6 Reasons Why Healthcare Data Warehouses Fail

By Steve Barlow

Enterprise data warehouses (EDWs) are notoriously difficult and expensive. 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.

However, when an EDW is done well, it can become a very important strategic asset. Healthcare administrators facing value-based payment models need ways to cut waste, reduce costs, and vastly improve care delivery. To do this, they need a measurement system for all the data their organization collects. A healthcare data warehouse is the answer.

The journey to building a data warehouse requires serious consideration of the risks—and finding ways to do things right the first time, thus limiting the ultimate risk of complete failure.

What makes the difference between a healthcare data warehouse project that fails and one that succeeds?

The goal of a data warehouse is to organize complex data from dozens of places into a single source of truth that makes it possible for an organization to succeed in the era of value-based care (i.e. cut waste, reduce costs, and improve care delivery).

Knowing what can doom a project is critical. Over the years, my colleagues and I have seen certain mistakes made time and again. The six ways to fail listed below encapsulate a majority of the missteps.
Six of the Most Common Ways a Healthcare Data Warehouse Project Fails

1. **A solid business imperative is missing.**

If you build it, they will come. True for a baseball field (à la “Field of Dreams” the movie); not so true when it comes to a data warehouse project. When the EDW initiative comes primarily from a technical perspective and there is no real solid business case, it’s hard to get after-the-fact buy-in from the business, clinical, and operations teams. 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.

Other related ineffective reasons for building an EDW include, “the competitor down the street just deployed a data warehouse” and “we need more metrics dashboards.” An EDW built on these premises will become an end in and of itself and fail to provide the actionable data needed by clinical teams. So, it will be ignored and fail to contribute anything meaningful to the organization as a whole.

For an example of this, I look to a time early in my career. My team and I were approached by our hospital’s key clinical strategists and told, “We need a measurement system to support clinical integration.” So we said, “Okay. We’re database people. We know how to build data models. We’ll go to work.” Then, we locked ourselves in a backroom for nearly a year and designed a data model we were really proud of. When we finally revealed our seemingly perfect data model to the clinical team, one of the doctors spoke up and asked “Hmmm…where is my diabetes patient or my diabetes registry? That’s what I need.” And we couldn’t say. It didn’t exist in our model. So, we had to go back to the drawing board and figure out a way to modify the model and provide the ability to design and build patient registries more quickly.

The lesson learned? Know your reasons for building, and constantly consult the people who will be using the data from EDW. Understand why data is needed and how the data will be used. Involvement of business and clinical teams in defining expectations is critical. We see this failure taking place at many organizations.

“Know your reasons for building, and constantly consult the people who will be using the data from EDW. Understand why data is needed and how the data will be used. Involvement of business and clinical teams in defining expectations is critical.”
Executive sponsorship and engagement is weak or non-existent.

This is probably obvious, but it’s not enough for executives to green-light a project and tell IT to let them know when it’s done. From the beginning, as the strategic importance of the EDW is defined, through deployment and governance, there must be significant senior leadership involved. The governing body that steers the EDW project should comprise senior leaders from various operating areas of the organization including clinical, finance, quality, strategy, and administration. Frontline clinicians and business staff must see that leadership supports the data warehouse with more than lip service. The governing body needs to be fully engaged in approving the budget, technical team, vendor partner, and priorities.

The graphic below shows how to organize the teams working on the EDW initiative. The senior executive leadership team provides the overall governance and prioritization of initiatives; however, this group still needs to be small enough for agile decision making. Moving from the bottom of the image to the top, the work group develops and refines the clinical content and analytics under the guidance of the clinical implementation team, who leads the implementation. Separately, the content and analytics team supports the development of clinical content.
Multidisciplinary, frontline teams composed of clinicians, technologists, analysts, and quality personnel should be involved throughout the design and deployment process to ensure that solutions meet their tactical and strategic requirements.

and provides feedback on the analytics they receive. Finally, the guidance team provides governance and oversight for all these teams and reports up to the senior executive leadership team.

When all of these teams are fully engaged — especially the executive team — the organization experiences significant success in clinical improvement. We know a CEO at a large healthcare system who called his EDW “one of the crown jewels of this company.” That’s how strongly this CEO felt about the strategic importance of the EDW. Teams saw that level of executive support, and it was reflected in the improvements and http://www.healthcatalyst.com/asthma-waste-reduction-clinical-improvement efficiencies accomplished by that organization.

Frontline healthcare information users are not involved from start to finish.

An organization can have a strong business case and great executive sponsorship but still fail if it doesn’t involve clinicians and other users on the frontlines 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 composed of clinicians, technologists, analysts, and quality personnel should be involved throughout the design and deployment process to ensure that solutions meet their tactical and strategic requirements.

When clinicians are involved from the beginning to the end, they tend to get very excited and invested in the process thus increasing its chances for success. For example, we worked with a clinical team to create a definition of an asthma patient for the EDW. At the end of a series of these meetings, one of the clinicians pulled a $10 bill out of his wallet and declared, “I bet 10 bucks that we have the best definition and criteria for pediatric asthma patients in the country.” The clinicians owned that definition, and they defended it. While it did evolve slightly over time, their early and consistent involvement in making those kinds of binding and definitional decisions made the process more efficient and easier to implement.
Boil-the-ocean syndrome takes over.

Trying to be all things to all people spells disaster. An IT team simply can’t deliver value quickly if it is bent on boiling the ocean — identifying every possible piece of data just in case it might be needed at some time in the future. Maintaining data over its lifetime is an expensive endeavor, and there are shelves of data elements that never get used.

Taking an iterative approach — choosing one clinical area at a time to focus on and gradually improving quality and cost — is what pays off in both the short and the long run. The key is to identify a few case studies with teams consisting of the clinicians and frontline personnel who will use the data and know the processes involved. This will allow the data architects to build a reasonable substrate of data because the costing data, satisfaction data, provider data, EMR data, etc. for that improvement project is a known entity.

This idea is displayed in the following graphic. On the left-hand side, there is the danger of capturing too little data and not being able to produce the needed reports and analytics. On the other side, the scenario of capturing everything means expensive and unused data that will just sit in storage. The happy medium between these two is the process described above: collect enough data to manage and measure clinical processes effectively using the input from the team.
We’ve seen the wasteful approach taken to the extreme. I was invited to a large reputable healthcare organization to review their work building their data warehouse. During the meeting, a number of medical informaticists, terminologists, and data architects pointed at a big wall covered in a written-out data model and said, “This is what we’ve done.” They had spent two years building this data model and it wasn’t even fully attributed (i.e. they had identified the entities in the data model only—they knew they needed provider data, patient data, diagnosis procedure billing data, but they hadn’t gone further than that). I suggested tactfully that they may want to consider a more agile, flexible way; two years just getting the data model, plus many more years to load that data model and writing the ETL routines is not sustainable. It’s a good way to lose executive support. They respectfully said they preferred to stay the course, but I’ve since learned that they’ve shifted to a more agile approach. I’m sure that’s working much better for them.

Being too idealistic (and not realistic) about what can be accomplished.

Starting the design of a data warehouse with a blank canvas and imagining that anything’s possible is asking for trouble. Sitting down with clinicians and brainstorming every possible esoteric piece of information you could ever want to measure results in a set of data that, while desired, may not even be available. Organizations need to consider the resources that are ultimately required to load the data.

Too often health systems 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 of what is at hand, they can put themselves on the road to failure. Instead, organizations should adopt a more pragmatic approach work with the data already available to them and add more as feasible.

Organizations need to consider the resources that are ultimately required to load the data. Too often health systems fail to consider what information they are currently capable of capturing through their financial, EHR, and other systems.
The graphic below underscores the importance of considering every step. The green box represents capturing data. In this phase, the organization needs to consider what data needs to be captured, ensure that the data is of high quality, and look at data capture processes in operational workflows throughout the hospital. Data provisioning is represented in the blue box, and here the organization needs to consider its data warehousing strategy and capabilities for building reports. In the orange box, the organization does the fun stuff—data analysis. This is where data is interpreted and evaluated for quality. Discoveries leading to improved care and efficiencies happen in this phase. Historically, health systems have focused too much on data capture. Instead, leaders should balance resources across getting data in, getting that data out, and turning it into actionable insights.

Not considering every step has consequences, as I learned when working on an improvement project for heart failure care. We worked with clinicians to identify a patient cohort for heart failure (similar to the process I described above with the asthma patients), and we decided to use those patients with an ejection fraction less than 40 percent. We were excited to roll ahead with the new care pathway. Then, we hit a snag.
When we went back through the data, we couldn’t find Left Ventricular Ejection Fraction (LVEF) in the source system. They were not electronically capturing LVEF in their system. Before we could even start, the organization needed to change the workflow process so that they captured LVEF in the EMR. That took another few months, and this project was paused until then. Again, health systems need to consider the data that is actually captured in the system rather than just assuming that everything they need is already there.

Worrying about getting data governance “perfect,” which immobilizes the project.

Making data governance a prerequisite for implementing a data warehouse is a mistake; an organization doesn’t need the have all the answers before starting. Waiting for a robust data governance policy to be in place will cause the project to die under its own weight. I recommend that a health system starts building the data warehouse, using high-level data governance principles to guide decision making.

These principles include data literacy, data stewardship, and data exploitation. Data literacy should include programs to increase user understanding of the data and ways for users to practice with the data. Data stewardship is the process of identifying and mobilizing the appropriate business or clinical steward of the various data domains. A data steward’s responsibilities are concerned with the authorization and quality of their respective data sets. Data exploitation policies should err on the side of getting data into the hands of knowledge discoverers trained to mine and interpret information. Too often, IT departments have too tight a clutch on data and release it in a slow trickle that hinders the usability of the data. An organization needs to empower its data analysts. Clearly, patient identifiable data needs to be handled carefully. But there are ways architecturally that a health system can build data systems that will remove PHI and still allow plenty of room for knowledge discovery.

The executive team and the guidance committee, mentioned above, should consider two dimensions of this issue: priority governance, which is the process of deciding what data should be made available first; and data governance, which establishes policies that ensure consistent data quality and definitions.
When it comes to governance, too often perfection is the enemy of progress. And while data governance is very important, it’s more efficient to put it further down the development path. Once the warehouse is built and an organization has data to govern and react to, the data governance and prioritization processes will accelerate.

Resources

ABOUT HEALTH CATALYST®

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

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

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

About the Author

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 over 18 years ago at 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. Mr. Barlow holds a BS from the University of Utah in health education and promotion.