Applying Business Intelligence Concepts to Medicaid Claim Fraud Detection

Leanndra Copeland
l-copeland@nvdetr.org
Nevada Department of Employment,
Training and Rehabilitation
Carson City, NV 89713, USA

Dana Edberg
dte@unr.edu
Department of Information Systems
University of Nevada, Reno
Reno, NV 89557, USA

Jeanne Wendel
wendel@unr.edu
Department of Economics
University of Nevada, Reno
Reno, NV 89557, USA

Abstract

U.S. governmental agencies are striving to do more with less. Controlling the costs of delivering healthcare services such as Medicaid is especially critical at a time of increasing program enrollment and decreasing state budgets. Fraud is estimated to steal up to ten percent of the taxpayer dollars used to fund governmentally supported healthcare, making it critical for government authorities to find cost effective methods to detect fraudulent transactions. This paper explores the use of a business intelligence system relying on statistical methods to detect fraud in one state's existing Medicaid claim payment data. This study shows that Medicaid claim transactions that have been collected for payment purposes can be reformatted and analyzed to detect fraud and provide input for decision makers charged with making the best use of available funding. The results illustrate the efficacy of using unsupervised statistical methods to detect fraud in healthcare-related data.

Keywords: business intelligence, government information systems, healthcare information technology, fraud, statistical analysis, unsupervised methods

1. INTRODUCTION

Publicly funded agencies in the U.S. face taxpayer demand to extend and improve services while also progressively lowering costs. These pressures are especially significant within the area of governmentally supported healthcare (Ryan, 2011).

The Centers for Medicare and Medicaid Services (CMS) estimate total U.S. health care spending in 2009 reached $2.5 trillion or 17.6% of gross domestic product (CMS, 2010). While it is difficult to pinpoint the exact amount of fraud in healthcare transactions, federal agencies estimate that from 3% to 10% of expenses are fraudulent, with 10% being the most accepted figure (Heaphy, 2011). Even using the most conservative estimate, it means that over $75 billion per year targeted for providing healthcare is stolen from taxpayer funds through fraudulent activities. To help combat this problem, the federal government has established Medicaid Fraud Control Units and State Program Integrity Units providing assistance and oversight for healthcare payment processes (CMS, 2011), but there is currently no evidence that these mechanisms have yielded significant improvement in fraud detection and protection.

Commercial enterprises frequently use business intelligence (BI) systems to help identify and control fraud (Han & Kamber, 2006; Kotsiantis, Koumanakos, Tzelepis, & Tampakas, 2006; Wegener & Rüping, 2010). BI systems consist of methods of gathering data, data storage (typically termed a "data warehouse"), and analytical tools such as visualization programs, statistical methods and data mining algorithms (Negash, 2004). The information gleaned from BI is used to provide decision makers with accurate, timely, well-formatted information. BI systems are a platform for organizing and analyzing data from disparate sources to provide meaningful information for decision makers (Davenport & Harris, 2007; Negash, 2004).

Health insurance organizations have applied this same technology successfully to monitor fraudulent activities (Sokol, Garcia, Rodriguez, West, & Johnson, 2001; Wang & Yang, 2009; Yang & Hwang, 2006). While authors emphasize the importance of using BI to support governmental decision making (Davenport & Jarvenpaa, 2008), actual implementation of these systems is problematic. Some government agencies experience challenges implementing BI due to the short-term funding cycles required by governing bodies, a lack of personnel with knowledge of BI, restricted funding, concerns about privacy, incomplete data, and poor integration of available data (Harper, 2004; Rosacker & Olson, 2008; Vann, 2004; Wilkin & Riddett, 2009).

A recent report from the U.S. Government Accountability Office (GAO, 2011) recommended that states make more effective use of information technology to help detect healthcare-related fraud, but also highlighted the difficulties experienced in creating and accessing a nationwide data repository for this function. The report emphasized the challenges in creating a fully integrated dataset to support inquiries from state agencies (GAO, 2011).

This paper addresses the use of BI technology for detecting fraud in a state's Medicaid payment system. We discuss how a state could use its existing data to create a data warehouse, and then illustrate how statistical methods could be applied to detect fraudulent Medicaid claims. Given the rising costs of healthcare, and the shrinking operating budgets of state government, it is of critical importance that fraud control units incorporate BI technology for effective fraud detection.

The next section of this paper presents existing research on methods of detecting fraud, highlighting studies aimed at identifying fraud in healthcare. The third section describes our study performed in one state using Medicaid claim data, and the fourth section briefly discusses the practical considerations of implementing BI to help detect Medicaid fraud.

2. METHODS OF DETECTING FRAUD

There are two main strategies for detecting fraud: auditing and statistics. Auditing strategies require the use of trained personnel to evaluate the process and/or product, while statistical methods rely on large data sets to identify potential anomalies. A summary of the strategies for detecting fraud is provided in Table 3 of the Appendix and each is described in the following sub-sections.

Auditing Strategies

When auditing healthcare systems, medical and claims experts are hired to review transactional claims on a case-by-case basis to identify anomalies based on the knowledge of those reviewing the claims (Yang & Hwang, 2006). Auditing strategies frequently use random stratification sampling methods to obtain samples from the spectrum of different claim types (Buddhakulsomsiri & Parthanadee, 2008), but auditing strategies cannot pinpoint

suspicious claims from millions of claims in a data set.

In a study of healthcare-related fraud, claims for durable medical equipment (DME) from two multi-county areas within a region served by an insurance carrier were analyzed (Wickizer, 1995). Part of the study utilized an audit strategy where nurse analysts reviewed DME claims to verify accuracy. Four different types of DME were examined in the study. As testament to the time factor in using audits, this study used a sample size of just 231 observations. The researcher examined twenty-one months (January 1990 – September 1991) of claims data in which four variables provided measurement for DME utilization: (1) order requests per month, (2) submitted charges per month, (3) Medicare-allowed payments per month, and (4) percentage of DME requests denied per month. Data on other covariates were gathered to control for external factors that could influence DME utilization, such as the number of hospital discharges per 1,000 Medicare beneficiaries. The findings showed that DME utilization management programs reduced the number of requests, submitted charges, and Medicare payments in three out of the four targeted DME items.

While auditing strategies tend to be accurate in finding fraud, they are costly and time-consuming to perform on the large number of transactions processed in the healthcare industry (Yang & Hwang, 2006). Thus, these strategies may not be feasible for detecting fraud in government organizations that are trying to make the best use of limited resources.

Statistical Strategies

Statistical fraud detection strategies rely on analytical methods such as correlation and regression to evaluate large data sets (Bolton & Hand, 2002). Some studies have pointed out that finding the source of fraud (insured, provider, etc) using statistical methods is far more efficient than analyzing individual claims (Ortega, Figueroa, & Ruz, 2006; Yang & Hwang, 2006).

Statistical strategies are classified as using either supervised or unsupervised methods (Bolton & Hand, 2002). Supervised methods require samples from both known fraudulent and non-fraudulent records in order to model the distinct characteristics of each. The data is labeled by human experts prior to processing through sophisticated computer data mining algorithms. Unsupervised methods, on the other hand, do not require any prior knowledge of the relative legitimacy of the data and the data is unlabeled. These two methods could be considered endpoints on a continuum of statistical strategies, with hybrid or semi-supervised methods sitting in the middle (Laleh & Abdollahi Azgomi, 2009). Semi-supervised methods use some data that is labeled, but also has unlabeled data that will be evaluated during program processing. The labeled data must be identified prior to input and requires pre-knowledge of fraudulent transactions for modeling purposes. To simplify the discussion of statistical methods, the next two sub-sections discuss studies concerning the two endpoints on the continuum.

Supervised Statistical Methods

A key issue in the use of supervised statistical methods is the need to identify fraudulent claims prior to using the data for further processing. An example of this constraint is a study using a multi-layer perceptron network to classify general practitioner (GP) physician profiles into categories ranging from normal to abnormal (He, Wang, Graco, & Hawkins, 1997). The study required an auditing portion to develop the supervised methods. Physicians, hired as expert consultants, identified 28 features which summarized a GP's practice over a year. The classified sample was used to train an automated classification system. The sample consisted of 1,500 randomly selected GP profiles from Australian physicians who participated in Medicare. The physicians serving as consultants classified all 1,500 profiles based on 28 distinct features before the sample was divided into two groups with 750 profiles for the training set and 750 profiles for the test set. The researchers concluded that a two-class neural network classification system was a viable method for detecting fraud. A problem with the method employed in the study is that it is not easily replicable in a governmental organization because of the expense involved with hiring medical experts to review such a large number of claims and create valid feature variables. In addition, the necessary software algorithms require personnel skilled in data mining operations.

Another study required meetings with medical experts to assist in developing a set of variables

used to discriminate between fraudulent and honest claims (Ortega, et al., 2006). This study examined 125 distinct features to four different areas/parties where fraud can occur: medical claims, the insured, the medical professional, and the employer. In addition, this study incorporated feedback results from each model input in the other three sub-models. As is common in studies detecting fraud, a full discussion of the results was prohibited due to a disclosure agreement between the authors and the insurance company that provided the data. However, it was proclaimed that the model accelerated detection on average 6.6 months earlier than standard audit strategies.

Another issue in the use of supervised statistical methods is the relative balance between legitimate and fraudulent transactions. Legitimate transactions far outweigh fraudulent ones in any practical dataset. Creating models from these unbalanced classes can cause misspecification (Bolton & Hand, 2002). Finally, the most critical issue is that supervised models cannot detect new types of fraud because the models are created from past fraud strategies (Bolton & Hand, 2002; Laleh & Abdollahi Azgomi, 2009). Despite these issues, supervised methods are widely used for fraud detection in healthcare and are supported by technologies such as neural networks , decision trees, fuzzy logic, and Bayesian networks (Li, Huang, Jin, & Shi, 2008).

Unsupervised Statistical Methods

Unsupervised statistical methods determine and tag outliers in a data set so that those outliers can then be marked for potential investigation. Unsupervised methods first use technology to identify potentially fraudulent transactions, and then afterwards require the use of expertise to determine the legitimacy of those transactions. The assumption is that fewer transactions will have to be investigated than supervised methods, and the investigation will be able to be performed by less costly personnel (Laleh & Abdollahi Azgomi, 2009). Unsupervised methods use clustering as a popular tool for detecting anomalous data (Bolton & Hand, 2002).

Using an unsupervised method, a study reviewed the medical insurance claims of 22,000 providers to test an electronic fraud detection (EFD) program (Major & Riedinger, 2002). The technique compared individual provider characteristics to their peers. The researchers recommended that provider comparison be grouped according to similar characteristics, such as the same organizational structure, specialty, and geographic location. The EFD developers examined 27 behavioral heuristics in five categories: financial (the flow of dollars), medical logic (whether a medical situation would normally happen), abuse (frequency of treatments), logistics (place, time and sequence of activities) and identification (how providers present themselves to the insurer). Validation of the model was yet to be determined at the time of publication, but the model did alert officials to over 800 suspicious providers and resulted in the launch of 23 investigations (Major & Riedinger, 2002).

Another approach to unsupervised fraud detection is a numerical phenomenon called Benford's Law. Benford's Law, sometimes referred to as the first-digit law, states that the first significant digit of some data follows a non-random pattern (Nigrini & Mittermaier, 1997). Benford's Law has been applied to fraud detection (for tax data) through the development of the Distortion Factor (DF) Model (Nigrini, 1992; Watrin, Struffert, & Ullmann, 2008). The DF model compares the first digit frequencies of observations in a data set to the expected frequencies of Benford's Law. The first digit distribution of many data sets follows Benford's Law, such as the one-day returns on the Dow Jones Industrial Average and the Standards and Poor's Index, street address, and many others data sets.

Unsupervised statistical methods may be more cost effective for government agencies than auditing or supervised statistical methods. Unsupervised methods use standard statistical processing, so it may be easier for government agencies to find appropriate personnel as compared to the data mining knowledge required for supervised methods. In addition, the initial detection of potential fraud is performed by technology, allowing for greater focus of expert time on those transactions that have a greater probability of fraud. An advantage of unsupervised methods is that they can detect new types of fraud. On the other hand, unsupervised methods require expertise in the initial development of potential factors that should be modeled for outliers as well as the creation of an appropriate dataset for processing. In addition, the efficacy of

unsupervised statistical methods is relatively untested in the literature (Bolton & Hand, 2002).

Purpose of this Study

This study contributes to the literature by exploring the use of unsupervised statistical methods for detecting healthcare fraud. The use of electronic healthcare claims lends itself to evaluation through BI tools such as statistically-based analytical methods. The goal of this study is to use statistical methods to improve the accuracy of detecting fraud, while minimizing the overall cost of system implementation for a government agency. This study was motivated by a practical need to create a simple and cost effective method of analyzing existing claims data to identify potential fraud. We demonstrate how that data might be reformatted and used to provide additional information for decision makers who are attempting to detect and control fraud.

The next section of this paper explores the use of unsupervised statistical methods to detect fraud in Medicaid claims for the state of Nevada.

3. APPLYING UNSUPERVISED STATISTICAL METHODS TO MEDICAID CLAIMS

Nevada is the seventh-largest state geographically, but with a relatively small population of 2.7 million people. About 10% of the Nevada population was enrolled in Medicaid in 2009 (as compared to an average 19% enrollment rate nationwide) and total Medicaid expenses for 2009 in Nevada were approximately $1.3 billion (Kaiser, 2009).

A recent spree in durable medical equipment, prosthetic, orthotic devices, and/or disposable medical supplies (DMEPOS) fraud in the state of Nevada prompted state authorities to explore whether BI might help the state become more effective at detecting fraud. DMEPOS is defined as equipment that is appropriate for in-home use and benefits the patient medically. DMEPOS may consist of items that can be used a repeated number of times or may be disposable supplies which are not reusable (NVHHS, 2009).

State authorities identified a particular DMEPOS item, disposable diapers, as being most appropriate for initial exploration. Diaper fraud is attractive to fraudsters because it is a high-volume item requiring relatively little medical expertise to process. Over the five year time frame of data used for this study, Nevada reimbursed 321 supplier companies for briefs, diapers and pads.

Data Used for Evaluation

The Nevada Department of Health and Human Services provided de-identified Medicaid claims data that linked provider, facility, and prescription claim transaction records over a five-year time period from January 2005 to December 2009. During this time, Medicaid reimbursed 693 DMEPOS supply companies for a total of $87,340,766. Data came from three different payer organizations, and was presented to us using three different formats. The data was delivered in comma delimited ASCII files.

Database Design, Extract, Transform and Load

Some researchers estimate that data preparation consumes 80% of the time in a fraud detection project (Li, et al., 2008; Lin & Haug, 2006; Sokol, et al., 2001). Database structures of raw claims data and electronic health records are designed to support financial transactions and health care delivery, rather than fraud detection or query development, and thus must be reshaped to support data analysis operations. We created a data warehouse from the data sources that could be used to support multiple inquiries, so data preparation for this project was time-consuming. We estimate that data preparation took about 85% of overall project time. However, once the data was loaded in a data warehouse, it could be accessed in a variety of ways for different analytical applications so we anticipate that future data preparation time will be significantly less than the original development.

A normalized database design was created to store de-identified data about patients, providers and claims. The data warehouse used for this study was used for additional studies, so it was critical to create an adaptable and flexible design. Claims were subdivided into provider, facility and pharmacy categories to facilitate faster data access. The data warehouse was implemented using Microsoft SQL Server 2008R2. Data from the three different payer organizations was extracted, transformed and loaded (ETL) using both SQL Server Integration Services and customized load routines. Since data formats differed among the three input sources, data had to be made consistent during the ETL process. The database contained a total

of approximately 46.7 million claim records for the five year period. The data of interest for this study was the characteristics of DME suppliers enrolled to provide services to Medicaid patients and their claims. This subset consisted of about 10 million claim records.

Data Analysis

After the data was loaded in the data warehouse, data analysis proceeded iteratively to identify appropriate features and evaluate the data.

The claim records used included detailed information such as company name, diagnosis codes, procedure (DMEPOS) codes, de-identified patient number, total charges claimed, etc. Every provider with a disproportionately high or low outcome for a given variable was assigned weighted points based upon total number of patients, total amount claimed, or length of company operation to ensure a variable did not disproportionately represent any provider with certain characteristics. A variable that resulted in high quality data was weighted more than variables with lower discriminatory power. Furthermore, the size of the company could have affected the outcome of a variable and was taken into consideration before assigning points. The assigning of weights will be addressed further in the results section.

Features that might help detect fraudulent activities were derived from hints in the literature and influenced significantly through discussions with Nevada state authorities. A profile consisting of 12 features was ultimately created for each of the 321 DME suppliers providing incontinence briefs, diapers, or pads. A hindrance in fraud control efforts is the expurgation of public discourse about new fraud detection techniques to prevent alerting fraudsters. If criminals gain knowledge of how detection systems work, this could occlude the efficacy of new ideas before opportunity to detect fraud arises. Thus, academic literature rarely reveals the features used to isolate fraud (Bolton & Hand, 2002). The features created for this study were largely original and cannot be revealed due to an agreement with authorities from Nevada. Besides the censoring of enforcement techniques, provision of data sets and complete discussion of fraud study results are a rarity in academic literature (Bolton and Hand, 2002).

After the 12 features were solidified, analysis proceeded in three steps. First, the DMEPOS supplier's behavior was measured for each feature. Second, suppliers were compared against other suppliers. If a supplier fell into the outlier range as determined by the upper or lower fifth percentile of any of the features, the supplier was assigned weighted points given the strength of the variable and the size of the supplier's transactions, as mentioned previously. Because the upper or lower fifth percentile cutoffs are assigned based on statistics, not logic, thought should be given to whether the statistical cutoff divides the groups into questionable and likely benign categories. Thus, the third step assists in this task by providing visualization through tables and graphs. Visualization techniques help show whether or not the feature variable divides the suppliers into useful categories with noticeable tails. If not, the weight for the feature variable is lowered. The more points a supplier has after all twelve feature variables are analyzed, the more suspicious that supplier looks. The next subsection provides more detail about two of the twelve feature variables used in the study.

Results: Diapers per claim

Medicaid rules limit patients to 300 diapers per month. The more diapers supplied per claim, the more money a fraudulent company can make. If a supplier consistently orders 300 diapers per claim for multiple patients, this means that most of their patients need approximately ten diapers a day. This equates to changing a brief nearly every two and half hours around the clock. Initially a patient will require more briefs because they need to stock up around the house and other frequented locations. For example, parents with newborn children have diapers in the car, living room, bedroom, etc. Adults needing briefs would go through the same transition and would require more in the beginning.

The results showed that suppliers whose average was over 272.5 diapers per claim fell into the 95 percentile. Using this metric, 16 companies were flagged as shown in Table below.

Table 1. Suppliers Flagged for Feature/Variable 1

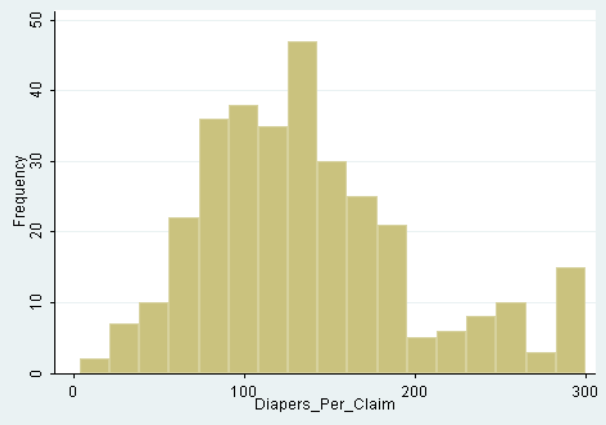
Supplier ID	Diapers Per Claim	Number of Unique Patients	Total Points
100503886	300	1	1 x 3 = 3
100505813	300	1	1 x 3 = 3
3302100	300	7	1 x 3 =3
100509994	300	45	3 x 3 = 9
3302448	300	152	5 x 3 = 15
100508690	300	138	5 x 3 =15
3302827	298	40	3 x 3 = 9
100510601	297	1	1 x 3 = 3
3302790	297	2	1 x 3 = 3
100500017	296	8	1 x 3 = 3
3302995	295	2	1 x 3 = 3
3302044	294	2	1 x 3 = 3
3302133	289	3	1 x 3 = 3
100501150	288	1	1 x 3 = 3
100509795	288	1	1 x 3 = 3
100503791	274	299	5 x 3 = 15

Before assigning points to the isolated companies, two things were considered: (1) some companies averaged a high number of diapers per claim, but only delivered to a few patients; and (2) the strongest variables separated suppliers into well-defined categories with a definite right tail.

Patients that need incontinence supplies have varying levels of bladder control; not all will need the maximum 300 briefs per month. To emphasize companies which consistently supplied a high average to numerous patients, a weighting system was implemented. The number of patients served determined the weight. Companies that supplied more than 50 patients were assigned five points. Companies with 20-49 patients were assigned three points, and companies with less than 20 patients were assigned one point.

Figure 1 illustrates the effectiveness of the feature variable at categorizing the suppliers. Figure 1 also demonstrates the power of visualization methods in BI, allowing a person to quickly see the anomalous suppliers. The histogram reveals that supplier behavior is roughly normally distributed barring the abnormal activity in the right tail. Because there is a distinct distribution, the variable successfully identifies a marked right tail; therefore, more emphasis should be placed on suppliers isolated with this variable. Variables that divide suppliers into many categories were weighted times 3 points. Revisit Table 1 to see the suppliers isolated by the diapers per claim feature variable.

Figure 1. Visualization Histogram for Diapers per Claim



The "diapers per claims" variable illustrates how an effective feature variable categorizes providers into a distribution with definite tails. Not all variables did such an effective job. The next section presents another feature that was less effective.

Results: Pre-authorization Requests

The number of pre-authorizations requested is an example of a variable with limited discriminative power.

If a supplier had no pre-authorization requests yet served many patients for an extended period, it may be considered suspicious. Fictitious organizations may prefer to limit their exposure to the system, whereas legitimate companies will likely require pre-authorizations at some point. Of the 321 DME companies that supplied briefs, 19 obtained pre-authorization for specialized orders.

Because only 6% of the suppliers needed a pre-authorization within the five year time frame, this feature variable did a poor job at categorizing the supplier companies. It essentially breaks the suppliers into two

categories with the vast majority never utilizing the preauthorization system. Any supplier that had a least one pre-authorization fell into the upper 95th percentile and was considered benign for this feature. This variable was not given much weight due to its inability to categorize the DMEPOS supply companies into many categories where a distinct tail can be seen; therefore, the variable is assigned a weight of one.

Next, the number of claims a supplier submitted is considered to further distribute points appropriately. Suppliers that submitted many claims were given more points. Suppliers that submitted less than 500 claims earned one point. Suppliers that submitted between 500 and 999 claims earned two points. Suppliers that submitted over 1,000 claims earned three points.

Table 2 on the top of the next page details the results. "Claims" was chosen as the weight in acknowledgement of the need to stack up on supplies after initial diagnosis.

Table 2. Supplier Points Based on Preauthorization

Number of companies	Range of Count of Claims	Range of Count of Patients	Total points received per company
288	1-477	1-82	1
7	548-826	65-152	2
6	1151-4145	79-414	3

Overall Results

Table 4 provided in the Appendix shows the suppliers that were flagged the most by the unsupervised statistical methods used for this study. The total amount of money spent on suspicious claims detected by this method totals to $449,100, or 5.9 percent of the total amount spent on incontinence briefs during this five year period.

After presenting the results to the state fraud surveillance unit, it was determined that three of the six suppliers flagged were fraudulent. Therefore, this method was believed by state authorities to have demonstrated its effectiveness in isolating suspicious suppliers.

Limitations

This is an exploratory study to help Nevada state authorities determine the applicability and effectiveness of BI for Nevada's Medicaid fraud detection. The results may be applicable only for this single state and may not be generalizable to the nationwide Medicaid claim population. Nevada's Medicaid population is significantly smaller than the national average and tends to contain more transient participants (Kaiser, 2009). There is little data available about the suppliers for Medicaid in the U.S., so we were not able to evaluate the comparability of Medicaid suppliers in Nevada to the rest of the U.S.

The point system applied in this study was used to explore the potential for relative weights in unsupervised methods. The weight system would need further evaluation to determine its most appropriate use.

4. CONCLUSIONS

There are three considerations if a governmental agency wished to implement unsupervised statistical methods for fraud detection. First, due to the dynamic nature of fraudulent activity, the way in which criminals commit fraud will evolve, as must the way in which the state goes about detecting it. The model presented here should be refined over time by dropping or adding relevant feature variables to continue being effective.

Second, a concern raised by state authorities about this procedure was that it only indentified companies that were no longer active; however, this method can easily be applied to real-time data to catch criminals before they go out of business. Real-time monitoring of provider behavior is a critical component of any medical fraud detection tool. This paper illustrates an effective method that could be incorporated with real-time claims data to achieve real-time business intelligence. The method presented is an analytic-based fraud detection tool that scores companies and isolates atypical providers.

Third, implementation of this model required knowledge of both standard statistical analysis and BI-related technology, as well as some limited knowledge about the Medicaid application domain. Creation of the data warehouse required expertise in database design and ETL,

while data access and analysis required skills in SQL programming and statistics. In order to determine the most appropriate feature variables, it was necessary to understand existing literature in healthcare fraud and to gather information from state experts in Medicaid claims.

Fraud is a perpetually changing enterprise. Once the state detects a scheme, it should implement detection tools that use supervised methods to rapidly spot future schemes with similar characteristics. This detection and scout method used by surveillance teams pushes criminals to constantly find new ways to steal money. The unsupervised statistical method presented in this paper should be used to continue scanning the data for new anomalies.

In these difficult financial times of shrinking state budgets and rising health-care costs, states need to target claims with a high probability of fraud so they can concentrate on stemming financial losses coming out of the taxpayers' wallets. Without implementing BI, the state will inevitably spend too much time reviewing honest claims.

This practical application of BI provides the opportunity for a government agency to reduce manpower and improve operational efficiency concurrently. The BI based analytic method explored in this study combines statistical methods with a data warehouse to turn data that is already available from claims processing into a new and powerful tool for detecting fraud.

5. REFERENCES

Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical Science, 17*(3), 235-249.

Buddhakulsomsiri, J., & Parthanadee, P. (2008). Stratified random sampling for estimating billing accuracy in health care systems. *Health Care Management Science, 11*(1), 41-54.

CMS. (2010). *National Health Expenditures 2009 Highlights*. Washington, DC: U.S. Department of Health & Human Services Retrieved from http://www.cms.gov/NationalHealthExpendData/downloads/highlights.pdf.

CMS. (2011). Medicaid Guidance Fraud Prevention. *State Program Integrity Support & Assistance* Retrieved July 14, 2011, from http://www.cms.gov/FraudAbuseforProfs/02_MedicaidGuidance.asp

Davenport, T. H., & Harris, J. G. (2007). *Competing on Analytics: The New Science of Winning*: Harvard Business Press.

Davenport, T. H., & Jarvenpaa, S. L. (2008). *Strategic Use of Analytics in Government*: IBM Center for the Business of Government.

GAO. (2011). *Fraud Detection Systems: Centers for Medicare and Medicaid Services Needs to Ensure More Widespread Use*. (GAO-11-475). Washington, DC: Report to Congressional Requesters Retrieved from http://www.gao.gov/new.items/d11475.pdf.

Han, J., & Kamber, M. (2006). *Data mining: concepts and techniques*: Morgan Kaufmann.

Harper, F. M. (2004). Data Warehousing and the Organization of Governmental Databases. In A. Pavlichev & G. D. Garson (Eds.), *Digital Government: Principles and Best Practices*. Hershey, PA: Idea Group, Inc.

He, H., Wang, J., Graco, W., & Hawkins, S. (1997). Application of neural networks to detection of medical fraud. *Expert Systems with Applications, 13*(4), 329-336.

Heaphy, T. J. (2011). Health Care Fraud Retrieved July 14, 2011, from http://www.justice.gov/usao/vaw/health_care_fraud/

Kaiser. (2009). State Health Facts Nevada: Total Medicaid Spending, FY 2009. *Individual State Profiles* Retrieved July 14, 2011, from http://www.statehealthfacts.org/profileind.jsp?cmprgn=1&cat=4&rgn=30&ind=177&sub=47

Kotsiantis, S., Koumanakos, E., Tzelepis, D., & Tampakas, V. (2006). Forecasting fraudulent financial statements using data mining. *International Journal of Computational Intelligence, 3*(2), 104–110.

Laleh, N., & Abdollahi Azgomi, M. (2009). A Taxonomy of Frauds and Fraud Detection Techniques. In S. K. Prasad, S. Routray, R. Khurana & S. Sahni (Eds.), *Information Systems, Technology and Management* (Vol. 31, pp. 256-267): Springer Berlin Heidelberg.

Li, J., Huang, K. Y., Jin, J., & Shi, J. (2008). A survey on statistical methods for health care fraud detection. *Health Care Management Science, 11*(3), 275-287.

Lin, J. H., & Haug, P. J. (2006). *Data preparation framework for preprocessing clinical data in data mining.* Paper presented at the AMIS Annual Symposium Proceedings.

Major, J. A., & Riedinger, D. R. (2002). EFD: A Hybrid Knowledge/Statistical Based System for the Detection of Fraud. *Journal of Risk and Insurance, 69*(3), 309-324.

Negash, S. (2004). Business intelligence. *The Communications of the Association for Information Systems, 13*(1), 54.

Nigrini, M. J. (1992). *The detection of income tax evasion through an analysis of digital frequencies.* Dissertation, Cincinnati, OH: University of Cincinnati.

Nigrini, M. J., & Mittermaier, L. J. (1997). The use of Benford's Law as an aid in analytical procedures. *Auditing, 16*, 52-67.

NVHHS. (2009). *DME Information Sheet*. (NMO 1115E 11/09). Carson City, NV: Retrieved from http://dhcfp.nv.gov/pdf%20forms/FactSheets/1115E.pdf.

Ortega, P. A., Figueroa, C. J., & Ruz, G. A. (2006). *A medical claim fraud/abuse detection system based on data mining: a case study in Chile.* Paper presented at the International Conference on Data Mining, Las Vegas, NV.

Rosacker, K. M., & Olson, D. L. (2008). Public sector information system critical success factors. *Transforming Government: People, Process and Policy, 2*(1), 60-70.

Ryan, K. (2011). Weighing the Impact of Cuts: Social Security, Medicare and Medicaid. *Public News Service*. Retrieved from News in the Public Interest website: http://www.publicnewsservice.org/index.php?/content/article/21180-1

Sokol, L., Garcia, B., Rodriguez, J., West, M., & Johnson, K. (2001). Using data mining to find fraud in HCFA health care claims. *Topics in health information management, 22*(1), 1-13.

Vann, J. L. (2004). Resistance to Change and the Language of Public Organizations: A Look at "Clashing Grammars" in Large-Scale Information Technology Projects. *Public Organization Review, 4*(1), 47-73. doi: 10.1023/B:PORJ.0000015651.06417.e1

Wang, J., & Yang, J. G. S. (2009). Data Mining Techniques for Auditing Attest Function and Fraud Detection. *Journal of Forensic & Investigative Accounting, 1*(1).

Watrin, C., Struffert, R., & Ullmann, R. (2008). Benford's Law: an instrument for selecting tax audit targets? *Review of Managerial Science, 2*(3), 219-237.

Wegener, D., & Rüping, S. (2010). On Integrating Data Mining into Business Processes. In W. Abramowicz & R. Tolksdorf (Eds.), *Business Information Systems* (Vol. 47, pp. 183-194): Springer Berlin Heidelberg.

Wickizer, T. M. (1995). Controlling Outpatient Medical Equipment Costs Through Utilization Management. *Medical care, 33*(4), 383-391.

Wilkin, C., & Riddett, J. (2009). IT governance challenges in a large not-for-profit healthcare organization: The role of intranets. *Electronic Commerce Research, 9*(4), 351-374. doi: 10.1007/s10660-009-9038-0

Yang, W. S., & Hwang, S. Y. (2006). A process-mining framework for the detection of healthcare fraud and abuse. *Expert Systems with Applications, 31*(1), 56-68.

Appendix

Table 3. Summary of Fraud Detection Methods Relevant to Healthcare

Detection Method	Description	Benefits	Drawbacks	Key Findings from Prior Research
Auditing	Medical experts review individual claims one-by-one. Claims are usually selected by a random sample, but could be a targeted sample. Relies on human expertise.	Accuracy, Comprehensive	Costly, time consuming, requires experienced personnel, inefficient	• Found that the best sampling method depends on what is being measured (Buddhakulsomsiri & Parthanadee, 2008)
Statistical: Supervised	Medical and claims experts identify known fraudulent and known honest claims. These claims are modeled to forecast unknown claims. Uses BI data mining tools such as neural networks and fuzzy logic	Proven technology in business fraud. Quickly pinpoints suspicious providers. Widely used.	Cannot detect new types of fraud. May identify legitimate claims as fraudulent. Requires expertise prior to detection of fraud. Requires knowledge of complex BI tools.	• Created sub-models with feedback connections (Ortega, et al., 2006) • Determined that two categories were more productive than four. (He, et al., 1997)
Statistical: Unsupervised	Statistical algorithms are used to identify outliers based on pre-defined categories. Filters out anomalous behavior from peer groups. Anomalous data is examined by claims experts to detect. Uses BI statistical tools such as standard T-tests, correlation, clustering, and regression.	Quickly pinpoints suspicious providers. Can detect new types of fraud.	May identify legitimate claims as fraudulent. Requires examination of claims after statistical evaluation. Requires knowledge of statistical methods.	• Identified key categories for health care fraud (Major & Riedinger, 2002) • Recommended use of clustering in data mining (Bolton & Hand, 2002) • Found Benford's Law applicable for fraud detection (Nigrini, 1992); (Nigrini & Mittermaier, 1997)

Table 4. Top Counts of Flagged Suppliers

	Feature Variable													
Supplier ID	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Number of times flagged**	**Total Net Pay Amt**
3302448	1	1	0	1	0	0	1	1	0	1	1	1	8	$146,160
100508690	1	1	0	1	0	0	1	1	0	1	1	0	7	$215,586
100509994	1	1	0	1	0	0	1	1	0	1	1	0	7	$33,930
100510770	0	1	0	1	0	1	0	1	0	1	1	0	6	$783
3302827	1	0	0	1	0	0	0	1	0	1	1	0	5	$44,053.5
100500017	1	0	0	0	0	0	1	1	0	1	1	0	5	$8,587