Health care subrogation:

combining science and art to maximize results

Ingenix is now OptumInsight™, part of Optum™—
a leading health services business.

Combating rising costs with subrogation

Often the success of a health plan or claim administrator is measured by how fast claims are paid once they are submitted. The additional time and expense associated with any claims cost management program must be weighed against the benefit of such programs. Accident-related claims typically account for 8–10 percent of claims volume. Reviewing claims for potential accidents, determining if other coverage exists, and coordinating benefits with accident insurers may not be the primary concern of a claims administrator. This can be an expensive oversight, particularly in today's economy, as total national health expenditures are expected to rise at a rate of two times the rate of inflation for the next decade, reaching a projected $4.3 trillion by 2017.[1]

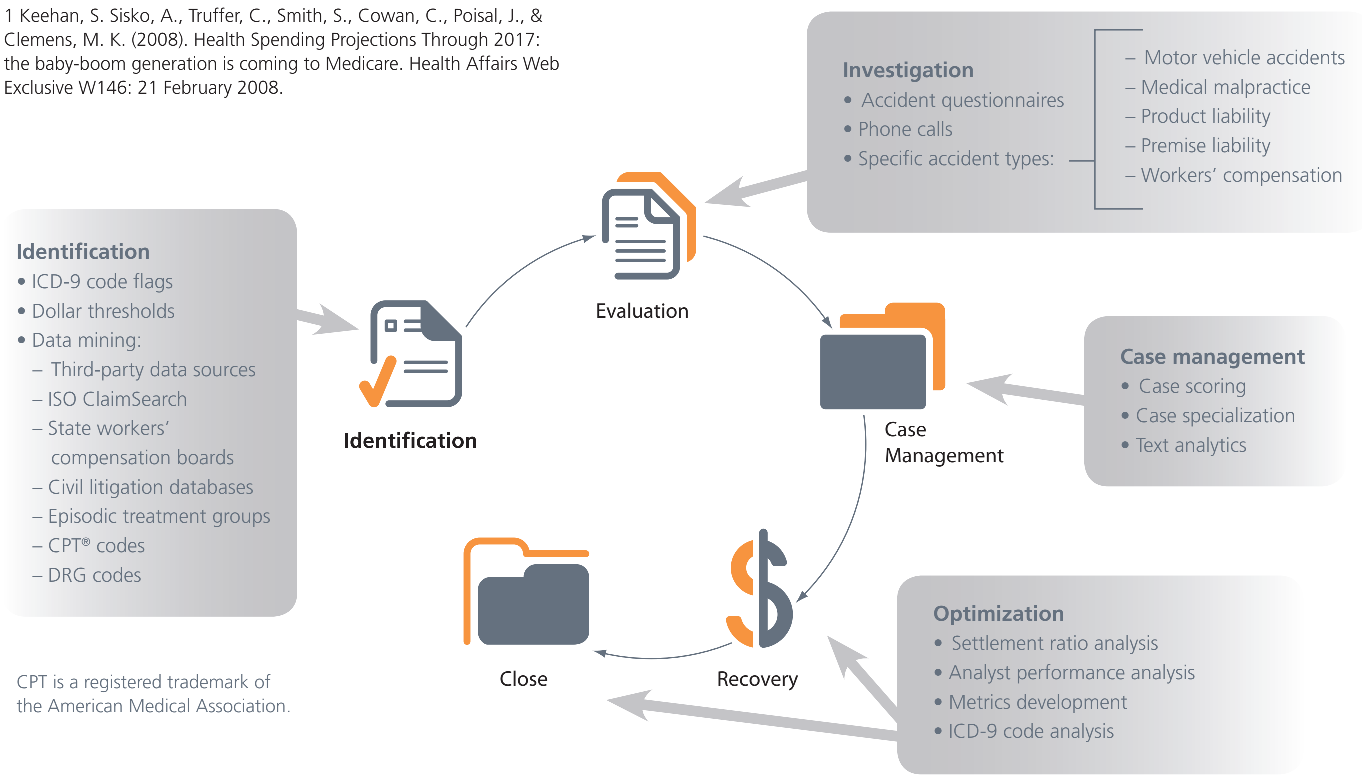
Containing health care costs is one of the most critical issues facing our country today. It has become a campaign platform for our government leaders, as the rapid increase of health care costs places everyone at extreme financial risk. Health care subrogation is just one of the weapons in a health plan's arsenal to combat the rising costs associated with financing accident-related health care claims. The most successful subrogation techniques combine the science of data analytics with the art of an individual's business acumen, knowledge, and intuition to identify and recover accident-related health care claims.

Subrogating health care claims

The rights of subrogation and reimbursement are provisions documented in one's health care policy. A technical definition of subrogation is the substitution of one person in place of another, with reference to a lawful claim, demand, or right, so that the one who is substituted succeeds to the rights of the other, in relation to the debt or claim and its rights, remedies, or securities. In layman's terms, subrogation means that a health plan can be reimbursed by the party whose insured was deemed responsible for the accident or injury. For example, having auto insurance or workers' compensation coverage pay for medical claims resulting from an accident, thus saving the medical insurance carrier from incurring inappropriate costs.

An effective health care subrogation program first requires the identification of accident-related claims. Subrogation teams mine post-adjudicated claims to find accident-related codes on claims and conduct an investigation that will ultimately generate recoveries from a third party for their clients. There are more than 20,000 injury and disease classification codes, or ICD-9 codes, and as many as 3,700 of them suggest a connection to an accident. These codes are related to motor vehicle accidents, medical negligence, work-related accidents, premise liability, and defective product injuries.

1 Keehan, S. Sisko, A., Truffer, C., Smith, S., Cowan, C., Poisal, J., & Clemens, M. K. (2008). Health Spending Projections Through 2017: the baby-boom generation is coming to Medicare. Health Affairs Web Exclusive W146: 21 February 2008.



CPT is a registered trademark of the American Medical Association.

The accident details are investigated along with contract language, applicable state or federal law, other insurance coverage, as well as additional theories of liability, to assess the probability of a successful recovery. If accident-related claims have already been paid by the health plan, the subrogation professional will negotiate and even litigate to secure reimbursement on the health plan's behalf.

All of these activities—from identification through optimization—are time and labor intensive, and many different statistical and analytic techniques can be used to facilitate each activity.

Telling the story

Good statistics involve principled argument that conveys an interesting and credible point. It should tell a story that an informed audience will care about, and it should do so by intelligent interpretation of appropriate evidence from empirical measurements or observations. Statistical analysis has a narrative role that allows us to sharpen the point to our story.[2]

Context is one key to the appropriate use of statistics. Let's use a simple example and ask the question, "how many motor vehicle traffic crashes occurred in the United States in 2007?" According to the National Highway Traffic Safety Administration's 2007 National Statistics report, the total number of traffic crashes in 2007 was 6,024,000. That number is generally accepted as fact, but what does it mean? Adding more data to the finding provides additional insight:

Year	Total crashes	Fatal crashes	Injury crashes	Property damage only
2006	5,973,000	38,588	1,746,000	4,189,000
2007	6,024,000	37,248	1,711,000	4,275,000
% Change (YoY)	0.85%	-3.60%	-2.05%	2.01%

The data reports that there were more accidents in 2007, but fewer fatalities and injuries. The questions that these numbers may trigger are:

- What factors might have contributed to an increase in total crashes?
- What changed from 2006 to 2007 to see a 3.60 percent decrease in the number of fatal accidents?
- Which types of accidents are more likely to result in a fatality or injury?

These questions might lead to the gathering of even more data (see table, bottom of page):

Although the total number of accidents decreased from 2006 to 2007, motorcycle fatalities and injuries showed a dramatic increase, yet overall fatalities dropped by 3.60 percent. Without an appropriate comparison or benchmark, a single number (i.e., 6,024,000 total traffic accidents) holds very little value on its own.

When we add data that allows us to make a comparison, the subsequent questions to "how many motor vehicle traffic crashes occurred in the United States in 2007" provide a deeper level of insight as well as an opportunity to affect change that may save lives. Extracted data is only as good as the logic used to gather it.

Analytics defined

Analytics is defined as "the science of analysis." Analytics is the extensive use of data, statistical, and quantitative analysis, explanatory and predictive models, and fact-based management to drive decisions and actions. Analytics are a subset of business intelligence: a set of technologies and processes that use data to understand and analyze business performance.[3]

The business intelligence questions being answered with analytics are typically: "why is this happening, what will happen next, and what is the optimal outcome?" But before one gets to

Traffic crash victims	Killed in 2006	Injured in 2006	Killed in 2007	Injured in 2007	% Killed change YoY	% Injured change YoY
Drivers	22,830	1,666,000	21,647	1,571,000	-5.46%	-6.05%
Passengers	9,156	709,000	8,657	692,000	-5.76%	-2.46%
Motorcycle Riders	4,810	88,000	5,154	103,000	6.67%	14.56%
Pedestrians	4,784	61,000	4,654	70,000	-2.79%	12.86%
Pedal Cyclists	733	44,000	698	43,000	-5.01%	-2.33%
Unknown	289	7,000	249	11,000	-16.06%	36.36%
Totals	**42,602**	**2,575,000**	**41,059**	**2,490,000**	**-3.76%**	**-3.41%**

[2] Abelson, R. Statistics as Principled Argument. Hillsdale: Lawrence Erlbaum Associates, 1995

[3] Davenport, T. H. & Harris, J. G. *Competing on Analytics: The New Science of Winning.* Boston: Harvard Business School Press, Boston, 2007.

those higher-level questions, often we must start with more basic investigative questions around "who" and "what." Common analyses include:

Descriptive analysis—Describes a data point's behavior by relating it to the "central tendency" of the data sample. This answers the questions of "what happened, how many, how often, and where?" Descriptive statistics enable the business to "benchmark" or standardize their measurements to determine whether increases or decreases are occurring and what they mean to their business processes and the impact on results. Common measures of central tendency include:

- **Mean**—The sum of the observations divided by the number of observations. This provides an "average" of paid claims per subrogation case, for example
- **Median**—The middle observation in ranked list of observations—exactly half of the observations are below and the other half are above
- **Mode**—The most frequently occurring observation

Trend analysis—Gathers information to spot trends or patterns in the data. The data used can be actual counts, averages, or transformed variables used to represent the data in a meaningful way. Trend analysis helps create a context and relationship between seemingly independent data points and gives a picture of behavior or performance over time that will help to inform business decision makers.

Correlations—A measure of the proportional relationship between two variables. A -1.00 is a perfect negative correlation (as one variable goes up, the other goes down) and a +1.00 is a perfect positive correlation (both variables move in the same direction). A correlation does not mean causation, but a well-designed experimental study allows for the inference of causation. Correlations can be extremely useful in an examination of how different case variables affect each other—like which investigative activities have the greatest influence on recoveries.

Predictive analysis—Uses historical data to determine the likelihood of future behavior. Examples of the techniques used to perform predictive analysis are: linear regression models; logistic regression; time-series models; survival or duration analysis; classification and regression trees; and machine learning techniques such as neural networks. One of the most well-known applications of predictive analysis is credit scoring, used by the financial sector to predict whether creditors will repay their debt. There are applications for credit scoring in subrogation, but instead of predicting whether a debt will be paid, it is used to predict the outcome or quality of the claims under investigation.

Data mining—Data mining is another vital component of case identification and information gathering and has become a standard within the industry. The definition of data mining is hard to pin down in that the multiple definitions focus on what it is—from the process of sorting through large amounts of data and picking out relevant information to the science of extracting useful information from large data sets or databases[4]—with less focus on the practical application of "how" it can be applied to a specific business scenario in a meaningful way.

Science vs. art

The case for science—Science involves rigor and discipline, validity, and reliability, as well as the testing of hypotheses. Utilizing a scientific method of investigation and introducing both independent and dependent variables into our study enable us to attribute a change to a specific set of variables that have been manipulated.

Advanced analytic techniques, such as statistical analysis, data mining, and predictive analytics, allow researchers to draw conclusions from the data and facilitate the forecasting of results and the optimization of processes. For example, many subrogation teams use an accident questionnaire that is sent to the member to get additional information about whether an accident occurred and whether a third party is liable for the cost of the medical claims paid out by the health plan. Suppose we would like to change the content of the accident questionnaire. What types of questions would we need to ask (and answer) to assure ourselves that we would be making the best decision? Some of the questions we might ask could include:

- What are the independent variables (the variables being changed)?
- What are the dependent variables (the variables being measured)?
- What sample size should be used to have enough power to detect an effect?
- How should the sample be selected to ensure the results can be generalized to a larger population?
- How are the results to be measured?
- Is there a baseline or benchmark to be used?

[4] Hand, D., Mannila, H., & Smyth, P. (2001). Principles of Data Mining. MIT Press, Cambridge, MA.

An important component within the science of data analysis is trying to avoid decisions based on human assumption. The human mind tends to suffer from a number of well-documented cognitive failings and biases that distort our ability to predict accurately. We tend to give too much weight to unusual events that seem salient. In context after context, decision makers who wave off the statistical predictions tend to make poorer decisions than a statistical model.[5] These human limitations make it important to automate components of the data extraction and analysis processes to produce consistent results.

A 2004 Opinion Research Corporation survey of executives found a clear opportunity to automate and improve decisions. Operational decisions were high-impact, but only a fraction had been automated, and maintenance costs and time-to-market were real problems. Specifically, the findings included the following:

- More than 90 percent felt that front-line operational decisions affected profitability
- About half had not yet automated about 25 percent of these decisions and nearly 80 percent still lacked automation for more than half of them
- Making the right decisions with a high degree of accuracy and precision was very important to more than 90 percent of respondents; more than 60 percent also rated consistency of decision-making and effective management as important
- 85 percent expressed difficulty in changing decision-making criteria in their systems, with more than 50 percent saying it took months to get business changes implemented in production
- 60 percent felt that automated decisions were carried out inconsistently or they were forced to deal with redundant logic in multiple systems to prevent this problem
- 70 percent of chief information officers/chief technology officers (CIOs/CTOs) didn't believe they were getting the most value they could from their data. The most common reason was that their inability to blend business rules with data prevented them from maximizing the value of their data

Using data to validate and automate decision making can provide a consistency and reliability that cannot be achieved with human decision making alone. Additionally, automating processes or tasks can increase efficiency and productivity, freeing your decision makers to focus their attention on activities that have an impact on maximizing results.

The case for art—Art is creative. It can be spontaneous, but also flexible and organic. It is influenced by the characteristics of the artist and is an expression of what is important and salient to that person. Knowing what questions to ask can be a science unto itself. Business experience and acumen, as well as intuition, play an important role in designing both the questions to be answered with your data and knowing which follow-up questions to ask once you have your data. In the accident questionnaire study mentioned previously, some of the questions that lend themselves to artistry are:

- Are there generational differences in response rates?
- What personality styles are more likely to respond to a questionnaire?
- How many questionnaires should be sent for optimal response?
- Are there regional differences in the answers received?
- What affect would the change have on members' response rates?
- What affect would the change have on whether a case is considered for additional investigation or litigation?
- Is the direction (positive or negative) of the change an important consideration?
- Would either outcome be acceptable (or unacceptable)?

Social science methods, such as the use of control and experimental groups, can be used to determine the effect of a change in business process prior to implementing that change on a larger scale. This allows the business to conduct an experimental study to determine the real impact before incurring any of the costs associated with making that change. By testing hypotheses and using real results to inform our decision-making, negative impacts are mitigated and the real value of a proposed change can be assessed more accurately.

Conducting a well-designed study of a particular business problem is critical to providing the correct context for your results. Both business (art) and analytics teams (science) must work together to provide the greatest impact, as the person providing the data may not have the complete operational picture. A typical study usually has the following components:

- **Define the problem statement or question**—What is being measured? Why is it important? What is the data going to be used for?

[5] Ayers, I. Super Crunchers: Why Thinking-By-Numbers Is The New Way To Be Smart. New York: Bantam Dell, 2007.

- **Population**—Determine the appropriate sample size to ensure reliability and validity of results as well as having enough power to detect an effect (if there is one)
- **Data extraction**—Determine the filters and parameters required to obtain the optimal data set
- **Data scrubbing**—"Sanitizing" the data by replacing or removing missing values, determining a strategy for handling outliers
- **Summary statistics**—Measures of central tendency that describe your entire data set
- **Descriptive statistics**—Correlations or trend analysis
- **Predictive statistics**—Regression techniques; time-series modeling, neural networks
- **Pilot/test results**—Conduct a pilot study, ideally with experimental and control groups, to determine if the change implemented produces measurable, statistically significant results
- **Implementation**—Make changes to code or reports; if process changes are required, training may also be necessary
- **Documentation**—What analyses were conducted? What was found? How were the results utilized, can the analysis be used again in a different study?
- **Evaluation**—Are the results consistent over time?

Once your analysis is underway, it is important to continue to ask these additional questions:

- Have all nuisance variables been controlled?
- Are the calculations correct?
- Is the data correct?
- What should you do with outliers? Are they real outliers or extreme data points?
- Are your assumptions correct?
- Are there other variables exerting an influence on the results?

Making informed decisions is difficult enough. Add the availability of terabytes of data, multiple data sources, and ambiguity to the equation and the decision-making process becomes even more complex. Humans, or "experts," are still required to make decisions about the factors studied and what they mean, both qualitatively and quantitatively. Several studies report the most accurate way to exploit traditional expertise is to add the expert evaluation as a factor in the statistical algorithm.[6]

[6] Ayers, I. *Super Crunchers: Why Thinking-By-Numbers Is The New Way To Be Smart*. New York: Bantam Dell, 2007.

Combining science and art

Combining your algorithms and equations (science) with traditional expertise (art) produces the best results. In a survey conducted by Teradata in 2004, 75 percent of the senior executives of top U.S. companies said that the number of daily decisions has increased over the past year, and more than 50 percent said that decisions are more complex this year than last year. The overwhelming majority of respondents—more than 70 percent—said that poor decision-making is a serious problem for businesses. The top casualties of poor decision-making are profits, company reputation, long-term growth, employee morale, productivity, and revenue.

Just because you have numbers or statistics to support your analysis, it does not automatically mean you have "good" or meaningful numbers. It is the responsibility of people to use critical thinking skills to evaluate the data for validity and reliability. We should not take information at face value. Results should make us ask who, what, when, where, why, and how questions.

In our motor vehicle traffic crash examples earlier, some of the additional questions that we should ask to better understand the data include:

- What does the data mean?
- Why did the number of motor vehicle traffic accidents decrease from 2006 to 2007?
- What factors contributed to the change?
- What other factors might have caused the decrease in fatalities year over year, even with an increase in motorcycle fatalities?
- Do all areas of the country show the same trends?
- Did all areas of the country have the same rate of increase in motorcycle fatalities and injuries?

In the context of subrogation, there are many factors that influence the outcome of a case. As mentioned previously, contract language, state and federal law, as well as insurance coverage play an important role in whether a case will result in a recovery or not. The ability to fully investigate a subrogation case is often dependent upon a member's response to an accident questionnaire. Many factors can affect whether the investigator will receive a response, including personal characteristics and socioeconomic factors.

Personal characteristics include things such as age, education, and gender. Socioeconomic factors include income, regional unemployment rates, and community involvement. For example, Medicaid populations tend to be lower income, more transient, and typically have a lower response rate. In contrast, Medicare populations are often older, often retired, and typically have a higher response rate. Additionally, there may be a generational effect that has an impact on whether a response will be received or not. For the first time in history, we have four generations in the work force at the same time.

Those four generations are:

- Traditionalists .... Born before 1946
- Baby boomers .... Born 1946–1964
- Generation Xers .... Born 1965–1981
- Generation Y or Millennials .....Born 1982–2000

Traditionalists tend to be loyal, feel it is important to "pay their dues" and may be more likely to respond simply because they were requested to respond. Baby boomers tend to be optimistic, idealistic, and don't share the same company "loyalty" as a traditionalist. Generation Xers tend to be skeptical, self-reliant, and impatient. Generation Y or Millennials tend to be realistic and have a sense of entitlement.[7]

All of these factors—personal characteristics, socioeconomic factors, generational differences, and many more—can exert an influence on our data when they occur independently. When these factors are combined, they may have another effect entirely. Analytics can facilitate the discovery of previously unknown relationships within our data.

There should never be a shortage of questions to ask ourselves or about our processes, such as:

- How can we improve case detection?
- How can we maximize subrogation recoveries?
- What factors influence response rates?
- What characteristics of a case would improve open rates?
- Which metrics would drive improved analyst performance?
- Which process improvements are most likely to increase revenue?

A subrogation case lifecycle brings a nuance to the application of analytical decision-making that is not present in many other types of businesses. Because the lifecycle of a case is dependent upon many external factors, such as court schedules and attorneys, it becomes imperative that we create tools and methodologies that allow the business to get ahead of the case lifecycle in order to make timely decisions that have a positive impact on results.

Maximizing subrogation results

Health care subrogation is in a position to impact health care in a dramatic—and positive—way. Analytics can have a powerful role in any organization, from proactively monitoring client results for problems and addressing issues before they become a cause for attrition, to retroactively mining data for patterns and additional revenue opportunities.

Analytics, both descriptive and predictive, are a necessity to ensure that decisions are based on more than intuition or a good idea. The leakage of social science methods from academics to the world of on-the-ground decision-making has had a tremendous impact on decision making in the business world. It is difficult to prove to ourselves that a change in our subrogation processes is prudent when the results often cannot be measured for many months. Using well-designed experimental studies—in combination with historical data—to make an educated decision based on the probability of an outcome, places the business operations manager in a more tenable position.

The use of data mining, analytics, and predictive modeling isn't the solution, it's just the start. The best subrogation teams will differentiate themselves in the marketplace with their ability to combine advanced analytics techniques, statistics, predictive modeling, and automation (science) with expert knowledge, intuition, experimental design, and interpretation of results (art) to maximize their results.

About the author

Marcia L. Moberg, Ph.D.

Dr. Moberg joined Optum as an operations research analyst in 2007. The SubroAnalytics® team she leads uses predictive analytics to identify additional subrogation and savings opportunities and provide data that enables the business to make better decisions. Moberg graduated from Minot State University with a BS degree in elementary education and a BFA degree in psychology. She holds an MS degree in psychology and a PhD in general experimental psychology from the University of North Dakota. Her experience in academia, as well as the high-tech computer and health care industries, has enabled her to incorporate academic principles, such as research and training, into business operations to help transform business processes and applications.

In addition to her current operational research, reporting, and project management responsibilities within subrogation, she has experience in curriculum development, process design and re-engineering, process implementation, and database and program

[7] Lancaster, L & Stillman, D. *Generation Gaps.* Twin Cities Business Monthly, Jan. 2002.

management. She has also presented at regional and national psychology conferences and trained national and international operations and sales teams.

Dr. Moberg is a member of the National Association of Subrogation Professionals (NASP) and serves on the NASP Benchmarking Committee for the 2009 Health Care Study.

About Optum

Optum is an information and technology-enabled health services business platform serving the broad health marketplace, including care providers, plan sponsors, life sciences companies, and consumers. Its business units—OptumHealth™, OptumInsight™, and OptumRx™—employ more than 30,000 people worldwide who are committed to enabling Sustainable Health Communities. For more information, visit www.optum.com.

OPTUM™ www.optum.com

13625 Technology Drive, Eden Prairie, MN 55344