduces upfront investment by 75 percent and allows organizations to get a sense of the potential return before committing additional dollars.

**Phase 2: Creating an Efficient Reporting and Distribution System for Healthcare Data**

Foundational Applications automate data distribution and data provisioning — the process of providing users with access to data. These applications put in place an efficient system of reporting and distributing data both internally and externally and become the analytics base for an organization. They also provide dashboards and basic registries for a broad range of clinical and operational conditions. With Foundational Applications, for example, users would be able to
see every population and every care process family’s performance across more than 100 clinical, financial, regulatory and operational metrics. Foundational Applications provide information about trends and patterns, although they are not designed to identify the root cause of these patterns.

Installing Foundational Applications begins with a site assessment in which opportunities are identified based on their alignment with the foundational apps. It’s equally as important to identify the application users, provide training, and assist in the data interpretation as needed.

**Phase 3: Mining the Healthcare Data Warehouse to Prioritize Improvement Initiatives**

Once the infrastructure for provisioning and distributing data is in place, an organization is ready to mine the EDW’s data to prioritize improvement efforts. This is where Discovery Applications come in. They allow users to prioritize analytics and improvement initiatives by helping organizations discover more detailed patterns and trends within the data. Once trends are noted, Discovery Applications help users pinpoint key areas and processes to focus on to reach cost and quality goals. For example, a Discovery Application might analyze data to determine which clinical care processes represent the greatest opportunities for an organization. Identifying processes that can yield the highest return is one of the greatest benefits of an EDW.

Discovery Applications also help an organization prioritize by identifying areas of variation and waste, discovering new cohorts, stratifying populations, pinpointing opportunities by payer mix, and analyzing readmission risk, among other things.

**Phase 4: Implementing and Sustaining Healthcare and Improvements**

Advanced Applications allow users to implement, track, measure, and sustain improvement initiatives identified as priorities. These applications might target a clinical process, such as a disease condition or procedure, or an operational support service, such as operating room workflow or pharmacy. Included in this stage are the data mart, applications, and visualizations specific to the clinical process or support service. Note: this stage can be rolled out incrementally, one advanced application at a time. Measurable returns can be realized within months of installing an Advanced Application.

Teams from within the organization assist with the success and customization by identifying Aim Statements, cohort definitions, and data metrics to be included in the visualization. Work group teams also identify improvement opportunities for either processes or outcomes. Guidance/leadership teams provide direction so projects align with the strategy and goals of the
organization, and a broad implementation team supports the implementation or improvement processes identified by the workgroup team.

This phased approach enables you to evaluate quantifiable improvements in financial performance and care quality before investing further in the enterprise data warehouse platform.
It’s important to consider total cost as well as how you want to use your internal resources when developing your healthcare data warehouse.

Should you build your EDW or should you buy one? There are reasons this is such a common question in healthcare. Some CIOs see vendors as potential partners who can help achieve success and mitigate project risk; others perceive vendors as intrinsically evil. People may feel that internal development teams fail to deliver on time and/or they deliver over-budget, while others have had tremendous success and consider software and BI development a core competency of their IT organization.
Historically, Why Were Healthcare Data Warehouses Built, Not Bought?

To give some context to this question, it may help to draw a parallel with the evolution of the electronic medical record (EMR). Early adopters of EMRs had to build their own, as there were few commercial vendor options. Then, the first generation of rudimentary commercial EMRs were launched. Second and third generation systems were subsequently developed by companies like Siemens, Cerner, and Epic and prompted by the EMR adoption model from HIMSS that gave vendors and customers a benchmark — a roadmap — for product development and acquisition.

Similarly, until recently no single vendor offered an enterprise data warehouse (EDW) solution for healthcare that could deliver quantifiable results and return on investment. Innovative organizations with the resources to support an extensive internal development effort had only one potential path: build it themselves.

Health systems without the staff, budget, or experience to build a centralized EDW were left with two main options:

- Data analyst heroism, where a small number of savvy analysts used whatever reporting or analysis tools they had at their disposal. Great people in these roles achieved excellent results, but their potential value was often underutilized because the analysts needed to spend too much time extracting data instead of analyzing it.

- Adopt best-of-breed analytics solutions to help address specific, siloed reporting and analytics needs.

The Case for Building a Healthcare Data Warehouse

While today’s healthcare analytics market has matured sufficiently to provide compelling “buy” options for companies, there are still plenty of organizations that prefer to build. And while the build vs. buy dilemma can escalate into a heated, political battle that divides the organization (typically IT vs. everyone else) in unhealthy ways, there are compelling reasons to build an EDW and equally compelling reasons not to.

Pros of Building Your Own EDW:

- **The Possibility of a Perfect Fit.** Custom development, in the hands of skilled software engineers, can yield excellent, tailored solutions.

- **Potentially Lower Initial Cost.** Organizations with a few cycles to spare can often get something in the way of healthcare analytics up very rapidly. When finances are a factor, “something” is often synonymous with “good enough for now.”
**Pride of Ownership.** Several healthcare organizations, including members of the [Healthcare Data Warehousing Association](https://www.healthdatawarehouse.org) have built successful data warehousing programs from scratch, starting in the late 1990s. Many of those organizations are now among a short list of organizations that are well positioned to address the needs of healthcare reform. Viewed by other systems nationwide as now possessing vital knowledge assets, they are justifiably proud of their accomplishments, vision, and forethought. However, several of those early pioneers are also realizing that their homegrown solutions are neither sustainable nor adaptable to new analytic use cases in the industry. The perceived value of these first generation EDWs is in decline.

**Cons of Building Your Own EDW:**

**Impact on Staffing.** Building an EDW requires software engineers who can be dedicated solely to the data warehouse build for many months if not years. Internal data warehouse projects are notoriously under-scoped, under-resourced, and delivered much later than planned. There is also an enormous shortage of experienced data and software engineers in healthcare and across all industries. This explains why, across all industries, 60 percent of internally developed EDW projects fail to meet expectations. Still, healthcare organizations with well-staffed and very agile IT governance may be able to make the build happen smoothly.

**Learning from Mistakes.** It’s good to learn from experience but when you’re working on a project as big as an EDW build, it’s often a one-time shot. So, from a technical perspective, one of the most significant cons to building your own EDW is that you’ll be learning lessons (i.e., making mistakes) that have previously been addressed by an experienced vendor. Without a starting point, including coding standards, naming conventions, and a proven approach to data architecture, developers often find themselves refactoring or re-engineering key parts of the data warehousing solution.

**Silos and Repeated Efforts.** Many internally developed data warehouse projects start under the radar and in response to a lack of agility in IT to meet the reporting and analytic needs of the organization. As such, their success can be culturally divisive and create silos of data that need to eventually be re-incorporated into an enterprise-wide strategy for data management.

**Inability to Adapt.** IT teams accustomed to strict waterfall project management approaches may also lack the agility necessary to adapt to healthcare’s rapidly changing vocabularies, standards, and analytics use cases. If a project is unable to deliver continuing value, this may
should we build or buy our data warehouse?

lead to delays, unmet expectations, and overall frustration with the resources that are being funneled into the effort.

the case for buying a healthcare data warehouse

as the healthcare analytics market has matured, buying a data warehouse solution for healthcare has become a viable option. the following are some important things to consider before choosing this option.

pros of buying an EDW:

- **Shortest Time to Greatest Value.** With the right technology, including adapters and accelerators for common source systems, the implementation of a healthcare data warehouse can now be accomplished in as little as 90 days.

- **Vendors with Experience.** The right vendor will have seen many different design approaches in practice, and they will know what works and what doesn't. Savvy health systems realize they benefit from a top-notch vendor’s real-world experience and prior investments in product development.

- **Project Resources.** Most IT departments today work hard to deliver on existing project commitments. New software development projects have to compete with prior commitments, such as ICD-10 or Meaningful Use, and may struggle to even get launched. Starting a project with an experienced vendor brings with it an infusion of resources to help get the work done when you can’t spare anyone else.

- **Lower Risk and Possible Customization.** As some of the best health systems in the country have learned, risks associated with the success of a data warehouse are often mitigated by working with a vendor. Additionally, vendors can often help build a tool that more accurately fulfills its desired purpose through custom development, which may also lead to faster internal adoption and the delivery of success earlier.

- **Lower Total Cost of Ownership.** While the initial price tag of a third-party software solution can seem high, the three- to five-year total cost of ownership is often much less than an internally developed EDW.

cons of buying an EDW:

- **Knowledge Transfer.** In a truly hands-off data warehouse implementation, there is a risk that when the vendor leaves, insufficient knowledge will have been transferred to the team that’s expected to provide operational support.
Tradeoffs. As with any enterprise software purchase, expect to encounter some tradeoffs between the absolute perfect solution for your environment and one that is very good. Health systems are already making these types of tradeoffs when they choose to configure off-the-shelf EMRs and adapt their processes rather than build an EMR outright from scratch to their exact specifications.

Not Enlisting Internal Resources. If your organization is fortunate enough to have some innovative in-house software developers, and you and/or your vendor don’t engage them in the data warehouse implementation, their talents could go under-utilized. Combined with the right knowledge transfer, these developers are often the people who help deliver unexpected future successes with your EDW.

Risk Factor. There is some risk associated with engaging any new vendor, or with engaging a new division or group within an existing vendor relationship. You are counting on the vendor to help your organization achieve its goals, but the fact that you haven’t seen that company succeed within your environment requires trust. Make sure you know what to ask your potential partner to ensure success.

Your Third Option: Buy AND Build a Healthcare Data Warehouse

Healthcare providers now have several commercial options for various types of healthcare analytics. But organizations that have been able to nurture one or more teams of effective, internal software developers still want to preserve this precious resource and deploy them strategically for a competitive advantage. This is leading to a build AND buy scenario — where an organization enlists a vendor to accelerate its implementation of analytics while also using internal developers to help them achieve even more, faster.

Here are some key points to consider when thinking about the buy AND build option:

Pros of Buying AND Building an EDW:

- **Rapid Implementation Time.** The right vendor will be able to get most of the data warehouse implemented quickly, allowing your IT leads to demonstrate early successes and keep project momentum high.

- **A Tailored Fit.** Internal software developers know your systems inside and out. A savvy vendor will include and empower your top IT performers, ensuring a smooth implementation and honing in on ways to meet immediate analytic needs.

- **Lower Overall Project Risk.** By choosing to leverage aspects of both the buy and the build strategies, you are positioning yourself for a successful project in the following ways:
› You are engaging some of your most valuable IT employees early in the process.

› Your vendor is contractually obligated to deliver on the agreed-upon terms and statement(s) of work.

› You’re giving yourself additional opportunity to innovate — a great EDW vendor will partner with you to empower your internal developers through access to integrated data and easily accessible metadata. With the plumbing taken care of by your vendor, your in-house engineers can focus on extracting even more value from your investment, developing their own predictive models, connecting to operational systems and bi-directional interfaces with other third party systems, and more.

Many EMRs are now providing interfaces to consume relevant external data, such as patient-level risk scores, to drive best practice alerts at the point of care. With access to the data in an enterprise data warehouse, a talented internal software developer could deliver a prototype solution to a similar use case in days rather than months or years.

Cons of Building and Buying an EDW:

- **Requires Existing Data-Driven Culture.** This approach is best suited to a data-driven culture that values analytics as a business differentiator. Organizations with a commitment to a higher degree of data literacy and data management skills are very successful with this type of data warehouse development.

- **Higher Cost of Ownership.** Taking a build-and-buy route results in a slightly higher total cost of ownership than the purely “buy” option. However, this increased expense is offset by the higher return on investment, which can be achieved through the optimal utilization of the EDW.

**Selecting the Right Healthcare Data Warehouse Approach**

Regardless of whether you choose to build, buy, or do a little of both, it’s important to consider both the total cost and how you want to engage your internal resources when developing your healthcare data warehouse. Think about how fast you want to realize value in the short term and how you envision developing your own analytic applications in the future. The early adopters of EMRs mentioned were also some of the first to build their own EDWs. Several of those early pioneers are now applying their experience with home-grown EDWs to the purchase and installation of hybrid solutions — commercial vendor EDWs that come pre-configured for rapid deployment.
and quick value, but that can also evolve and be maintained by local IT organizations if desired.

For many organizations, the “hybrid” option of buy AND build provides a way to achieve the most value, while also mitigating many of the risks associated with large in-house software development efforts. You get the flexibility and empowerment of building a system on your own but without the risks, and you get the benefits of a commercially supportable solution.
Embarking on an assessment with the knowledge of key, general criteria can help you determine whether a vendor has the philosophy, experience, and viability that can lead to a successful outcome for your organization.

Every healthcare analytics company claims it will help providers use data to improve care and lower costs. But, as a healthcare provider looking to invest in an EDW, it’s important to carefully assess vendors and solutions up front to ensure you’re getting what’s best for you and your organization.
General Criteria for Assessing a Clinical Analytics Company

Embarking on an assessment with the knowledge of key, general criteria can help you determine whether a vendor has the philosophy, experience and viability that can lead to a successful outcome for your organization. By evaluating vendors according to the following, you can narrow down the list considerably.

1. Completeness of Vision

What lessons does the vendor bring from the past healthcare analytics market and how have they adjusted their current strategy and products accordingly? Can they bring lessons from other industries that are more advanced in their adoption of analytics? What is the vendor’s understanding of the present market and industry requirements? What is the vendor’s vision of the future for healthcare analytics? Look for vendors who can clearly outline how they have evolved to meet — and anticipate — industry needs.

2. Culture and Values of Senior Leadership

It’s no cliché, but rather the precise truth: the overall culture of a company starts at the top. Get to know the senior leadership of the vendors you are evaluating. Insist on meeting several members of their executive team. When interacting with individual members of the vendor’s team, ask yourself, “Would I be excited to hire this person into our company?” If the answer is consistently “no,” it may be a good idea to look at other vendors. More than technology is required to leverage analytics to drive real, sustainable change. Cultural transformation will be required throughout your organization. If your culture and values don’t mesh with your vendor’s, you may encounter significant roadblocks to success.

3. Ability to Execute

Does the vendor have a track record for delivering value and satisfaction to their clients? Review analyses by independent firms including KLAS, Gartner, Chilmark, and the Advisory Board. Also contact three to five of their referenced accounts directly, preferably ones that are similar in size and demographics to your own organization — and shy away from referenced accounts that have been pre-screened and selected by the vendor. Ask these references very simply: How satisfied are you with the vendor’s products, services, and overall value? If you had to do it all over again, would you?

4. Technology Adaptability and Supportability

The reality is, in today’s connected world, all businesses, including healthcare, move at the speed that their software can adapt — either fast or slow — to new processes and business models. Therefore, the underlying engineering and architecture of a vendor’s software is critically important.
You must peel back the covers of the vendor’s products and evaluate their software engineering for modern design patterns like object-oriented programming, service-oriented architectures, loose coupling, late binding, and balanced granularity of software services.

Glossing over this assessment is akin to buying a multimillion-dollar office building without assessing the modularity of the walls and soundness of the foundation, plumbing and electrical systems. How fast can the system adapt to the market and your unique needs for differentiation? Data standards, vocabularies and analytics use cases are changing rapidly in the healthcare industry, literally every day, with no signs of slowing down. You want a vendor whose software engineering can keep up. Analytic agility is critical because executives in your organization can’t wait weeks or months anymore for a new report to inform a critical decision. The industry is changing too fast.

5. Total Cost of Ownership

The best solution in the world is of no value if it’s not affordable. To assess affordability, you must understand the total cost of the vendor’s solution. Measuring total cost of ownership (TCO) is easy. You simply add up the three-year labor costs, licensing fees (including third-party), support fees, and hardware costs associated with a vendor’s solution. Many old-school analytics vendors require a significant upfront investment with no guarantee of value for two years or more. Your TCO over three years should be evenly distributed, not front-end loaded, and your contract should be structured with escape clauses if the vendor’s solution cannot prove value in the first year. In today’s market, clients should expect initial value from analytics vendors in less than six months, preferably only three. If a vendor cannot or will not commit in their contract to this timeframe for delivering value, look elsewhere.

6. Company Viability

Will the vendor be around in nine years (the average life span of a significant IT investment)? If not, can you live without them? Take advantage of evaluations by neutral third-party analysts like Gartner, Chilmark, KLAS, and The Advisory Board. What are these analysts saying about the vendor’s prospects in the market? Is the vendor in solid financial shape? What’s their monthly burn rate vs. income? How many days cash on hand do they maintain? What does their sales pipeline look like? Does its executive leadership team jump from one company to the next or does their track record indicate longevity and success? How much is the vendor spending on its sales staff relative to its engineering and product development staff? The best products are often supported by lean sales staffs. That’s because great products sell themselves.
Technology, a Key Consideration for Analytics Success

Technology is vital to the success of an analytics initiative. Here’s what to look for:

1. **Data Modeling and Analytic Logic**

   Different vendors’ analytics solutions feature different data models. Which data model they use can have a significant effect on the cost, scalability and — especially — the adaptability of your analytics solution to support new use cases. Rapidly adaptable and very flexible, a bus architecture is the best data-modeling option for healthcare. Most vendors utilize a healthcare-specific enterprise data model at the heart of their solution, but these enterprise data models are difficult to load and map initially, and slow to evolve subsequently, particularly when faced with new use cases and source system data content. These enterprise data models come in the following three basic flavors, so be aware of them:

   1. **Dimensional star schema**
   2. **Enterprise 1st, 2nd, 3rd normal form**
   3. **I2B2**

   Over-modeling data is the single most significant contributor to data warehouse and analytics failure in healthcare. Our advice is simple: stop modeling your data and start relating it. Relating data is what analytics is all about.

   In addition to the issue of data modeling, the analytic logic associated with the content of data marts and reporting is critically important. To learn more about the role of data modeling and “binding” data to business and clinical logic in healthcare analytics, read about the Late-Binding™ data warehouse in chapter 4.

2. **Master Reference/Master Data Management**

   The ability to incorporate data from new and disparate sources into your analytics solution requires significant expertise in master data management. What is the vendor’s strategy for managing unique patient and provider identifiers? How does the vendor accommodate international, national, regional, and local master data types and naming conventions? Do they support mappings to RxNorm, LOINC, SNOMED, ICD, CPT, and HCPCs? How tightly does the vendor bind your data to the vocabularies that change regularly? The tighter the binding, the less flexible the analytic design will be to accommodating changes in the vocabulary and analytic use cases based on those vocabularies.
3. **Metadata Repository**

An effective metadata repository is the single most important tool for the widespread utilization and democratization of data in an organization. Look for a vendor that provides a tightly integrated, affordable, simple repository with their overall analytics solution. The most valuable content in a metadata repository is not computable — the most valuable content is subjective data that comes from the data stewards and analysts who have interacted most with the data. Look for vendors that have the ability to maintain this subjective data through a wiki-style, wisdom-of-crowds contribution model.

A web-searchable metadata repository should provide information such as the source of the data, how often it is updated, examples of the data, natural language descriptions of the physical data tables and columns, any known data quality issues, and the contact information for the associated data steward. The ability to quickly establish the origins and lineage of data in a data warehouse is also a critical component to an effective repository.

Data warehouse and analytic vendors tend to operate in one of two extremes:

1. They either oversell very complicated and expensive metadata repositories that require an overwhelming level of support and maintenance in return for a declining return on investment; or

2. They offer no solution for metadata management, which is disastrous to a long-term analytic strategy.

Your goal should be to find a vendor that offers a simple, low-cost, pragmatic solution between these extremes.

4. **Managing “White Space” Data**

Does your data warehouse and analytics solution offer a data-collection alternative to the proliferation of desktop spreadsheets and databases that contain analytically important data?

White space data is the data collected and stored in desktop spreadsheets and databases that is not being collected and managed in primary source systems, especially EMRs, or it is being collected in clinical notes and must be manually abstracted for reporting and analysis. This desktop data fills in the missing “white space” of analytic information that is important to the organization. For example, these desktop data sets are commonly found in support of Joint Commission reporting, internal KPIs, finance analytics, and clinical researchers. It is not unusual for healthcare organizations to have hundreds of
these desktop data sources that are critically important to the analytic success of the organization.

However, because the data resides on desktop computers and shared drives, it cannot be integrated with other mission-critical analytic data being stored in the enterprise data warehouse from the primary source systems. Data synergy suffers as a result.

White-space data also poses information security risks. Data warehouse and analytics vendors must provide a tool for attracting the management of white space data into the content of the EDW. Look for a white space data management tool that is web-based, as easy to use as a spreadsheet or desktop database for the collection of data, and that makes it easy for end users to convert and upload their existing desktop data sets. Also, look for a security model in the EDW that allows for the isolation and stewardship of these white space data sets.

5. Visualization Layer

The best data warehouse and analytics solutions include a bundled visualization tool — one that is both affordable and extensible if licensed for the entire organization.

However, the analytics visualization layer is very volatile. The leading visualization solution today may not be the leader tomorrow. Therefore, look for a data warehouse and analytics vendor that can quickly and easily decouple the underlying data model and data content in the data warehouse from the visualization layer and swap the visualization tool with a better alternative when necessary.

Also realize: a single visualization tool will not solve all of your organization’s needs. Data analysts will want to use a variety of tools to access and manipulate data in the enterprise data warehouse. The underlying data models in the data warehouse must be capable of supporting multiple visualization tools at the same time. Ask vendors if their data model is decoupled from the visualization tool. Does the data model support multiple visualization tools and delivery of data content?

6. Security

As always in healthcare IT, the privacy and security of patient data is paramount. Ask these important questions of a potential analytics vendor about security:

- Are there fewer than 20 roles in the initial deployment? Contrary to popular belief, more roles can actually lead to lower security and will definitely lead to higher overhead administrative support costs.
Does the solution employ database-level security, visualization-layer security or some combination of both? The vendor’s solution should support both.

What is the vendor’s model for protecting patient-identifiable [protected health information (PHI)] data and the more sensitive subsets of PHI that are typically defined at the local state level, such as mental health data, HIV data, and genomic/familial data?

What type of tools and reports are available for managing security and auditing access to patient identifiable data?

7. Extract, Transform, Load (ETL)

A robust ETL process — how analytics technology extracts data from source systems, applies the required transformations, and writes data into the target database — is fundamental to the success of your chosen solution. Ask vendors to demonstrate how their ETL measures up in terms of reliability, supportability, and reuse. At present, Microsoft’s ETL tool — SQL Server Integration Services (SSIS) — is by far the most cost-effective ETL tool in the market, offering the highest value per dollar.

8. Performance and Utilization Metrics

As you implement and continue to use an analytics solution, you will need to generate metrics about who is using the system, how they are using it, and how well the system operates. Can the vendors’ solution track basic data about the environment, such as user-access patterns, query response times, data-access patterns, volumes of data and data objects? This kind of information will be essential to you as you refine and organize the data content and analytics services you provide from the data warehouse.

9. Hardware and Software Infrastructure

Does the vendor use Oracle, Microsoft, or IBM for its hardware and software infrastructure? These three are the only viable options in today’s healthcare market and data ecosystem. Hadoop and its associated open source tools is not an appropriate analytic and data warehousing infrastructure option at this time in healthcare (except for genomics).

Microsoft is the most integrated, affordable, and easiest to manage of these technology platforms and now makes up 70 percent of all new sales in the analytics and data warehousing market, across all industries, far outpacing Oracle (its closest competitor) in new sales. Microsoft’s parallel data warehouse platform can scale to the petabyte
level, far beyond the largest data warehouse needs in the healthcare provider space. Ask any data engineer who has worked extensively on either Microsoft or Oracle, which platform he or she believes is the easiest to use and most efficient for delivering quick, adaptable analytics solutions. The answer will almost certainly be Microsoft.

Cultural Change Management

A vendor’s solution must also include processes and real-world experience for helping manage sustainable change driven by analytics in your organization. Sustainable cultural changes are required to turn the insights from data into improvements in patient care and reductions in cost. But nothing is more politically or culturally disruptive than the spotlight of analytics, not even the deployment of an EMR. Choose a vendor that has been in the trenches of cultural transformation driven by the enlightenment of data.

Your Analytics Roadmap: The Healthcare Analytics Adoption Model

Health Catalyst joined with other thought leaders in the analytics industry to lead the development of a Healthcare Analytics Adoption Model (for more details, see chapter 4). The model outlines eight levels of analytics adoption an organization passes through as it gains sophistication in using its data to drive improvement. Following this model with discipline will lead to the successful adoption of analytics in an organization, both culturally and technically. Also use this model to evaluate vendors’ capabilities in each level and have the vendor demonstrate its products and services for each level as well.

Consider the model as a roadmap for your organization to measure your progress of analytics adoption. Ask yourself, “How fast do we want to achieve the highest levels of adoption in this model?” With the right analytics vendor as a partner, organizations can achieve Level 5 within 18 months of implementing and following the model, and some organizations can make it in 12 months. Level 7 is easily achievable within 24 to 30 months.
We provide the tools healthcare organizations need to use their data effectively.

Health Catalyst is a healthcare data warehousing and analytics company founded by industry veterans. Health Catalyst solutions are designed to respond to the rapidly changing nature of today’s healthcare industry and adapt to our clients’ unique goals. Our guiding philosophy is to be systematic through the services, analytic applications, and Late-Binding™ Data Warehouse we provide, so healthcare organizations have the tools they need to effectively use their data to make well-informed improvements and decisions.
Key products: Late-Binding™ Data Warehouse, Foundational Applications, Discovery Applications, and Advanced Applications.

Late-Binding™ EDW: Agile and Iterative, Providing Quick Time to Value.

Our Late-Binding™ Data Warehouse is one of a kind in the healthcare industry and was built on the following guiding principals:

1. Minimize remodeling data in the data warehouse until the analytic use case requires it. Leverage the natural data models of the source systems by reflecting much of the same data modeling in the data warehouse.

2. Delay binding to rules and vocabulary as long as possible until a clear analytic use case requires it.

3. Earlier binding is appropriate for business rules or vocabularies that change infrequently or that the organization wants to lock down for consistent analytics.

4. Late binding in the visualization layer is appropriate for what-if scenario analysis.

5. Retain a record of the changes to vocabulary and rule bindings in the data models of the data warehouse to provide a self-contained configuration control history. This can be invaluable for conducting retrospective analysis that feeds forecasting and predictive analytics.

Components:

Metadata Engine

The Health Catalyst Metadata Engine powers the generation and automated loading of Source Marts into the Health Catalyst Platform.

Data Acquisition Engine

The Data Acquisition and Storage subsystem of the Health Catalyst platform supports the optimized extraction, transformation, and loading (aka storage) of data from a number of source systems into Source Marts. The Data
Acquisition and Storage subsystem manages all the Source Marts of the data warehouse. Tools include Source Mart Designer to facilitate the rapid creation of custom source marts. Also included is the Source Mart Engine, a metadata-driven ETL engine that manages the efficient extraction, transformation, and loading processes of the EDW.

Data Bus Architecture

Health Catalyst’s Late-Binding™ Data Bus Architecture allows binding to be delayed until needed, which allows data to remain in its original, undiluted, unaltered form.

**Foundational Applications:**

Foundational Applications automate data provisioning and distribution and enables broad use of the data warehouse by providing dashboards, reports, and basic registries across clinical and operational areas. These applications help clinical, financial, and operational teams understand the data relevant to their specific areas of focus. Each Foundational Application consists of a near real-time data mart and one or more analytical applications that provide advanced analytics and drill-down capabilities in an easy-to-use web and mobile-accessible format. Foundational Applications provide significant information, historic trends, and patterns to a broad audience across a health system and are built on a common flexible architecture allowing for expandability.

**Discovery Applications:** Discovery Applications allow users to discover patterns and trends in the data that inform prioritization, inspire new hypotheses, and define populations for management to help health systems and clinicians pinpoint the areas they should focus on to achieve operational, financial and quality goals. These applications inform organizational prioritization efforts based on variation, discovery of new cohorts, selection and stratification by comorbid condition, analysis of payer mix, and predication of readmission risk, among other things. Discovery Applications are built on a flexible architecture that allows for models, algorithms, rules, definitions, and hierarchies to be easily editable and expandable by clinicians.

**Advanced Applications:** Advanced Applications provide deep insights into evidence-based metrics that drive improvement in quality and cost reduction
through managing populations, workflows, and patient injury prevention. The technology component includes a data mart, applications, and visualizations specific to the clinical process. Health Catalyst provides a cross-functional team of clinical and technical resources that can help clients organize, implement, and deploy a clinical improvement program. The application allows the hospital's clinical improvement team to define both clinical and financial improvement objectives and customize the metrics and visualizations to best fit the processes of the institution.
Good data has always been at the heart of improving medicine. With the advent of digital health, though, the sheer volume of information has created a new challenge for healthcare leaders: how to transform terabytes of raw data into meaningful improvements in the quality and cost of care.

For a healthcare analytics strategy to succeed within an organization, it must start with an enterprise data warehouse. Many healthcare organizations expected the deployment of an EHR would be the answer. While it’s clearly a necessary step towards data-driven delivery of care, the EHR alone is not enough without an EDW that enables an enterprise-wide, consistent view of data from many sources.
The winners in U.S. healthcare won’t get there by accident. They’ll distance themselves from the also-rans by leveraging robust, comprehensive, and accessible data warehouse platforms to guide them to excellent performance in cost and quality in their delivery of care.

Reviewing the conceptual framework for adopting a healthcare EDW is one thing. But how does this concept stand up to the real world? The following case studies illustrate successful implementation and user adoption.

**Case Study 1: Getting an EDW Up and Running in 90 days at Indiana University**

**Background**

Good data has always been at the heart of improving medicine. With the advent of digital health, though, the sheer volume of information has created a new challenge for healthcare leaders: how to transform terabytes of raw data into meaningful improvements in the quality and cost of care.

IT leaders at Indiana University Health (IU Health) — an academic health system with 18 hospitals, 3,300 staffed beds, and more than 3,700 physicians and allied professionals — realized that fostering a genuinely data-driven culture meant pushing beyond the limits of IU Health’s existing technologies. Even powerful tools like IU Health’s Cerner electronic health record (EHR) needed to be augmented with new healthcare analytics and cost-accounting tools capable of connecting the dots between IU Health’s mountains of clinical data and the much-needed operational and financial insights it lacked.

What IU Health needed was an enterprise data warehouse.

**Traditional vs. Late-Binding Data Warehouses**

The problem was that traditional EDWs can take years to deploy across large enterprises, and IU Health is the largest health system in its state. Worse, through trial and error, IU Health discovered a critical flaw in “early-binding” data warehouse architectures commonly used in other industries. Early-binding EDWs impose business rules on data at the outset. But healthcare evolves, organizations evolve and rules

Health Catalyst guaranteed they could deliver a useful and useable enterprise data warehouse in 90 days and they delivered. To my knowledge, the scope of this accomplishment is unprecedented in the industry.

— Bill McConnell
CIO and Senior VP
Indiana University Health
change. The net effect is poor agility and lengthy time to value. Crucial near-real-time clinical data analysis is almost impossible.

“IU Health had previously struggled to create a data warehouse that would serve as the organization’s source of truth,” said Bill McConnell, IU Health’s Senior Vice President and CIO. Those early attempts ended with a few data stores each with a partial picture of the system.

Success in Just 90 Days

So IU Health turned to Health Catalyst to develop a platform based on its Late-Binding™ Data Warehouse architecture. The company offered IU Health an outstanding value proposition with a no-risk pledge. It promised to create a data warehouse and integrate IU Health’s Cerner EHR within 90 days. And it delivered.

Before even hitting the deadline, Health Catalyst successfully loaded 14 billion rows of data into the EDW — fully 10 years of clinical data from across IU Health’s network. Clinical events, encounters, lab and radiology, and other patient data were included, as were key pieces of IU’s performance management, revenue cycle, and patient satisfaction data.

Collaboration has been key to the EDW’s successful implementation. EHR vendor Cerner gave Health Catalyst rapid and complete access to its Kansas City-based database and provided ongoing support to ensure that Health Catalyst’s work would not tax IU Health’s system.

Three Months to an Executive Dashboard

Within 90 days of the EDW’s completion, IU Health put in place a newly developed interactive dashboard — effectively a window into the EDW. It provides IU Health’s leadership with the daily operational insights they need to solve the quality/cost equation. It offers visibility into key operational metrics and trends so that IU Health’s leadership can easily track the performance measures critical to controlling costs and maintaining quality.

IU Health’s operational dashboard went live with unprecedented speed and provides a near-real-time view into the data versus its previous manual reporting that was a 30-day retrospective.

The executive dashboard currently extracts data from the EHR and costing systems. IU Health plans to integrate additional source applications such as corporate billing and patient satisfaction in the near future. Health Catalyst’s agile architecture and methodologies can easily be applied across IU Health’s departments to analyze, track, and measure clinical, financial, and patient experience outcomes.
Expansion Plans

Ultimately, IU Health’s EDW will add Health Catalyst Advanced Applications that will enable deep-dive analyses into such granular topics as workflow, hospital-bed turnaround times, the success of various cardiac care techniques, and hospital-acquired infections.

The project is proving a critical first step toward IU Health’s goal of capturing and analyzing clinical data to improve the quality of patient care while reducing costs.

Case Study 2: Using the EDW Broadly Throughout the Health System: Texas Children’s Hospital

Texas Children’s Hospital is internationally renowned for caring for children in the United States and other countries. U.S. News & World Report ranked it the nation’s fourth best children’s hospital and the best in Texas in 2012. Founded in 1951, Texas Children’s resides in the largest medical complex in the world, the Texas Medical Center in Houston, and provides primary and tertiary care for children through its hospitals, affiliated practices, and health plan.

In 2006, with value-based reimbursement on the horizon, Texas Children’s began to examine its quality improvement program, carefully evaluating its data management capabilities. To succeed under a value-based system, the hospital’s leaders knew they needed the ability to analyze and better manage specific populations of patients, especially the most costly patients with chronic problems such as asthma. They also sought to identify areas of inefficiency and waste in their care programs but lacked the hard data to pinpoint the suspected problems and to uncover other hidden inefficiencies and safety issues.

To address this challenge, Texas Children’s launched an overall quality and safety strategy in 2006. The goal was to develop a comprehensive and integrated enterprise-wide data management infrastructure. The key foundational element involved implementation of an EHR to collect raw clinical and financial data from across its enterprise. This was a critical first step in the data management strategy to transform data

Before the data warehouse, if someone in the organization wanted data from IT to answer a question, they would submit a report and there was this back and forth that was so frustrating for the clinicians, who really didn’t want to have to communicate with us. Now with the data warehouse as a hub with near real-time data, the clinicians can all see the data at the same time and they can very easily spot the data they’re looking for.

— Myra Davis, MSE
Senior Vice President of Information Services
Texas Children’s Hospital
into meaningful information to guide clinical quality interventions and waste-reduction efforts.

Meeting Expectations

The clinicians expected that the EHR would readily provide data they could use for individual patients and populations of patients. “Our clinicians thought that the EHR was a silver bullet to get the data they needed [for quality improvement], and they blamed IT when the information wasn’t forthcoming,” recalls Myra Davis, MSE, Senior Vice President of Information Services for Texas Children’s Hospital. “The comment I would hear is, ‘I can’t get the right data from them,’ or ‘They don’t understand what I need from them.’ It created nothing but frustration.”

Leaders of quality, clinical, and IT departments at Texas Children’s knew the solution was to nurture a truly data-driven clinical culture at the hospital and develop an EDW to help meet the expectations of the clinicians.

To help nurture this new data-centric culture, Texas Children’s decided to roll out a distinctly different, clinically-driven EDW and new methodology to help it measure care and population health outcomes. Beginning in September 2011, the hospital worked with Health Catalyst to implement a clinical and analytic framework that included:

- Implementation of an adaptive data warehouse platform and advanced analytics to collect data from systems inside and outside of the enterprise. This allowed users to report on a variety of short-term operational and clinical metrics.

- Development of permanent, integrated teams of clinicians, technologists, analysts, and quality personnel to identify areas for improvement in care processes and build evidence-based care guidelines into the care delivery workflow.

- Implementation of a measurement system infrastructure to better track and interpret iterative improvement, a tactic that Texas Children’s found critical to sustain improvements.

Early, Iterative Wins Propel Adoption

Like many large health systems that have strong clinical leadership and priorities, Texas Children’s Hospital had historically been challenged to decide on care improvement opportunities that would have the greatest impact. The data warehouse and its data discovery application enabled Texas Children’s to prioritize its quality improvement programs. The application is based on a Pareto Analysis — a statistical technique used to make decisions by identifying the limited number of tasks that will produce the most significant overall effect. The exhaustive analysis of variation and
resources consumed led leadership to focus first on improving the quality of the hospital’s asthma program.

As a first step, a permanent cross-functional workgroup consisting of physicians, nurses, and experts in patient safety, quality improvement, finance, and IT was assigned to assess and manage acute asthma from the time of arrival in the ED to discharge. Improvement opportunities became evident from the very first meeting. The team identified that a higher-than-normal volume of chest X-rays was being administered to asthma patients. Rather than request an analyst’s report to explain the cause, as they would have in the past, the team immediately used the EDW’s dashboards to drill down into the near real-time X-ray data.

To their astonishment, the team recognized that, as a group, Texas Children’s Hospital physicians were ordering chest X-rays for 65 percent of their patients, when evidence-based practice called for an X-ray in only 5 percent of cases. This discovery was an early win for Texas Children’s Hospital to reduce unnecessary and unsafe testing and resource use and align more fully with evidence-based care guidelines.

Texas Children’s IT team traced the problem to a faulty order set within the hospital’s EHR and quickly rewrote the order set to reflect best practices. Within six months, the number of chest X-rays ordered for asthma patients had declined by 15 percent; today it has declined by 35 percent.

As the chest X-ray results show, the integrated three-system approach of measurement, team-based improvement, and best practices is rapidly changing the culture and transforming care delivery at Texas Children’s Hospital. The integrated hospital workgroups now actively use the approach to improve patient care by asking better questions about how care is delivered, uncovering the root causes of variation, and delivering evidence-based care. Davis explains: “Before the data warehouse, if someone in the organization wanted data from IT to answer a question, they would submit a report and there was this back and forth that was so frustrating for the clinicians, who really didn’t want to have to communicate with us. Now with the data warehouse as a hub with near real-time data, the clinicians can all see the data at the same time and they can very easily spot the data they’re looking for.”

Following on early successes, such as the asthma chest X-ray program, Texas Children’s Hospital began actively working to deploy the three-system approach to other care process families, including appendectomy, spine surgery, pneumonia, diabetic ketoacidosis, and more. The hospital even plans to expand the program beyond hospital-based care to include its primary pediatric practices and clinic-based care. On the operational side, Texas Children’s Hospital is using executive dashboards that feed into the EDW’s ability to create near real-time reports from data previously siloed within
multiple enterprise applications — the EHR, enterprise resource planning (ERP), and other business systems — to guide executive decision-making.

While the below results describe only the Labor and Productivity dashboard, others addressing length of stay and additional key operational areas are currently in use and more are under development.

Positive Results

Use of the EDW and healthcare analytic applications has produced numerous positive results at Texas Children’s Hospital. The following examples reflect just a fraction of the EDW-healthcare analytics projects underway at the hospital. As a result of direct benefits associated with just these four Texas Children’s Hospital projects — two clinical improvement and two operational efficiency projects — as well as efficiency gains related to EDW versus EHR reporting, Texas Children’s Hospital realized estimated $4.5 million in direct benefits in one year.

Impact of EDW vs. EHR Reporting

Texas Children’s Hospital has significantly improved efficiencies with EDW information delivery. On average, each EDW report costs 70 percent less to build than an EHR report. And, because the EDW visualizations enable end users to quickly and easily drill down into the data, one visualization is, on average, equivalent to 10 EHR-generated reports.

Labor Productivity

The Labor Productivity Advanced Application built on top of the EDW provides Texas Children’s Hospital with a global view across its various cost centers and enables increased operational efficiency management of hospital-wide departments as it relates to volume, FTE budgeting and actual premium dollar (overtime, temp labor, etc.). The application provides correlated data that aids in the efficiency and management of FTE monitoring and flexible budgeting.

In the past, reports were manually pulled from a number of sources, such as the enterprise resource planning (ERP) system, the EHR, and other administrative service sources, and then forwarded on to a third-party hosting vendor, which coalesced the data. Savings and benefits as a result of using the EDW-based application include:

- Near real-time reporting, which enables Texas Children’s Hospital to make timely staffing and budgeting decisions.
- Greater than an estimated $120,000 annual savings associated with automated data integration versus previous manual data pulls and data aggregation.
Increased adoption of a data-driven approach to labor management by unit managers due to ease of use, resulting in fewer hours spent obtaining internal application support.

Radiology

The Radiology Advanced Application built on top of the EDW helps the radiology department improve its outcomes through data-driven information. The framework being developed to support radiology will be similar across other departments, so the savings and improvements multiply as Texas Children’s Hospital expands to other clinical services.

Implementation is recent and quantitative metrics associated with the benefits have not yet been fully established. However, there is an estimated $150,000 annual savings in just the preliminary automation of a “single source of truth” report versus previous manually assembled reports that had differing data. The savings does not include all the individual staff members who were reviewing numerous reports each month, trying to decipher what data was correct. Nor does the quantitative data to date include measures on order timeliness, referrals, billings, or production efficiencies. These additional benefits are being benchmarked and will be quantified in the near future.

Some of the metrics implemented and/or being considered for future analysis include:

- **Order Timeliness**: Time from order to order completion, and time from order completion to order report to physician. Improving order timeliness helps drive more responsive patient care and improved efficiency and satisfaction of healthcare providers.

- **Referrals**: Outbound versus inbound referrals. Inbound referrals allow for an integrated view of the patient across the care continuum within the EHR.

- **Patient Satisfaction**.

- **Financial**: Cost per test compared to revenue generation.

- **Billing**: Orders completed versus order billed.

- **Production**: Volumes. Understanding volumes assists with resource planning and drives efficiency improvements.

Asthma and Appendectomy Advanced Healthcare Analytics Applications

Two additional clinical projects using the EDW and its applications are asthma and appendectomy. Asthma is the most common chronic disease of children. An estimated 80,000 children in Houston alone suffer from asthma. In 2011, asthma accounted for 3,000 emergency department visits and 800 hospital admissions at Texas Children’s Hospital.
Improvements include:

- Established and drove usage adoption of evidence-based order sets.
- Identified waste associated with the unnecessary use of chest X-rays for asthma patients.
- Decreased and sustained unnecessary chest X-ray usage by 35 percent by using the three system approach.
- Increased usage of an EHR-based asthma action plan by 90 percent of physicians.

Appendectomy is the most common indication for abdominal surgery in children. It accounts for a third of all childhood admissions for abdominal pain. Fewer than one-third of children have “classic” presentation.

Improvements include:

- 93 percent standardization on mono-therapy antibiotic.
- 85 percent order set usage adoption.
- Decreased complications and reduced lengths of stay.

Maintaining Transformational Change

Keys to Texas Children’s Hospital maintaining the transformation change based on the EDW three-system approach include:

- Organizing permanent, integrated workgroup teams consisting of physicians, nurses, IT, quality and patient safety, quality improvement, clinicians, and business analysts that are responsible for a clinical program or clinical services over the long-term.
- Integrating critical elements of evidence-based practices into the delivery of care.
- Establishing baseline measures, AIM statements with measurable goals, and on-going review of results versus targets. Outcome and balance metrics are included.

Texas Children’s Hospital now has a solution that integrates data management with evidence-based practice, operational data, and financial metrics — and enables clinicians and hospital leaders to understand the larger scope of care delivery. With data-driven insights at their fingertips, Texas Children’s Hospital clinicians can continue to raise the bar of quality care.
APPENDIX: FURTHER READING

Success Stories

Health Catalyst has collected a number of examples of customer successes with data warehousing and analytics. A detailed list can be found on the Health Catalyst website on the Success Stories tab (www.healthcatalyst.com/customers). Included are the following:

How to Increase Professional Billing Charge Capture with Healthcare Analytics

The demand on hospital coders continues to rise — and even more so with the ICD-10 rollout. At the same time, health systems want to make sure professional billing charge captures are accurate. Learn how North Memorial Health System leveraged their hospital enterprise data warehouse with the Health Catalyst Professional Billing Module to increase the number of provider notes with sufficient clinical data for billing, increasing their monthly net income and improving their hospital coding staff productivity by 25 percent. [http://www.healthcatalyst.com/success_stories/increase-professional-billing-charges-with-healthcare-analytics/](http://www.healthcatalyst.com/success_stories/increase-professional-billing-charges-with-healthcare-analytics/)

Using Advanced Analytics to Manage Population Health in Primary Care Clinics

The need to effectively manage the health of populations is largely driven by the fact that 5 percent of the population accounts for 50 percent of healthcare costs. Being able to identify these patients, provide high-quality care, and reduce their utilization is a pressing goal for many of today’s primary care providers (PCPs). Learn how one organization used healthcare analytics to meet this challenge. [http://www.healthcatalyst.com/success_stories/using-advanced-analytics-to-manage-population-health-in-primary-care-clinics/](http://www.healthcatalyst.com/success_stories/using-advanced-analytics-to-manage-population-health-in-primary-care-clinics/)

Streamlining Radiology Operations and Care Delivery Through Analytics

Texas Children’s Hospital radiology practice administrators were dedicating several hours each week to performing manual report reviews, which interfered with time they were able to spend streamlining operations and delivering patient care. After adopting a healthcare analytics system, the organization is now able to spend more time improving their operations and increasing patient satisfaction — while also reducing costs by an estimated $400,000. [http://www.healthcatalyst.com/success_stories/how-to-improve-radiology-operations-through-health-analytics/](http://www.healthcatalyst.com/success_stories/how-to-improve-radiology-operations-through-health-analytics/)
The Problem of Measuring Spinal Surgery Outcomes and How Analytics Can Help

Spinal problems are a common issue with a profound impact on healthcare costs. Faced with the high cost of surgical spine care in an industry transitioning to value-based payments, health systems need analytic solutions to evaluate the effectiveness of surgical interventions. Learn how one medical center used analytics to:

1. Build a spine registry;
2. Drive patient engagement through patient portal usage;
3. Integrate data from a multitude of quality of life surveys; and
4. Develop a platform to measure spinal surgery outcomes.


Using a Hospital Enterprise Data Warehouse to Reduce CAUTI Costs

A mere 10 weeks from the decision to implement a Patient Injury Prevention Advanced Application, the health system was running its first reports. Health Catalyst estimates the organization will reduce their associated surveillance activity, including collecting, recording, analyzing, interpreting, and disseminating, by 80 to 90 percent over their former manual tracking process. http://www.healthcatalyst.com/success_stories/hospital-data-warehouse-supports-efficient-enterprise-level-infection-surveillance/

Building a High-Quality Cancer Care Delivery System

An EDW allowed the hospital to track which protocols were being used by which clinician as well as the incidence of emergency department or urgent inpatient admissions after receiving a specific treatment protocol — a useful indication of the protocol’s relative effectiveness. http://www.healthcatalyst.com/success_stories/building-a-high-quality-cancer-care-delivery-system/

Population Registries Kick-Start Rapid-Cycle Clinical Process Improvement

After implementing an EDW, Texas Children’s Hospital overcame previously cumbersome processes for defining patient cohorts, analyzing populations and implementing improvement programs that drove measurable and sustainable improvement. http://www.healthcatalyst.com/success_stories/population-registries-kick-start-rapid-cycle-clinical-process-improvement/

How to Reduce Healthcare Labor Costs by Using an EDW and Analytics

In the shift to value-based care, Texas Children’s Hospital began to see its bottom line decline and needed an immediate solution to address labor and

Improving Population Health Through the Use of Data and Reporting

Using a Late-Binding™ Data Warehouse, one large healthcare system switched from manually pulling together reports with varying data to having near real-time data available when they needed it. “The results,” says one of the health system’s executives, “enable our care coordinators to drive preventive care and ultimately lower our population health costs.” [http://www.healthcatalyst.com/success_stories/population-health-enabling-patient-engagement/](http://www.healthcatalyst.com/success_stories/population-health-enabling-patient-engagement/)

Hospital Acquired Infections – How to Reduce Surveillance Waste

Adopting an enterprise data warehouse had a number of positive affects for this medical center. Not only did they create ways to more effectively discover and treat certain hospital-acquired infections, the organization was also able to develop five streams of quality improvement: infectious disease, population health, cardiovascular, neuroscience, and oncology. [http://www.healthcatalyst.com/success_stories/hospital-aquired-infections-90-percent-reduction-in-surveillance-waste](http://www.healthcatalyst.com/success_stories/hospital-aquired-infections-90-percent-reduction-in-surveillance-waste)

How to Reduce 30-Day Heart Failure Readmission Rates

Thirty-day heart failure readmission rates were targeted as a key area for improvement. With the help of an enterprise data warehouse, one health system was able to obtain near real-time analysis, providing insight into measurable results and a way to track progress. The results continue to exceed the organization’s expectations and goals. [http://www.healthcatalyst.com/success_stories/reduce-30-day-heart-failure-readmissions/](http://www.healthcatalyst.com/success_stories/reduce-30-day-heart-failure-readmissions/)

How to Reduce Sepsis Mortality Rates by 22%

Healthcare organization Multicare reduced septicemia by 22 percent, leading to a $1.3 million cost savings in the same period by using an EDW and advanced analytics. The resulting success prompted the organization to tackle additional areas of improvements. [http://www.healthcatalyst.com/success_stories/how-to-reduce-sepsis-mortality-rates-by-22/](http://www.healthcatalyst.com/success_stories/how-to-reduce-sepsis-mortality-rates-by-22/)

How to Reduce Unnecessary Early Term Deliveries

Seeking a path to sustainability, North Memorial Health Care started with a Late-Binding™ Data Warehouse. Using their new technology, the organization was able to determine which area of improvement they should focus on first:

**Other Resources**
Articles, documents and other references included in this ebook can be accessed through hyperlinks embedded within the book or through the following URLs.

CONTRIBUTORS

Dale Sanders, Senior Vice President, Editor

Before moving to healthcare, Dale Sanders worked in the military, national-intelligence, and manufacturing sectors, specializing in analytics and decision support. Previously, Mr. Sanders was CIO for the National Health System in the Cayman Islands, CIO with Northwestern University Medical Center and regional director of Medical Informatics at Intermountain Healthcare where he was chief architect of their enterprise data warehouse. Mr. Sanders is a founder of the Healthcare Data Warehousing Association.

Steve Barlow, Co-founder & SVP, Client Operations

Mr. Barlow is a co-founder of Health Catalyst and former CEO of the company. He oversees all development activities for Health Catalyst’s suite of products and services. Mr. Barlow is a founding member and former chair of the Healthcare Data Warehousing Association. He began his career in healthcare more than 18 years ago with Intermountain Healthcare, and acted as a member of the team that led Intermountain’s nationally recognized improvements in quality of care delivery and reductions in cost.

Bobbi Brown, Vice President, Financial Engagement

Bobbi Brown started her healthcare career at Intermountain Healthcare before moving to Sutter Health and, later, Kaiser Permanente, where she served as Vice President of Financial Planning and Performance. Ms. Brown also acted as team lead on the installation of planning software at Ascension as Vice President of Financial Engagement. Ms. Brown regularly writes and teaches on topics including value-based purchasing, payer-provider collaboration, revenue cycle management, and Medicare reimbursements.
Dan Burton, CEO

Dan Burton serves as CEO of Health Catalyst. He became involved with Health Catalyst when it was a three-person startup. Mr. Burton is also the co-founder of HB Ventures, the first investor in Health Catalyst. Prior to Health Catalyst and HB Ventures, Mr. Burton led the Corporate Strategy Group at Micron Technology. He also spent eight years with Hewlett-Packard in strategy and marketing management roles. Before joining HP he was a consultant with the Boston Consulting Group, where he advised healthcare systems and technology companies.

Mike Doyle, Vice President

Mike Doyle joined Health Catalyst in May 2013 as Vice President, and he has been connected with the organization’s senior leadership team since 2006. Prior to Health Catalyst, Mr. Doyle led the Enterprise Data Warehouse (EDW) program at Allina Health, growing the EDW from a clinical improvement initiative into a strategic asset spanning Allina’s 11 hospitals and 100+ clinics. Mr. Doyle previously held roles as a database administrator, web programmer, data architect, and business intelligence developer.

Eric Just, Vice President, Technology

Eric Just joined Health Catalyst in August of 2011 as Vice President of Technology, bringing more than 10 years of biomedical informatics experience with him. Previously, he managed the research arm of the Northwestern Medical Data Warehouse at Northwestern University’s Feinberg School of Medicine. While working as a senior data architect, Mr. Just helped create the data warehouse technical foundation, determined new ways to extract and load medical data, and led the development effort for a genome database.
Michael McCuistion, Technical Director

Michael McCuistion joined Healthcare Catalyst in December 2012 as Technical Director. Previously, he spent more than 17 years at PeaceHealth, a nine hospital Catholic Healthcare System in the Pacific Northwest. While there, Mr. McCuistion managed the Enterprise Data Warehouse and Business Intelligence team. Prior to working with PeaceHealth’s EDW, he managed the system-wide Help Desk for six years and served in desktop support capacities for the health system.

Chris Rains, Data Architect

Chris Rains joined Health Catalyst in September 2012 as a data architect. Prior to coming to Catalyst, he was a senior data analyst, lean process coach and project manager at Bank of America. While there, he used SAS and the Teradata enterprise data warehouse to identify opportunities for improvement. Previously, Mr. Rains worked in database analytics and Java web app development roles in various industries.

Russ Staheli, Vice President, Analytic Applications

Russ Staheli started as an intern before becoming an outcomes analyst at Intermountain Healthcare. Prior to joining Health Catalyst in October 2011, he served as a management engineer programmer analyst for Duke University Health System in Performance Services supporting Infection Control and Epidemiology efforts. While there, Mr. Staheli also worked as an external consultant to advance the analytical work of the Duke Infection Control Outreach Network (DICON), a collaborative of more than 30 community hospitals.
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 large amounts of data utilized in population health and ACO projects supporting over 30 million unique patients. Health Catalyst platform clients operate over 135 hospitals and 1,700 clinics that account for over $130 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, Children’s Hospital of Wisconsin, Crystal Run Healthcare, Indiana University Health, Kaiser Permanente, Memorial Hospital at Gulfport, MultiCare Health System, North Memorial Health Care, Partners HealthCare, Providence Health & Services, Stanford Hospital & Clinics, and Texas Children’s Hospital. Health Catalyst investors include CHV Capital (an Indiana University Health Company), 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.