Healthcare Informatics: Analytics

Jason Burke
SAS Institute, Cary, North Carolina, U.S.A.

Abstract
In healthcare, we are just at the beginning of seeing how analytics—whether biased toward the business or clinical sides of the spectrum—can transform our ecosystem. Analytics will give us better guidance on how to control costs—not just line items, but the hidden and true costs of healthcare. Analytics will help us identify and dismantle old assumptions about the way healthcare is delivered—not relying on gut instincts and hearsay, but real evidence. Analytics will allow us to determine not only the treatments that produce the best outcomes, but the real factors that determine optimum treatment efficacy and cost. The hidden gems in healthcare will surface, and with them will come better lives for patients everywhere.

INTRODUCTION
Evidence-based medicine. Personal electronic health records. Disease management. Personalized medicine. These terms, among many others, reflect a rapidly growing change within the health sciences ecosystem—a transformative shift toward more information-based decision making related to patient care and healthcare cost management. For decades, the efficiencies and improvements attained in other industries through the adoption of information technology have largely been missing in healthcare, an ecosystem mired in paper records, administrative overhead, and labor-intensive business processes.

But that is all changing. The sustained rise in healthcare costs, consistent problems in patient safety, highly expensive prescription drugs, and inconsistent treatment outcomes have all contributed to a new drive toward making better use of the tremendous volumes of information flowing through the ecosystem. Whereas, historically hospitals have looked to expansions in service lines and facilities to drive top-line revenue growth, analytics that provide business opportunities in utilization, cost containment, and quality control are now seen as critical enablers of bottom-line financial performance. Health plans that have relied on relatively simple business rules to determine the appropriateness of reimbursements are now looking to advanced analytical models to identify previously undetected fraudulent claims activities and patterns. Drug researchers, struggling to find ways to bring innovative and safe therapies to market faster and cheaper, are aggregating tremendous volumes of data covering many years of research to look for biomarkers that can accurately predict drug safety and efficacy in named patient populations. Across the board, electronic data, whether business based or science based, are now seen as the fuel to power the engines of business and clinical analytics driving the evolution of patient outcomes and wellness.

But those growing volumes of data contain a hidden burden: How can we efficiently and effectively manage such large and disparate volumes of information? How do we make it useful? With information flowing from every corner of the healthcare ecosystem, how do we prioritize which data are most important, and how can we simplify the inherent complexity down to something with which educated human beings can make rational decisions? Anyone seeking an easy answer to this dilemma will be disappointed.

Modern information technology, especially in the areas of data integration, data quality, data management, and advanced analytics, holds the key to unlocking the power of this information and the corresponding business and scientific transformation contained within. Advanced analytics and information management sit at the center of the new health enterprise—an information-driven business and science-informed medical practice that can dramatically reduce healthcare costs and improve health outcomes for all patients. But only if organizations embrace them.

ANALYSIS PARALYSIS
The healthcare industry is no stranger to technology—hospitals have invested millions in medical devices for decades, for example. But capabilities with respect to information technology—electronic data collection, management, quality, analysis, and reporting—are reasonably new. Extensive paper-based forms, change-
averse physicians, tightly controlled business processes, overtaxed nurses, and business demands on self-funding investments have conspired to inhibit the proliferation of information technology in much of the healthcare sector. But as the industry has sought to better understand its deep-rooted problems in cost, quality, safety, and outcomes, a growing recognition has emerged that information technology must be a priority for every health enterprise.

And yet, when we speak of advanced analytics in healthcare, it is not uncommon to hear a list of excuses why the industry is not ready for them:

- **More technology.** Many people argue that until the industry has had more time to implement more technology that collects information electronically, there is little use in investing in advanced analytics.
- **More integration and standards.** Some people argue that, because the industry has historically lacked data standards that facilitate information aggregation and sharing, any insights that might be derived from their existing data would be of questionable value.
- **Data privacy.** Inevitably, some people will question the appropriateness of using personal medical information outside the context of care for that particular patient; Health Insurance Portability and Accountability Act is usually cited.

As organizations consider analytics-oriented projects and hear these concerns, it is quite easy to fall into “analysis paralysis”—continuously trying to find ways to overcome issues that cannot be overcome without doing the projects that elicited the issues in the first place. Organizations will always need more technology, but we have a lot today. We will always need better integration and deeper support in standards, but we have standards and integration models that are proven today.

We should always be holding patient data privacy at the forefront of our minds, but we have many ways of protecting patient privacy, while also allowing us to pursue improvements that will inevitably benefit those patients.

The question should not be whether to take on analytics as a corporate priority; the question should be how. And the answer is surprisingly simple, residing in the neonatal and pediatric units of every hospital in the world. Newborn babies, infants, and toddlers physiologically develop along a predefined biological path, one that serves to gradually bring new biological systems online and grow the systems already online until the person reaches adulthood. It is a long-term process, but one with clearly defined steps and associated personal abilities. Such is this case with analytics, as they are born and grown inside companies.

### ANALYTICAL MATURITY AND OBJECTIVES

The term *analytics* may be one of the most overused and misunderstood terms in the business community today, with the possible exception of *business intelligence*. Every software application that has the ability to run reports with numbers in them also has the capability of providing analytics. Any person who has taken a statistics course is absolutely capable of performing whatever analytics are needed for an organization. Even the definition of the term *analytics* is used in one context to describe web reporting, while in another context it describes the most obscure statistical methodology imaginable.

When we use the term *analytics* in healthcare, we are using it to mean something very specific:

Analytics are the complete series of integrated capabilities needed to provide progressively deeper statistical insights into health-related information.

We are describing capabilities—a capacity that can be found or learned within organizations and individuals. Those capabilities should be complete, meaning they cover all of the needed areas of information access, integration, quality, storage, management, interpretation, and governance. Those capabilities are also progressive, meaning that the simpler capabilities need to be in place to enable the more sophisticated capabilities to operate. They are statistical in nature, not merely mathematical. And they are progressively deeper, meaning the insights derived from higher-order analytical capabilities offer greater value than those of lower-order capabilities.

### THE EIGHT LEVELS OF ANALYTICS

So what are these capabilities that organizations and individuals need to have? There are eight levels of analytical capabilities that any organization or person needs in order to fully address the challenges in healthcare (Fig. 1):

1. **Standard reports.** Answer the questions: What happened? When did it happen?
2. **Ad hoc reports.** Answer the questions: How many? How often? Where?
3. **Query drilldown.** Answer the questions: Where exactly is the problem? How do I find the answers?
4. **Alerts.** Answer the questions: When should I react? What actions are needed now?
5. **Statistical analysis.** Answer the questions: Why is it happening? What opportunities am I missing?
6. **Forecasting.** Answer the questions: What if these trends continue? How much is needed? When will it be needed?
7. **Predictive modeling.** Answer the questions: What will happen next? How will it affect my business?

8. **Optimization.** Answer the questions: How do we do things better? What is the best decision for a complex problem?

Let’s use a hypothetical example to illustrate how these various capabilities are developed and used. Christopher Regional Hospital is struggling to understand why its operating margins are decreasing. In particular, the cardiac service line, typically associated with a good contribution margin, has seen a decline over the last six months. Hospital administrators decide to undertake an analytics initiative in order to understand what is causing the decline in contribution margin.

1. **Standard reports.** The problem was identified when the administrators consulted a series of standard reports that the hospital uses on a monthly basis. They realized the cardiac service line’s contribution margin was down from the same period last year.

2. **Ad hoc reports.** These standard reports did not show all of the performance indicators the administrators needed to understand the issues, so they asked for a few specific reports related to the inpatient and outpatient volumes, as well as statistics trending by fiscal period over the last year. These reports showed that patient visits were not down, and the service line was fairly busy, and so the problems were not simply a matter of getting more patients. The administrators asked that these *ad hoc* reports become a part of the hospital’s standard reporting environment.

3. **Query drilldown.** To explore the problem further, hospital administrators asked the head of the cardiac service line to investigate the problem and report back at the next service line committee meeting. He formed a series of hypotheses about potential causes, and then explored a number of other factors related to contribution margin, such as length of stay, payer, and service utilization. The reporting environment allowed him to use a web browser to dynamically divide, subset, and report on these business metrics. During his query drilldown work, the administrator noticed something peculiar—there were a lot of reimbursement denials for patients having a “rule out myocardial infarction (MI)” code. The director checked with the billing department and was told that they are being denied because the hospital will not receive payment for “rule out MI” patients unless they are coded as an observation status and not as an inpatient.

4. **Alerting.** In order to understand what was happening with these patients, the director set up an alert that was sent each time a patient received the “rule out MI” code and had a status of an inpatient. Those alerts were sent to the case manager and the nurse manager of the cardiac unit so that they could monitor the care, to understand what was happening.

5. **Statistical analysis.** The director then decided to analyze the data for this particular diagnostic code. Running a series of statistical analyses on this patient population, he found a correlation between inpatient status, longer lengths of stay, and days of the week the patient arrived in the emergency room. In particular, he noticed that patients who were admitted in an inpatient status, with lengths of stay between 3 and 4 days, who had come to the emergency room with chest pain on a Friday or a Saturday were most of the denials. This finding was reaffirmed when following up on the alerting procedure: the nurse manager of the cardiac unit knew that these patients required a stress test, and stress tests were not performed in the hospital on Saturdays or Sundays. Patients who had come to the emergency room with chest pain on a Friday or a Saturday were most of the denials. Patients were admitted to stay through the weekend to receive their stress tests on Monday; standard procedure was to schedule the stress test within 24 hours, allowing the patients to be placed in a cardiac Intensive Care Unit (ICU) bed and not be admitted. The director also realized when speaking with the nurse manager that more and more “rule out MIs” were taking up beds in the cardiac ICU.
6. **Forecasting.** After hearing the findings in the service line committee meeting, the administrators wanted to know what the impact of this trend would be over a longer time horizon. They constructed a forecast of patient admission, diagnosis, length of stay, and payments over the next 24 months based on the past 4 years of hospital data, census data for their region, and projections from several medical institutions. The forecast showed that their hospital’s patient admissions around this condition were expected to grow 56% each year for the next 2 years due to the recent closing of area hospitals. It also showed that their relatively small drop in contribution margin today could easily grow into a bigger problem in the next 18 months unless they found a way to address both the availability of stress testing on weekends and the problem this condition presents on utilization of beds in the cardiac ICU. This analysis highlighted the patient throughput issues that were just starting to develop, affecting the quality of care for the cardiac patients.

7. **Predictive modeling.** The hospital leaders now wanted to understand the value of more timely treatments and care focused on acuity. The administrators wanted to know whether receiving stress tests within 24 hours, and removing cardiac rule-outs from the emergency room to an observation unit, were beneficial from a quality, safety, and financial perspective. They had staff construct a statistical model that could predict patient outcomes. By using historical patient information, the statistical model could predict the likelihood of death, readmission, disease progression, and long-term costs based on the timing and the treatment that was administered. This model showed opening an observation unit and staffing stress testing on weekends could decrease this condition’s returns to the emergency room visits; reduce medical errors and negative outcomes, including death; and create more cardiac ICU beds, increasing throughput for critical cardiac patients and reducing expenses by staffing by acuity. They also found out something else.

8. **Optimization.** During the predictive modeling exercise, the hospital analysts used data mining software to look for trends in the electronic medical records that might have an impact on outcomes. They uncovered a previously unknown trend: patients under the care of one physician in the hospital were 32% less likely to be readmitted after a cardiac catheterization. The administrators contacted the clinical chief of cardiology and informed him of the analysis. When the clinical chief questioned the physician about her treatment strategies, she indicated that she required her patients to follow up with a cardiovascular exercise and diet program. She had her staff follow up on her patients to make sure they completed the program. Predictive modeling showed that patients that participated in that program, as well as received timely treatment, had 61% fewer admissions in the following 2-year period.

The preceding example illustrates how health insights are derived through a successive series of steps. Each step provides vital information needed to make the next step feasible and effective. It would have been quite difficult, for example, to know what statistical analyses needed to be run if the organization did not already have some direction from the query drilldown. Each step is also more complex than the former, requiring deeper analytical skills, better data, and more coordinated engagement within the organization.

**BUSINESS VS. CLINICAL ANALYTICS**

The numbers and types of analytics that can be applied to healthcare are practically endless, constrained only by the creativity of the human mind to ask intelligent questions and define mathematical inferences. When comparing health and life sciences to other industries’ use of analytics, one characteristic stands out as somewhat unique: the questions, data, and decisions involved include traditional types of business information—sales, operations, etc.—but also include scientific information and interpretation. This distinction brings an additional level of complexity to looking at analytics as a transformative engine for healthcare. Whereas many software tools and people skills applicable to other businesses can be applied equally to health and life sciences, there are a variety of science-oriented capabilities, Fig. 2, that are not as broadly available. And for many business insights in healthcare, it is the combined view of both business and scientific information that enables more educated decisions.

Any analytical solution in health and life sciences exists on a continuum of business-focused to clinically focused analytics (Fig. 2). Some types of insights—assessing profitability of a business unit or providing a standardized financial reporting environment—mainly involve the use of information from business units, systems, and knowledge workers. Other insights, such as the safety dimensions of a drug therapy or the outcomes of a clinical research study, reside more clearly in the realm of patient information and scientific interpretation. Between these two extremes lies an entire continuum of analytical applications that provide a unique view into the operations of a health enterprise, the management of patient populations and diseases, and the primary determinants of costs, quality, and outcomes.
With such diversity and breadth of scope, it is difficult to develop any taxonomy of analytical capabilities that adequately conveys all of the various analytical dimensions of the ecosystem. However, at the highest level, any healthcare-related analytical application can be said to target at least one of three main business imperatives:

1. **Improving organizational performance.** These analytics focus on the financial and commercial performance of the organization. Profitability and performance management are commonly cited issues in this area.

2. **Lowering healthcare costs.** These analytics focus on cost avoidance, active cost management, cost reduction through improvements in efficiency, and other aspects of operational improvement. Detection and prevention of healthcare fraud, as well as activity-based costing initiatives, often fall into this area.

3. **Improving health outcomes.** These analytics focus on improved patient outcomes, including areas of patient safety and treatment efficacy. Clinical research of novel drugs and therapies can be considered a component of this area.

Obviously, there is a close interdependency between these three business imperatives, and it is not uncommon for improvements in one area to impact the other areas. But for an organization considering a new analytics project, these imperatives represent a critical aspect of project scoping. For an analytics program to be successful, absolute clarity needs to exist in the program’s intended objectives. Programs with scopes intending to cover all three imperatives are unlikely to be successful; the data, objective measures, people, and impacted business processes are dramatically different. But for targeted analytical programs, the impact to each of these imperatives can be equally dramatic.

With these imperatives and the business-clinical continuum as context, let us look more deeply into some representative programs.

**IMPROVING ORGANIZATIONAL PERFORMANCE**

At its core, healthcare is a business just like any other industry. Profitability is linked to efficiency and competitive differentiation. Innovation breeds opportunity. Quality can command premium pricing. Customer relationships determine long-term revenue potential. These and many other business principles serve as the underpinning for a portfolio of analytical capabilities that can have a tremendous impact on the quality, safety, and cost of healthcare delivery.

Aside from issues related to clinical outcomes (considered separately in a subsequent section), there are three general categories of organizational performance analytics:

- **Financial:** Analytical capabilities related to revenue, operating costs, and investments. Typical topics include financial management, revenue cycle optimization, and profitability analysis.
- **Operational:** Analytical capabilities related to the way an organization operates internally. Typical topics include utilization management, human resources, and other enterprise-wide competencies.
- **Commercial:** Analytical capabilities related to the way an organization sells, markets, and interacts with its customer and partner base. Typical topics include customer targeting, retention, and sales and marketing effectiveness.

The following table highlights some of the more common performance-oriented analytical scenarios found across the healthcare ecosystem today.
Let us look at one real-world example of an organization using analytics in this way. A large, nonprofit teaching hospital cares for over 40,000 inpatients and 750,000 outpatients every year. Like any large hospital, its databases and reports are incredibly diverse, covering financial, quality, and patient satisfaction information. But it was quite difficult for executives to see the information in a way that would facilitate better decision making. For example, what was the relationship between patient-related services and productivity in the different care units?

The hospital developed a performance management system utilizing balanced scorecard concepts to draw correlations between measurements from the different areas of their organization. By combining data relating to 50,000 patient encounters a year from 29 different sources, the system analyzes and distributes operational metrics to executive management, physicians, and frontline employees throughout the hospital. Whether in a patient care setting or in the business office, about 900 employees at all levels can see how their actions, individually and as a whole, affect the organization and its patients. Over thirty metrics detail key financial information, length of stay, and patient satisfaction measures. Select physicians and nurses even have their own scorecards, enabling them to share information relating to finance, productivity, workload, and quality indicators with their colleagues.

Some department heads now routinely use the system to gauge hospital performance, and spot anomalies relating to length of stay, spikes in certain diagnosis-related groups, and procedural delays. As an example, operating room efficiency and productivity are impacted significantly by start times. If more cases start on time in the morning, then efficiency and productivity increase. So by identifying patterns relating to delays in operating room start times, physicians can have a better understanding of the impact of specific test orders and procedures. Using this approach, the hospital can also identify best practices.

The key to unlocking the value of performance solutions is in the identification and measurement of key performance indicators (KPIs). In this area, healthcare actually has a fair amount of information already available from which to draw. Organizations such as the Joint Commission (http://www.jointcommission.org), the Agency for Healthcare Research and Quality (http://www.ahrq.gov), and the National Quality Forum (http://www.qualityforum.org) have completed extensive work in identifying critical healthcare KPIs, which every healthcare enterprise should monitor. These measures cover the gamut of business and clinical analytics, and are themselves the subjects of entire books. The following table gives a small sample of these types of KPIs:

When considering a performance management program and its associated KPIs, it is important to keep several things in mind:

1. KPIs should be empirically measurable. The idea in performance management is to make better decisions based on real data.
2. KPIs should be linked directly to a business objective. Simply measuring performance with no intended action based on the measurement is a waste of effort.
3. KPIs are the indicator, not the problem or the solution. The goal in performance management is to use KPIs as a detection tool for the business; the root cause analysis of a failing indicator is a separate process often involving different analytical techniques.
4. Measuring and reporting are not the end goal. The reason to use analytics in performance management is not to create easy-to-read dashboards (though good reporting is required). Rather, the end goal is to be able to predict and optimize performance along those specific performance dimensions.

As you might imagine, one of the more significant areas to predict and optimize is costs.
LOWERING HEALTHCARE COSTS

Many people would argue that the concept of lowering healthcare costs is a measure and benefit as opposed to a business imperative. In some sense, this is probably true. Many analytics initiatives related to organizational performance or healthcare outcomes are justified by illustrating the impact on health-related expenditures in R&D, marketing, reimbursements, and other sources. But considering the central role that rising costs are playing in healthcare market dynamics, it is also useful to consider cost reduction programs as a separate topic in its own right. In fact, many organizations place a greater emphasis on cost management initiatives than performance or outcomes, as the organizational “pain” is so acute.

One example of analytics applied in this way pertains to healthcare fraud. Historically, an organization’s ability to detect abusive or fraudulent activity has been limited to solutions that are called rules engines. Rules engines—a simplification of the “alerting” tier in our taxonomy—maintain an inventory of known fraud and abuse schemes, and can draw a sample from the collection of all healthcare transactions to look for situations that violate the rules. There are problems here, though:

1. The rules engine is usually only collecting a sample of records; it is not looking at every transaction. So low-volume fraud transactions are likely to slip through undetected.
2. When a rules engine detects a deviation, it is only detecting a violation of a rule. There are countless reasons why that rule violation might occur that are perfectly legitimate. There is little ability to understand the actual likelihood a given incident is indeed fraudulent, or what the real financial impact of that violation might mean for the organization. As such, fraud investigators spend time investigating a large number of incidents that waste time and money.
3. The volume of false positives (transactions that are identified as fraudulent but actually are not) creates a situation where investigators have very limited bandwidth to investigate fraudulent activities. As such, they tend to focus their efforts on high-dollar

<table>
<thead>
<tr>
<th>Financial</th>
<th>Operational</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Net revenue, profit/loss, and contribution margin per health system, facility, service line, condition, and physician</td>
<td>• Nurse-to-bed ratio overall and per nursing unit</td>
</tr>
<tr>
<td>• Actual and overtime expenses to budgeted per health system, facility, service line, and department</td>
<td>• Attrition overall, per role, and per department</td>
</tr>
<tr>
<td>• Cost per service line, condition, physician</td>
<td>• Throughput red alerts per health system and facility</td>
</tr>
<tr>
<td>• Nursing travelers’ expense per health system, facility, service line, and department</td>
<td>• Length of stay per health system, facility, service line, condition, and physician</td>
</tr>
<tr>
<td>• Ambulatory surgeries per health system, facility, and surgeon</td>
<td>• Percentage of 11 A.M. discharges per health system, facility, service line, nursing unit, and physician</td>
</tr>
<tr>
<td>• Clinic and Emergency Room visits per health system, facility, and clinic</td>
<td>• Number of tests performed by test, health system, facility, service line, condition, and physician</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Quality and Safety</th>
<th>Nursing Care</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Percent of patients who received recommended care, and percent of process measures met for specific conditions by health system, facility, service line, nursing unit, physician, etc.</td>
<td>• Percent of inpatients with a hospital-acquired pressure ulcer</td>
</tr>
<tr>
<td>• Percent of patients who received recommended surgical infection prevention, and percent of process measures met by health system, facility, service line, nursing unit, physician, etc.</td>
<td>• Rate of urinary tract infections with catheters</td>
</tr>
<tr>
<td>• Infection rates for specific conditions by health system, facility, service line, nursing unit, physician, etc.</td>
<td>• Rate of bloodstream infections with catheters</td>
</tr>
<tr>
<td>• Medication errors per 1,000 orders</td>
<td>• Ventilator-associated pneumonia</td>
</tr>
<tr>
<td>• Pediatric IV infiltration rates</td>
<td>• Smoking cessation counseling for acute myocardial infarction, heart failure, and pneumonia</td>
</tr>
<tr>
<td>• Psychiatric patient assault rate</td>
<td>• Number of registered nurses per patient day and number of nursing staff hours (registered nurse, licensed vocational/practical nurse, and unlicensed assistant personnel) per patient day</td>
</tr>
<tr>
<td>• Percent of major surgical inpatients with a hospital-acquired complication and death</td>
<td>• Nursing work life scores related to participation in hospital affairs, foundations for quality of care, manager ability, staffing and resource adequacy, and collegiality of nurse-physician relations</td>
</tr>
<tr>
<td>• Inpatient falls per inpatient days with and without injuries</td>
<td></td>
</tr>
</tbody>
</table>
transactions. Fraud and abuse schemes consisting of low-dollar transactions are unlikely to trigger an investigation.

4. In order for a rules engine to know about a fraudulent activity, the activity must have been previously detected and codified into a rule. So rules engines always suffer from a delay between someone finding a fraud scheme, someone else coding into a rules system, and when organizations then start implementing the updated rules. These delays can span from months to years. And even a slight modification to an existing fraud scheme can render its existing rule useless.

5. Generally speaking, there is a complexity ceiling for rules engines. Abuse patterns that do not easily lend themselves to the relatively simple structure of a codified rule—for example, a complex matrix of collusion—are not easily detectable by rules engines.

So the smart fraudster is one that generates low-dollar transactions at a relatively low and distributed volume so as to not show up in the rules engine sampling. Of course, there will always be a pattern to the activity, but as long as the volume doesn’t look suspicious and the dollar value is low, the likelihood of detection remains small. Even if the fraudulent scheme is detected, by the time it becomes a rule deployed within healthcare institutions, the fraudster will be already working the next scheme.

The use of analytics for fraud detection is not encumbered by these shortcomings. An analytical model looks for patterns, not rule deviations. As opposed to sampling a small group of transactions, an analytical model can be applied to every single transaction flowing through a healthcare system, even in real time. It is less likely to produce false positives, because the detection is not based on some arbitrary comparison criteria, but rather on a model. Indeed, an analytical model can even be applied to the targets identified by rules engines to help separate the real signals from the noise. Note also that the delay between fraud identification and deployment of a detection process is much lower, because the analytical model can separate the real signals from the noise. Note also that the delay between fraud identification and deployment of a detection process is much lower, because the analytical model can be “tuned” in real time based on real-world experience.

Analytical fraud models also offer the ability to detect fraudulent or abusive patterns, whose complexity is larger than what can easily be understood by the human mind. For example, a collusion scheme might actually involve three people working in four different organizations and generating six different types of claims. It is quite difficult for an investigator to ascertain a pattern across so many variables and permutations. But by applying data mining algorithms that examine each of the variables, relationships, and interdependencies, a pattern of social “connectedness” between these three people can emerge that is markedly different from those of their cohorts.

Fraud is only one example of the potential contribution of analytics to cost management and reduction. Other examples include:

- Portfolio optimization—Ensuring that high-value products and services are progressed, and stopping investments in products and services predicted to be unviable or unprofitable.
- Product development and pricing—Balancing market opportunity to pricing strategy in order to maximize profit (not necessarily price), minimize leakage, and ensure competitiveness.
- Case management/readmission prediction—Identifying and targeting patient cost factors and proactively introducing interventions to avoid the costs.
- Activity-based costing and management—Understanding where an enterprise is actually incurring costs in specific activities, as well as identifying and predicting the outcome of potential changes.

Regardless of the specific focus area, all of these analytical opportunities revolve around a common concept of using prediction to maximize investments in the right areas of the business, and minimize exposure to areas that extend costs without sufficient upside.

Let’s look at another real-world example. One of the Blue Cross and Blue Shield insurance providers in the United States sought to gain a better understanding of costs and cost overruns within its organization. Traditional accounting systems provide departmental budgets, but beyond budget line items, those systems typically lack the analytical sophistication needed to look at financial information along different cost dimensions: corporate vs. departmental processes, predicted vs. accrued costs, etc. In this particular case, the company’s finance managers could perform basic budget variance at the corporate and department levels, but they could not break out costs by product or individual lines of business, or even scrutinize contributions to activity costs at the program level.

The lack of process-oriented insights, which is far from unique to this particular insurer, has several ramifications. First, there is no way of ensuring that a given customer contract would be profitable. In one case, executives believed the company was losing money on a particular contract, but had no comprehensive way to evaluate all of the costs that supported this agreement. Profitability calculations can be very complex, as the way that costs are accrued (i.e., by individuals) differs significantly from the way that pricing is structured (e.g., by family, employee, child, etc.). Second, this lack of visibility at the contract level is further complicated by the fact that the administrative costs of plans differ between large group plans, small group plans, individual
products, and government plans. And third, even if a single contract is unprofitable, unless there is a mechanism to ensure cost management on an ongoing basis, future contracts will suffer similar problems.

To address these concerns, the insurer adopted a particular analytical framework focused on activity-based costing and management. This solution pulls data from its corporate and departmental account systems, enabling planners to explore and predict cost data at all different levels of the business. For many programs, budgets need to actually be set at the individual business activity level. As opposed to simply watching budget reports for cost deviations, the analytical solution allows them to actively monitor projected costs and compare actual costs with the allowable rates for each activity. It becomes much easier to recognize that potential cost overruns and institute changes to the company’s activities fall within acceptable cost parameters.

Moving from reactive to proactive, this same approach is also used to help managers decide when to modify the pricing plans for contracts, or when to restructure expenses by outsourcing or consolidating tasks. This concept provides real value when considered at the enterprise level. Departmental budgets often mask the costs associated with cross-functional business activities such as corporate sustaining activities, sales and marketing, and product service and support. Using this approach, an organization can even establish and measure corporate benchmarks as part of a broader transformational strategy.

Ultimately, the solution allows the company to evaluate costs for every business process and to identify areas for implementing cost-saving measures and company-wide process improvements. Marketing and underwriting can see how product costs vary depending on the size and type of the group plan; department managers can see what activities performed truly at cost; and executives can look at administrative costs for specific lines of business as a ratio to revenue, claims paid, or other financial metrics.

In this scenario, notice the many different levels of costs being impacted:

- Budget management—Ensuring programs and business units stay to budget.
- Activity management—Ensuring efficiency and consistency in execution.
- Work reduction—Spending less time manipulating data, improving decision speed.
- Profitability—Ensuring that contracts and pricing are actually profitable.
- Cost arbitrage—Facilitating outsourcing to lower-cost fulfillment options.

The breadth of scope and impact is one reason why reducing healthcare costs can stand on its own as a business imperative, and not simply reflect the outcome of other improvement programs.

**IMPROVING HEALTHCARE OUTCOMES**

Beyond any of the organizational productivity or cost issues presented so far, the Holy Grail in healthcare analytics is all about improving the health of patients. Though providers, payers, researchers, manufacturers, and policy makers can disagree on most other aspects of the healthcare ecosystem, everyone can agree that we need more health in healthcare. Analytics provide the means of finding that health, and bringing it to scale.

Health-oriented analytics are not new. Every prescription drug discovered and marketed was required to demonstrate safety and efficacy via statistical models before being approved for use. Epidemiologists rely on computational methods in studying the progression of diseases in populations. Analytical models and methods are used to explore the estimated 3 billion chemical base pairs and 20,000 genes in the human genome. In truth, advanced analytics have been a long-standing tool in understanding the scientific bases of biology and medicine.

Despite this analytical heritage, the hardest health-related questions—the ones that will likely have the biggest impact toward the lives of patients—are just now rising to the surface. Why? There are several reasons:

1. **Scientific advancement.** Our understanding of the fundamental nature of human physiology, genetics, and disease is reaching a point where we can more directly apply the learnings toward making patients better.
2. **Convergence.** The health-related improvements needed in any one health market—providers, payers, researchers—require access to information in the other markets. As greater transparency and openness to collaboration unfold, new opportunities to apply advanced analytics arise as well.
3. **Cost.** The dramatically rising costs of healthcare around the world have brought a new level of discrimination in treatment efficiency and effectiveness. Rather than continuing to fund expensive treatments with unclear outcomes, a better fiscal policy would focus on funding the best treatments with the best cost structures.

So what types of opportunities exist now? The following table highlights some of the main categories of outcomes-oriented analytics. It is important to note that this list is just a sampling. The ability to positively impact the treatment programs and corresponding patient outcomes is limited only by the creativity of the
human mind to develop novel ideas and test them. As you will see in the examples that follow, there are many different ways to look at health outcomes.

Consider the following real-world example. A professor of epidemiology at a large Canadian university is studying the risk factors and relationships between cardiovascular disease and cholesterol levels, lifestyle, diet, childhood experiences, and environmental factors. The research includes studying the incidence and distribution of diseases by combining massive amounts of data from many disparate sources: nutrition databases, healthcare delivery data, community demographic information, national patient surveys, and information supplied by Health Canada and the U. S. Centers for Disease Control and Prevention. By bringing these large, disparate information resources together, the professor has been able to develop statistical models that establish the relationship between high cholesterol intake at an early age and chronic diseases later in life. For example, the evidence indicates that cardiovascular disease has a long latency time, and that problems begin in early childhood. By looking at a variety of factors such as cholesterol intake, height, weight, and level of physical activity, the professor can predict childhood risks of adult cardiovascular illness and death.

So how can analytics like that be put into action? In one case, a large U. S. health services company uses data mining to help identify people who would benefit from preventive care services. The data mining software predicts hidden relationships in millions of member records to 1) determine patient risk levels; and 2) develop more targeted intervention and prevention plans. Large volumes of clinical and operational data are applied in statistical models that can predict the members in greatest need of support programs. In addition, by identifying high-risk patients and implementing preventative actions against future conditions, health problems and treatments can actually be avoided.

In another case, a western European health research institute sought a way for doctors and patients to predict problems and treatments can actually be avoided. To do this, they developed two analytics-based solutions: one

<table>
<thead>
<tr>
<th>Prospective Performance and Intervention</th>
<th>Optimized and Targeted Treatments</th>
<th>Adverse Event Avoidance</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Disease and population management</td>
<td>- Biomedical informatics</td>
<td>- Adverse event detection and prediction</td>
</tr>
<tr>
<td>- Expedited research</td>
<td>- Clinical and patient decision support</td>
<td>- Biosurveillance</td>
</tr>
<tr>
<td>- Member and patient programs</td>
<td>- Clinical program compliance</td>
<td>- Clinical alerting</td>
</tr>
<tr>
<td>- Pay for performance</td>
<td>- Clinical research analysis</td>
<td>- Clinical quality performance analysis and reporting</td>
</tr>
<tr>
<td>- Provider performance</td>
<td>- Evidence-based clinical decision support</td>
<td></td>
</tr>
<tr>
<td>- Provider/physician profiling</td>
<td>- Clinical performance management</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Patient adherence</td>
<td></td>
</tr>
</tbody>
</table>

enables rheumatism patients to examine the factors that affect their health, and the other helps rheumatologists select the best treatment for each patient. A statistical methodology called factor analysis is used to test various treatment hypotheses with the support of randomized trials. The goal is to find the scientifically best combination of various lifestyle factors for each individual case. Patients choose what they would prefer to improve, such as pain, and up to four independent lifestyle factors, such as exercise, diet, and sleeping habits, which might conceivably reduce pain. The system then generates a test plan that the patient prints out as a diary of what should be tested each day. At the end of the period, the program calculates which combination of factors provided the best lifestyle impact.

Over 50 clinics have been involved in the collection of these data from patients and physicians: more than 26,500 patients and 144,000 consultations have been collected and analyzed. The predictive model uses patient data as the basis for predicting the results of various treatments such as different drugs and drug combinations. In other words, instead of a physician relying on a generic treatment plan that may or may not be the best option for a particular patient, the system assists doctors in choosing the best treatment for each patient based on that patient’s profile. This approach has the added benefit of avoiding the costs associated with expensive treatments or drugs that are unlikely to be a good fit for a given patient.

Unfortunately, in many cases, humankind has yet to find cures or even strong treatments for diseases. It is in these areas where analytics can have a dramatic impact. For pharmaceutical and biotechnology companies, the development of a new medicine can take more than a decade—a decade of great expense for the developer and of prolonged anxiety for patients in need of new treatments. Meanwhile, the escalating cost of research and development, accompanied by the increasing complexity and expense of human clinical trials, threatens pharmaceutical innovation and drives up healthcare costs.

Using analytics, it is possible to develop simulations of human clinical research (clinical trials simulations)
that can considerably expedite the development of novel therapies, and save millions of development dollars. Simulated clinical trials are virtual replicas of actual clinical trials. The simulations take data models developed from actual clinical trials and develop clinical scenarios and putative trial results that take into account variability caused by treatment effects, survival times, adverse events such as the occurrence of headaches or nausea, and other events that occur during trials. Researchers might perform 10,000 simulations of a single scenario, which means an entire trial can generate millions of observations. Using this approach, a U. S. biotechnology company is actively exploring new treatments for HIV infection, hepatitis, various forms of cancer, inflammatory diseases such as rheumatoid arthritis, and many others. Its solution, which also takes advantage of grid computing technology, can produce as many as 1 million patients for 2 terabytes of data in a single simulation. The company has the ability to do over 10,000 replicates in an hour, and that rate is growing.

**SUMMARY**

With all of the different dimensions of analytics covered so far, it is logical to ask where an organization begins in its journey toward deeper insights. As you might expect, there is no single correct answer—the journey depends on the traveler and the desired destination. In watching many organizations move through their analytical maturity process, one key theme emerges among successful companies: start where you are. As mentioned earlier, it is easy to fall prey to analysis paralysis, constantly evaluating abilities instead of instituting programs to change those abilities. You will never have enough electronic data, it will never be clean enough, there will always be conflicting priorities, and you will never have enough of the right resources to focus on the project. But in all likelihood, you have enough of all of the above to get started.

Another trap some organizations fall into is “boiling the ocean.” It is impossible to take on all of these analytical challenges and opportunities at the same time. And even if you could, it would not be advisable. Organizations and their knowledge workers need time to learn, determine what makes sense for their company, find the pitfalls and build bridges over them, and institutionalize those learnings for future projects. The intelligent enterprise identifies a significant, clearly defined business challenge and puts an agile team together to demonstrate success. With a successful project completed, the ability to launch the next initiative is much easier, and the people involved in the first project are now tour guides on the journey.

Using this approach, leading organizations often create an analytical center of excellence (ACE) within their enterprise. An ACE consists of a small, cross-functional group of people who help the many different parts of the company successfully leverage analytics. ACEs provide several benefits:

- Provide internal consulting and training on how to apply analytics to business problems
- Serve as business sponsors in the identification and management of data integration, quality, and management activities that support analytics
- Establish clear accountability for the advancement of analytical competencies within the enterprise
- Ensure consistency in the selection and use of enterprise architecture, analytical tools, and solutions to support analytics across the enterprise

Whether you are starting your first analytics initiative or finishing your 100th project, you hopefully are getting a sense of the curiosity and excitement many people feel about analytics: there is always another question to ask with hidden gems in the answers. In healthcare, we are just at the beginning of seeing how analytics—whether biased toward the business or clinical sides of the spectrum—can transform our ecosystem. Analytics will give us better guidance on how to control costs—not just line items, but the hidden and true costs of healthcare. Analytics will help us identify and dismantle old assumptions about the way healthcare is delivered—not relying on gut instincts and hearsay, but real evidence. Analytics will allow us to determine not only the treatments that produce the best outcomes, but the real factors that determine optimum treatment efficacy and cost. The hidden gems in healthcare will surface, and with them will come better lives for patients everywhere.

**ACKNOWLEDGMENT**

I am grateful for the contributions of Cindy Berry and Rick Pro, both from SAS’s Health and Life Sciences organization, to the content and review of this entry.