



HEALTH CARE FRAUD REPORT



VOL. 15, NO. 4

FEBRUARY 23, 2011

The Changing Face of Health Care Fraud Detection—Predictive Analytics



BY JULIE MALIDA

Historically, health care payers in both the public and private sectors have relied upon business rules and outlier models (sometimes also called anomaly detection methods) to mine claims data for fraud, waste and abuse patterns.

For example, a pathologist that is submitting bills for routine office visits might be a rule violation. After all, pathologists analyze tissue samples; they don't have routine office visits. But business rules are easily "gamed" once the fraudster knows what violations the payer is watching for in the data and deftly determines how to evade the rule that is in place.

Similarly, if an ear, nose, and throat (ENT) physician is performing 10 times the number of chest X-rays as other ENTs in the same geography, that can be considered an "outlier." But these outlier and rules-based methods have many false positives.

Malida is the principal for Health Care Fraud at SAS, in the Fraud and Financial Crimes global practice. She has devoted more than 27 years to the health care industry, focusing on managed care and cost containment in medical claims. Malida also is a Fellow of the Society of Actuaries and a member of the American Academy of Actuaries. She can be reached at (312) 240-0083, extension 58809, or Julie.malida@sas.com.

In the ENT example, preliminary scrutiny might reveal this ENT is located in a rural area and is the only office for 100 miles with X-ray equipment, so all the X-ray referrals come to that office. The investigator has wasted precious time only to move on to the next case. This is known as a "false positive."

Since the size (by dollars) of the problem of health care fraud is 100 times bigger than the fraud problem in the financial services sector, health care payers are realizing they could never hire enough experienced investigators to chase all the fraud. Hence, the prioritization of leads for fraud detection is of utmost importance, so those precious investigators are spending time on cases likeliest to be truly fraudulent.

The best arguments for using powerful analytical, user-friendly approaches come from payer successes:

- An insurer saved \$11 million in just its first year, and investigator productivity climbed 30 percent as activities that once took hours now take minutes.
- One state prevented \$14 million in Medicaid fraud and detected an additional \$27 million in fraudulent claims, leading to indictments.
- A national insurer's fraud investigators now quickly stop payment on complex fraud cases. The solution frees up experienced investigators to work on highly complex cases.

Predictive Modeling as an Added Fiduciary Guardian

Predictive modeling provides more accurate fraud scheme discovery than rules and anomaly detection methods. By using data mining capabilities to detect statistical patterns in previous, known fraudulent behavior, models can assign a propensity score on new claims with similar characteristics. This enables investigators to prioritize those new claims with the highest probability of fraud.

Predictive modeling is also superior to rules-only approaches because the investigator need not define the specific criteria that can later be evaded. Instead, the data will drive the identification and weighting of parameters that contributed to the fraudulent pattern. In this way, the false positives can be reduced, sometimes dramatically, allowing the investigator to scrutinize the most important cases first.

Examples of accurate health care fraud scoring supported by predictive models include:

- Alerting investigators to a potentially fictitious or “phantom” clinic. Using predictive models, a payer can analyze the data from prior phantom clinics that were opened for a short period of time, experienced high use, and then shut down. New claims exhibiting similar patterns can be identified before shutdown occurs.
- Uncovering stolen benefit identification numbers by identifying similar methods to those used previously to manipulate numbers in identity theft cases.
- A physician treating the same family members multiple times per week for the same chronic injury or illness, including the children. A predictive model can be built to “laser in” on which cases like this have had the highest statistical propensity for fraud in the past, so that every case like this is not unnecessarily chased by an investigator (e.g. acute influenza might make sense, but chronic low back pain may not).

Using varied predictive modeling techniques provides the best accuracy (lift in accuracy of scoring) and most appropriate method for the situation in question (depending on missing data, large number of variables, etc.). Incorporating varied methods, and the capability to compare and determine the most accurate analysis, provides the best mechanism for reducing false positives.

Common methods include regression, decision trees, and neural networks, each of which has subcategories underneath. There is not one method that is superior in performance to others in all cases. Testing various methods also provides the ability to frequently reevaluate the methods to prevent the analysis from degrading over time. All models degrade over time as the likelihood of the fraud scheme changing increases.

Investigations and resolution of improper payments provide a valuable source of training data to improve the accuracy of predictive models over time. Once that feedback loop is made available, the models become smarter. An audit trail is created as the model is phased into production and the preceding model is retired. It is important to monitor the performance of predictive models on an ongoing basis and not just put them on autopilot.

As the feedback from investigations is incorporated over time, payers can become less reliant on an investigator’s “expert judgment” on the weight to place on various rules and outlier models. They become more confident in the predictive models’ interpretation of the underlying data.

What starts as 90 percent reliance on a human’s expert judgment, and 10 percent data driven, can evolve to 90 percent reliance on the predictive power of the data and 10 percent on expert judgment.

That is not to imply there won’t always be a role for expert judgment in the investigation process, but rather that the focus of the role will change from identification of initial leads to investigation of those better leads.

Enabling Events and the Government Landscape

Provisions in the Affordable Care Act (ACA) are aimed at curtailing health care fraud and abuse in government programs, including support for the use of pre-

dictive analytics. In September 2010, President Obama signed into law the Small Business Lending Act, which includes implementation of predictive modeling techniques to combat fraud and abuse in government health insurance programs.

The legislation created one of the largest single investments in health care fraud management solutions. The Centers for Medicare & Medicaid Services (CMS) is a pioneer in this effort. CMS plans to implement predictive modeling for health care improper payments in 10 states by July 2011. After the first year, a progress report to Congress will determine expansion to other states.

This focus has led to the federal government leading state and local governments in adopting predictive analytics to combat fraud. The TechWeb research paper *The State of Fraud in Government*, published in October 2010, summarizes a survey of 327 federal, state, and local government decision makers and offers valuable insight.

- 27 percent of those from federal government organizations describe themselves as proactive in using predictive analytics, compared to just 12 percent of those from state and local government.
- More than half (56 percent) of participants say that predicting the likelihood of fraudulent events or behavior is a benefit they seek from their investment in data analysis. However, only 23 percent currently employ data mining and predictive analytics and just 28 percent analyze, model, and score potential risks and threats.

The survey found that only 10 percent of organizations are very satisfied with their current software and services for enabling predictive analytics, and 33 percent are somewhat satisfied.

Interestingly, only a minority of participants say their organization is using tools from recognized data mining and predictive analytics tool providers; the majority are using systems from the market-dominant BI, spreadsheet and database software providers.

This suggests that organizations may currently lack the appropriate tools to accomplish more advanced data mining and predictive analytics objectives.

Specific to health care, the survey offered participants a choice of 15 frequently encountered fraud, abuse, or improper payment instances that involve benefit, billing, health care product, insurance and Medicaid fraud.

The top three purposes for which the largest percentages are implementing data analysis and predictive analytics are:

- To uncover cases of billing for services not provided, not necessary, or at too high a rate (34 percent).
- Hidden conflicts of interest, kickback, or other relationships (33 percent).
- Claims irregularities, such as duplicates or too many claims to one provider (32 percent).

While many state and local governments are still figuring out a role for predictive analytics, there are some states and municipalities at the forefront, employing predictive analytics to model and score fraud risks, uncover patterns, and anticipate damaging incidents or behavior.

- As part of its efforts to combat Medicaid eligibility fraud, North Carolina will use predictive modeling

to create prioritized lists of potential fraud scores. The models generate a fraud risk scorecard from historic data. Risk factor scores are totaled to determine the overall fraud risk of an applicant or recipient.

- The Illinois Inspector General's Department of Healthcare and Family Services transformed its Medicaid program using SAS®, including predictive analytics, to identify overpayments and prevent improper payments to health care providers.

The anticipated gains in preventing fraud, waste, and abuse are likely to have far-reaching impacts in the way fraud is perpetrated and curtailed.

As government-funded cash flow to suspicious providers is curtailed, the threat of fraud activity from those same providers to private insurers, or self-funded programs, may increase. This will occur at a time when the formerly uninsured are now entering the market, further exacerbating the situation.

The time is now for all potential payer targets to step up analytics efforts in health care fraud detection.

Closing Thoughts

There is no one detection technique that is capable of systematically identifying all health care fraud, espe-

cially in the fastest-growing areas involving collusion, identity theft, and organized crime. It is only through multiple approaches that a highly organized scheme can be detected without the need for an inside informant.

A hybrid approach, which integrates rules, anomaly detection, powerful predictive analysis capabilities, and modeling of relationships between entities using social network analysis, is key to recognizing fraud and stopping it while it occurs.

The United States attorney for the Southern District of New York recently compared past efforts to prevent health care fraud to preventing auto theft while leaving the keys in the car with the doors unlocked and the engine running.

We can either wait for the next fraud ring to grow so large that eventually they are caught or become more proactive in detecting them and stopping them. Only through multiple data-driven and advanced analytical techniques, working in concert, will payers have the opportunity to achieve this goal.

Reproduced with permission from BNA's Health Care Fraud Report, Vol. 14, No. 4, Feb. 23, 2011, The Bureau of National Affairs, Inc. (800-372-1033) <http://www.bna.com>.