# Improving Patient Care and Clinical Outcomes by Analyzing Healthcare Data for More Actionable Results

## The Problem with Traditional Data Analysis Tools

It may be an understatement to say that the medical field gathers an abundance of quantitative data — patient information; insurance claims; symptom, procedure, and pharmaceutical data; vital statistics; treatment outcomes; even progressive disease factors.

Most of this information is collected and referenced in the short-term by caregivers and insurance companies to help identify and treat diseases with the best possible outcomes. Healthcare data may assist a physician in identifying borderline breast cancer or liver disease cases. It may allow a hospital to reduce its readmission rates for patients who undergo heart surgery. Or, it may help an insurance company increase the number of claims that are automatically adjudicated, thus reducing its processing costs.

The healthcare industry also recognizes the long-term value of such data to help improve patient care, optimize processes, lower costs, recognize trends and predict patient outcomes. As with many industries, though, collecting the data is the easy part. Analyzing the data — thoroughly, completely, and accurately — can be a daunting challenge for many reasons. In most cases, the primary vehicles used for mining healthcare data include traditional statistical tools and techniques such as reductions, modeling and simulations. While these methods often produce results with a high degree of accuracy, they tend to be time-consuming, often taking weeks or even months to produce results. These methods also require data analysts with both statistical and medical expertise.

Other shortcomings that have plagued traditional methods of data mining include missing data and increasingly complex data sets. These issues can prevent healthcare providers from realizing the true value of the data they are diligently collecting. For example, missing data points in a data set render linear statistical approaches ineffective, as these tools cannot effectively compensate to fill in the blanks. Data sets that contain too many records, or more variables than the statistical tool can accommodate, may require omission of records or variables to reduce the size or dimensional complexity of the data set. The natural bias that comes from determining *which* variables should be omitted makes the analysis more subjective. Although the choice may be based on traditional assumptions about which factors are relevant to the outcome, the fact that assumptions are made at all narrows the field of possible results.

## Using Artificial Intelligence to Find Hidden Relationships in Data

Traditional data mining techniques, while still viable and necessary for many business intelligence applications, fall short when it comes to the enormous, complex data sets that modern technology allows the medical industry to compile. What is needed is a statistical tool that can analyze large, multi-dimensional data sets — quickly, easily and accurately.

Fortunately such tools do exist, in the form of powerful artificial intelligence (AI) applications, such as RapAnalyst™. Developed by Raptor International, RapAnalyst goes beyond data mining into the realm of analytics. By using advanced AI techniques, RapAnalyst digs deeper to find hidden relationships between variables in any data set.

RapAnalyst is based on neural network technology. Neural nets simulate the operation of the human brain by connecting hundreds or thousands of simulated neurons grouped together in much the same way as the brain's neurons connect. Neural nets are particularly effective for predicting future events when the networks have a large database of prior examples to draw on.

A self-organizing map (SOM), also employed by RapAnalyst, is a unique type of neural network that constructs a topology map from many variables. Based on the perception systems used by mammals, which cluster similarly functioning neurons together, a SOM groups similar data points together and visually displays the data so that relationships between variables can be seen easily and clearly.

## The Power of Artificial Intelligence to Improve Healthcare

How can AI technology benefit the medical community? Imagine conducting a study such as the breast cancer research done by the University of Wisconsin Hospital (Wolberg, 1989-1991). Collecting the information would still take months or years. However, using a neural network application such as RapAnalyst, the research team could reduce the time spent *analyzing* the data to hours, instead of days or weeks. The team could analyze the entire data set, rather than omitting records where there are missing values. No assumptions would need to be made about which attributes were relevant, as the tool would be able to accommodate numerous data dimensions. And, the predictive accuracy would be equal to or higher than could be achieved using traditional methods — high enough for reliable decisions to be based on.

In another case, RapAnalyst was put to the test with a data set from a large pathology laboratory. The data initially supplied consisted of 1320 randomly extracted patient records containing 24 attributes (including 11 blood readings, 10 liver conditions, age and gender). Raptor’s analysis team read the entire data set into RapAnalyst. When asked to predict whether any given patient had normal or abnormal liver functionality, the software quickly provided the results and did so with an average of 99.3 percent prediction accuracy.

RapAnalyst is able to make predictions by “training” the data. The program uses a holistic unsupervised neural network to group the data points based on similarity. It learns more about the data as it goes, regrouping data points as needed. All attributes are included in the training process with no weight being attached to specific attributes (e.g., the software makes no assumptions about which attributes are more relevant).

Once the training process is complete, RapAnalyst presents the user with a series of “Knowledge Filters” – visual representations of the relationships between the attributes. The Knowledge Filters contain “nodes” (hexagonal markers representing one or more data points) which are grouped based on proximity and similarity. Nodes close together are as similar as they can be considering all the attributes in the data set; nodes that are farther apart represent data points that are more dissimilar. Natural boundaries form as similar nodes group together into clusters.

By comparing each attribute’s Knowledge Filter, observers can quickly identify the relationships between the attributes and see them visually on the computer screen. These relationships can be seen even by those who have no experience relevant to the data set. In the liver disease example, the resulting node clusters suggested that at least three attributes (figure 1) were closely correlated with abnormal liver conditions (figure 2).

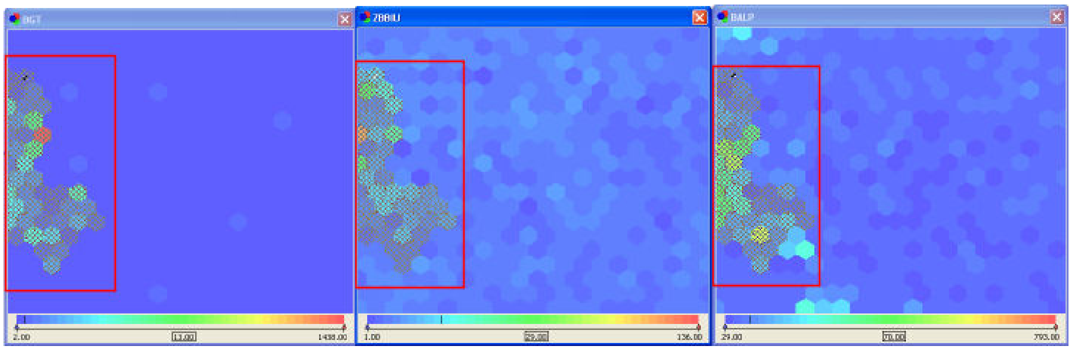

Figure 1 – Three attributes show similar clusters of nodes from the high end of the scale.

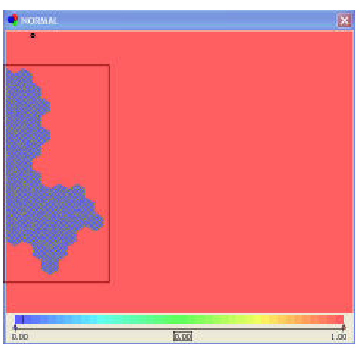

Figure 2 – The “Normal” attribute window displays abnormal liver conditions in blue.

## Using RapAnalyst to Improve and Predict Future Care Needs

In addition to identifying contributing disease factors, patient care data is often mined to learn which factors may help improve care, shorten rehabilitation time, save money or reduce process inefficiencies. Case in point, a major east coast hospital that desired to reduce its readmission rates for patients who had undergone CABG (coronary artery bypass grafting) surgery.

RapAnalyst was given a data set containing 1292 records, representing all CABG surgeries at the facility over a two year period. Over 20 attributes were read into the software. During the investigation of the resulting Knowledge Filters, several attribute windows immediately presented themselves with distinct areas of similarly colored nodes. When the attributes were further analyzed using RapAnalyst's Quick Predict tool, the data showed that patients who had spent a minimum of eight hours in ICU and at least 24 hours on a ventilator exhibited almost a zero percent readmission rate (figure 3).

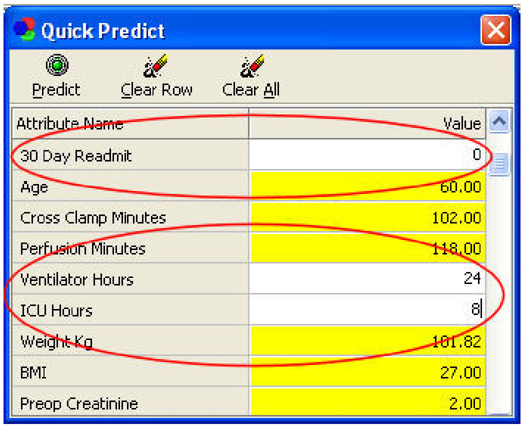


Figure 3 – RapAnalyst's Quick Predict tool can help predict future care needs.

These correlations suggest that CABG surgery patients significantly benefit from an infusion of oxygen after surgery, as well as from the acute care provided by ICU attendants. Such conclusions may be newsworthy to some, common sense to others. However, without the data analysis to support the assertion, a hospital may not deem it necessary to provide such intense care for these patients, let alone a special facility devoted to post-surgery care in addition to ICU. Such rationale seems more feasible when one looks at the data, and then compares a $10,000 price tag per patient for ICU and ventilation services, versus $30,000 (and up) per readmission.

## Putting RapAnalyst to Work in Your Healthcare Environment

In addition to the relatively small data sets mentioned here, RapAnalyst has proven its ability to work with data sets as large as 50,000 rows and 125 variables, or even larger data sets depending on the processing power of the desktop computer it is running on. With this sort of bandwidth, a powerful, easy-to-use tool such as RapAnalyst is a valuable addition to anyone's statistical toolkit. It proves especially beneficial to those in the healthcare industry due to its accuracy, its speed and its ability to allow users to question traditional assumptions in search of more in-depth knowledge.

To find out more about how RapAnalyst can help your organization improve patient care, or to learn more about Raptor International's expert consultations and custom software, visit us online at (website) or contact us at ###.