CHAPTER 1

Business Intelligence
An Introduction

When I tell people what I do for a living they respond one of two ways. First, "Business intelligence, isn't that an oxymoron?" Oh, first time I have heard that! So funny. The second response is: "What?" Complete with a blank stare on their face.

I almost always qualify it with something like "You know, reporting and analytics." That usually seals the deal. It's not completely accurate but in these instances I am okay with good enough.

Many definitions of business intelligence (BI) exist; the most well-known is "The right information to the right person at the right time in the right way." This is my least favorite because it implies a factor of luck. Perhaps the oldest was written by H. P. Luhn in 1958: "The objective of the system is to supply suitable information to support specific activities carried out by individuals, groups, departments, divisions, or even larger units. . . . To that end, the system concerns itself with the admission of acquisition of new information, its dissemination, storage, retrieval, and transmittal to the action points it serves." The one I use most often is: *BI is the integration of data from disparate source systems to optimize business usage and understanding through a user-friendly interface.*

Data warehousing is a companion phrase to BI. The well-documented best practice for BI is to create a data warehouse. A data warehouse is exactly what it sounds like, a place where a lot of data resides. Good data warehouses have a strong organization system, like the card catalogs from libraries of the past. Without that strong organization system, healthcare companies find themselves digging through their data warehouse for data, not an optimized method for certain. To be clear, business intelligence is not an IT (information technology) activity. But it does require support from your IT group for the more technical aspects of data warehousing. We address more of these in Chapter 5.

The truth is that simple definitions don't really do business intelligence justice. True BI, good BI, is an enablement mechanism to provide IT leaders and hospital executives the best information possible to improve their ability to make informed decisions. BI helps organizations go from management by instinct to management by data. BI isn't just a capability, although certainly it provides capabilities; when done well BI can become the life-blood of your organization, providing your organization with key performance indicators that help manage revenue cycle management, quality and safety indicators, or outcomes associated with diabetes management, to name a few. Few healthcare organizations treat BI as life-blood. But as you will see throughout these pages, when they do, the results are nothing short of stellar.

What BI Isn't

BI isn't reporting, it isn't analytics, it isn't data warehousing, and it isn't dashboards. All of these things individually do not make a BI program, but put them together and that is exactly what BI is. Business intelligence enables all of these. BI is greater than the sum of its parts. You may question why BI enables

data warehousing, but the truth is that you don't need a data warehouse if you don't intend to analyze data or report from it. BI is an industry and a skill set, but BI isn't the group you go to that will provide you the knowledge or intelligence about your organization. Good BI means putting valuable information at the fingertips of many businesspeople, not just a lucky few.

Is It Really Worth It?

I received a call one early January day. For consultants, those are the calls that are the most intriguing, because usually it means that someone really needs help, contemplated their next steps over the winter holiday, and waited until the new budget year to make the call. This call came the first business day back from break, and the caller was a director of IT. She was looking for an "objective" voice that had knowledge of both tools and BI programs. Two days later I was sitting in her office learning all the details. The most important question she asked me was: "Does anyone really do this? It seems so complicated and hard to find the right resources, is it really worth it?"

The answer is yes; many organizations have done BI and done it well. They have found the proverbial gold at the end of the rainbow, where all the work they did brings them the value from the data that they needed.

Do You Need BI?

If your organization uses data to make decisions then the answer is yes. If your organization *wants* to use data to make decisions then the answer is yes. If your plan is to hire a team of really smart analysts then the answer is no, because BI is

meant to deliver information to a broad audience. The degree to which you have to invest and create your BI program is what should vary.

"Do you need BI?" is a great question, and one every person who is in charge of a BI initiative should ask themselves often, and here's why:

- If you don't, someone else will, usually around budget time.
- If you aren't asking then you probably aren't thinking about how to make your program the most relevant for your organization.
- These programs are expensive; between tools, resources, and time they cost money. You have to make sure that you are providing the value that matches the investment. If you aren't, then what are you doing?

Ask yourself these questions at least twice a year, and depending on how your organization is structured, have a prepared statement or a PowerPoint ready when these questions are posed to you by someone else.

Healthcare Information Environment

To "do BI" you will have to organize your data for usage. Odds are, as you read this, your hospital or clinic has data stored somewhere. That data comes from a transactional system like an electronic health record (EHR) or a financial system. The data on its own is not user-friendly for the majority of businesspeople. If the goal of BI is to put better information into many businesspeople's hands you must take the time to organize your data to ensure that it's easy to use and provides the most value. That is where a traditional data warehouse comes in.

When working in healthcare I do make a few modifications to the traditional data structures you see in other books on the

topic. For example, the analytics sandbox and audit control sections are critical to healthcare organizations, but maybe not as necessary for retail. Each of them provides a method to allow your more sophisticated analysts access to the data that is granular. The analytic sandbox provides your analysts a "play space" to create predictive models that can help you adjust staffing in your emergency department without an impact to regulatory reporting. The audit control environment (ACE) provides a one-stop-shop for both internal and external auditors to see the data and the path the data took to validate your approach for anything from JCAHO (the Joint Commission for the Accreditation of Healthcare Organizations) reviews to medical records reviews for public health documentation.

The first thing you should know about your data environment is that it is unique to your organization and should be created based on the needs and wants of your hospital or health plan. As you construct your information environment, important key criteria need to be kept in mind. These environments are built to optimize stability and data usage for your organization. Some methods of shortcutting the process exist, but few deliver the capabilities that are promised during the sales cycle. We review these methods in Chapter 5, but for now let's look at the baseline healthcare information environment that I recommend.

Let's start at the beginning, or in this case, on the left side of Figure 1.1. The source systems in healthcare do vary, but they generally follow two categories: Clinical and Financial. In theory, anything that has data can be a source (e.g., Excel), but as you consider what you bring into your data warehouse you need to ask yourself a basic question: "Yes, you can, but *should* you?" Every industry is buckling under the weight of data, prompting interest in "the Cloud." But not all data is equal, only data that provide valuable insights should be stored in your data warehouse.
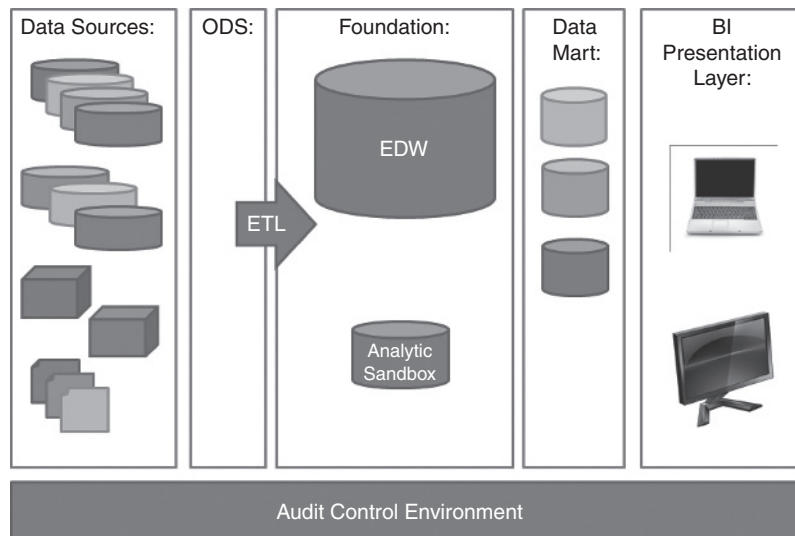
FIGURE 1.1 Healthcare Information Environment

The next step in the environment is the staging area, often referred to as the *operational data store*. Named for the type of data that generally appears, for healthcare we find that the staging area provides an important two-stage function. The first is to isolate activity against your transactional system. In other words, you don't want your EHR performance to slow down so that you can run a report, and staging areas help provide that buffer. The second function is to maintain a completely untouched version of the data that can be used if there is a data load failure. This prevents the need to go back to your transactional systems (EHR or financial system) to pull the data again. This also protects the transactional system. We take a copy of this data and store it, untouched, in the audit control tables. These tables enable support for the internal and external audits that are routinely done in healthcare. Although not necessary, these control tables make audits much easier to manage because all of the required data and supporting documentation

needed for audits (e.g., transformation scripts that outline the changes that were made to the data) are included.

The next section is the extract, transform, and load, or ETL, portion of the healthcare information environment. ETL is absolutely critical for healthcare. In its raw form most healthcare data is unusable for the average businessperson. The "transformation" part of ETL is the application of business rules so we can aggregate things like encounters and deduplicate provider and patient records. Once the business rules are applied we place this data into an organized set of data tables within the data warehouse itself where the most granular level of data exists, such as the claim level or patient level. The function of a data warehouse is to aggregate the most granular data up into summary level data to improve the ability to use the data. For example, in its most granular form a claims record can have many rows (think about this in the context of Excel). Each row may represent a different status of the same claim, such as submitted, paid, reversed, or rejected. Most claims go through many different statuses before it's considered a "final paid" claim. If the average businessperson is only interested in the percentage of rejected claims, then having to go through this granular data is time consuming. The organization of your data warehouse removes the need for a businessperson to have to do that by aggregating (in this case summing) the number of rejected claims for a specific time period or place based on business rules.

Within this environment an analytic sandbox can also exist. These sandboxes are the "play space" for your analysts and they should not allow the average businessperson access. These sandboxes give analysts a chance to test models and create new business rules or functionality without exposing the activity to everyone. This allows for innovation using data. As we move right in Figure 1.1 you see the data mart section. Data marts represent logical subject areas, such as claims or encounters, any subject area that you report on frequently; data marts are

created to improve performance (response time of reports). In other words, most of your standard reports will access data contained in these data marts. This is not a required method, but if you know that you will have a high degree of usage of reports, these logical subject areas are quite helpful in managing the overall performance of the system.

Finally, we have the BI presentation layer. This presentation layer represents the BI products that you have bought (e.g., MicroStrategy, Cognos, Tableau, SAS) to allow your businesspeople access to the data. This would include products that provide reports, dashboards, or even ad hoc analysis to all users of the data. I will make a somewhat controversial statement and say that your BI presentation layer will likely have multiple tools. Some tools are better at certain things than others, and if you have a strong analytic component of your BI program, your standard BI product will likely not meet all of your needs.

Every BI program is different because every organization is different—this is not a one size fits all. If you've seen one BI program you have seen one BI program. Critical similarities, such as the need for a data warehouse, ETL (in some form), and a method of report distribution exist, but the rest is the art of good BI.

The art of BI is probably the most difficult thing to master. It is often the function most overlooked by organizations. I have seen hundreds of BI programs, and each one is different, but the ones that have considered the "softer" side of delivering valuable information to their businesspeople are successful. They have considered what it takes, who it takes, and haven't compromised on quality, or backed down on the need for good architectural standards. They are a testament to what good BI programs *do*. Good programs have strong BI leadership that is aligned and empowered. They have a strong and dedicated staff, they don't go for the "easy" button, and

they understand that the foundation of everything starts with data modeling.

Data Modeling

I have struggled many times to explain or define data modeling to a businessperson. When business leaders are trying to decide whether they should invest in this thing called a *data model*, where the deliverable, at least from their perspective, is a drawing, it isn't that complicated. But when they get down to what data modeling can do for their BI deployment and how it does it, it gets tougher.

Most people understand the concept of data hierarchies, and the idea that some data cannot be summed. Much of that understanding is the foundation of data modeling. When you do data modeling right, you define the hierarchal nature of the data. A great example is time: Year, Quarter, Month, Week, Day, Hour. For most healthcare organizations you have some type of organizational hierarchy such as Department, Unit, and Floor. The other part of data modeling is identifying the "facts" and "dimensions." Simply, the things that can be summed and the dimensions are attributes of the facts. For example: A dimension of a patient or member is their unique identifier, so you can't sum that. You can sum the total number of patients or members, which is a fact. The vast majority of data modeling is simply organizing your data in the appropriate hierarchies, facts, and dimensions.

The art of data modeling is to know when to do what, and how to create the relationships between these. In healthcare the relationships in our data are complex; in data warehousing we call it a *many-to-many*, which happens over and over again (think about the relationship between a patient, physician, diagnostic code, member, address, etc.).

I will stop here, at the risk of getting too much into the detail to confuse the point. The point is, data modeling for healthcare is complex and requires a special skill set. We discuss this further in Chapter 5.

The Don'ts

Much of this book focuses on what you *should* do, but here are four things you absolutely should not, under any circumstances, do:

1. *Never* make a consultant the leader of your program. Yes, I said it. And, yes, I am a consultant. But for much longer I was a BI practitioner in an organization a lot like yours, and I have made this mistake. Here's the simple reason why this will never work—consultants are not employees. Sure, consultants want your program to succeed but their reasons for that are not aligned with what is best for *your* organization—they are aligned with what is best for *the consultant's* organization. Consultants will strive to be kept around and will find ways to ensure that happens. It doesn't matter how much you like them or trust them, and it doesn't mean that they set out to mislead; it just means that you are not aligned with each other, and that is never a good thing. Instead, invest in a BI leader who is on your payroll and reports up to the executive who has the most on the line for a successful BI program.

2. Don't ignore or forget about your staff. This could probably be said for any program, not just BI. But at the end of the day BI is your intellectual property (IP), and the only way you get good IP is to have really smart, dedicated people thinking about your business and the data. Few things are as powerful as a highly internally motivated staff that will do whatever it takes to get the job done the right way.

3. If it sounds too good to be true, it is. Didn't our mothers teach us this? Well, please forgive me, my vendor friends, but when a vendor tells you that you can install and plug in your data and be up and running in a few days, take that with a giant grain of salt. Yes, *technically* this is true, but just because you *can* doesn't always mean you *should*. What you will get are some reports, and that can help get some executive support, but what you won't get are the things that are the hallmarks of good BI programs: things like a reuseable data warehouse, good data-quality processes, and automated ETL scripts.

4. Never de-emphasize the importance of a good data model. If I had a dime for every time I have seen this I would be rich. This is highly correlated to the "easy" button phenomenon. Again, sorry my vendor friends, but you usually cannot buy a 100 percent ready-to-go data model. If you do, please know that the value comes in the customization (yes, customizing a prebuilt template), so it can jump-start your program; it just won't get you over the finish line.

If you have decided that you need BI and are ready to start on this journey, here are the hallmarks of good BI programs. They:

- Have good leaders who are employees.
- Have solid executive sponsorship.
- Are willing, or have, invested in the right tools and people to get the job done.
- Are committed to data governance.
- Focus on all aspects of delivering BI.
- Maintain excellent communication out to the business community during any build cycle.

If you do these things, you are well on your way. The rest of this book is dedicated to each of these hallmarks of BI programs,

referred to as the *Tenets of Healthcare BI*. If you have started your program and already have some of these, you can skip to the chapters that will be most helpful to you. If you are starting from scratch, you'll benefit from reading this book from cover to cover and joining our online collaborative community for further discussions.