University of Nevada, Reno

Detecting Fraudulent Suppliers of Incontinence Briefs

A thesis submitted in partial fulfillment of the
requirements for the degree of Master of
Science in Economics

by

Leanndra Copeland

Dr. Jeanne Wendel/Thesis Advisor

May, 2011

University of Nevada, Reno
Statewide • Worldwide

THE GRADUATE SCHOOL

We recommend that the thesis
prepared under our supervision by

LEANNDRA COPELAND

entitled

Detecting Fraudulent Suppliers Of Incontinence Briefs

be accepted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

Jeanne Wendel, Phd.., Advisor

Sankar Mukhopadhyay, Phd. , Committee Member

Dana Edberg, Phd., Committee Member

Anna Panorska, PhD., Graduate School Representative

Marsha H. Read, Ph. D., Associate Dean, Graduate School

May, 2011

Abstract: **An unsupervised fraud detection approach that flagged anomalous companies was applied to Medicaid claims data. Durable medical equipment suppliers of incontinence briefs were examined for suspicious behavior using claims data from January 2005 to December 2009. Twelve statistical variables were created and applied to individual companies. For every outlier range a company fell into, it was assigned a point. After examining all variables, the points were tallied. High scoring companies were referred to an investigative unit.**

The method flagged six suspicious companies, of which half were confirmed by the surveillance unit as identified fraudulent cases. The ability of this statistical approach to detect fraud validates the method as a valuable tool in pre-investigative analysis of durable medical equipment suppliers. This method allows evaluators to scan large populations of insurance claims in search of unusual provider behavior. The net payment amount of the flagged companies totaled $449,100, or about 5.9 percent of total payment made to fund incontinence supplies for Medicaid patients.

TABLE OF CONTENTS

LIST OF TABLES

Detecting Fraudulent Suppliers of Incontinence Briefs

Chapter 1: Introduction and Literature Review

The Centers for Medicare and Medicaid Services (CMS) estimate total health care spending in the US in 2008 reached $2.3 trillion or 16.2 percent of gross domestic product (GDP). Federal agencies tracking health care fraud estimate that anywhere from three to ten percent of expenses are fraudulent, with ten percent being the most accepted figure (U.S. Attorney's Office, Western District of Virginia). This implies that $230 billion a year is wasted due to fraudulent activities.

At the state level, Nevada's aggregate personal health care spending was 0.7 percent of national expenditures in 2004, at $10.9 billion spent on personal health care (NHE factsheet, CMS, 2010). At the 10 percent fraud estimate, about $1.9 million goes to illicit transactions in Nevada.

Fraud in medical cases may or may not include the actual delivery of health care, whereas abuse is the over-prescribing of unneeded medical treatment. Abuse by service providers can deliberately place patients at a higher risk by the performance of unnecessary procedures, invasive testing, and/or prescribing certain drug therapies (Li, Keui-Ying, Jionghua, and Jianjun, 2008). The Senior Health Insurance Counseling for Kansas (accessed 6/12/2010) defines fraud and abuse as follows:

> Fraud: Fraud is intentional deception or misrepresentation that an individual knows to be false or does not believe to be true, when that individual knows that the deception could result in some unauthorized benefit to himself or herself or some other person.

> The most frequent kind of fraud arises from a false statement or misrepresentation that is material to entitlement or payment under the Medicare program.
>
> Abuse: Incidents or practices of providers that are inconsistent with accepted and sound medical, business, or fiscal practices. These practices may directly or indirectly, result in unnecessary costs to the program, improper payment, or payment for services that fail to meet professionally recognized standards of care, or that are not medically necessary. Abuse involves payment for items or services when there is no legal entitlement to that payment and the provider has not knowingly misrepresented the facts to obtain payment.

The health care industry traditionally uses four types of controls: (1) claims processing with both human and automated audits and edits; (2) prepayment medical review; (3) post-payment utilization review; and (4) audits (Sparrow, 1998). Filters in the billing transaction process are a basic step in fraud control and prevention. Filters check certain variables or known ratios that indicate fraud. Examples of transaction-level tests include testing eligibility, price monitoring, and checking diagnoses/procedure codes relationships (Sparrow, 1998). Although this approach can detect certain types of fraud, its effectiveness is limited because detection relies on simple inquires that perpetrators circumvent overtime. Sophisticated criminals read the rulebooks and attend seminars to avoid detection through billing; they generally consider correct billing a minor inconvenience.

1.1.a Challenges to stopping fraud

There are many challenges to stopping health care fraud. First, social acceptance of the government and insurance companies as targets of fraud and society's degree of trust in medical professionals makes fraud control particularly difficult. Also, the database structure of raw claims forms need to be put into a usable format often by knowledge computer personnel. Third, the immense volume of patient records and reimbursement claim forms can challenge law enforcement. Another hindrance in fraud control efforts is the expurgation of public discourse of new fraud detection techniques to prevent alerting of the opposition. If criminals gain knowledge of how detection systems work, this could occlude the efficacy of new ideas before opportunity to detect fraud arises. 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).

Due to the adaptive nature of fraudsters and the underdevelopment and potential underuse of fraud control systems, contemporary revelations in fraud are more a matter of luck rather than systematic detection. The use of statistical sampling techniques offers a solution to circumvent the tedious task of individual examination of records on a case-by-case basis. By itself, statistical analysis cannot determine the perpetration of fraud; instead, the analysis simply alerts compliance officers to anomalous observations for suggestive investigation.

1.2 Durable Medical Equipment, Prosthetics, Orthotics, and Supplies (DMEPOS)

DME is defined as equipment that is appropriate for in-home use, benefits the patient medically, and can be used a repeated number of times. Medical supplies are disposable supplies that cannot be reused and are of no use in the absence of a medical condition. Four DME regional carriers (DMERC) process claims and provide reimbursement for DMEPOS. The CMS maintains a list of which HCPCS (Healthcare Common Procedure Coding System) codes are under DMERC jurisdiction and which are under area carrier jurisdiction (Medicare Claims Processing Manual). HCPCS codes are codes used to report the medical, surgical, diagnostic services and procedures, DMEPOS, and medical procedures performed or used and in need of reimbursement from government or private health insurance (GM Associates, accessed 1/24/2010).

When prescribing expensive DMEPOS items, physicians must submit a certificate of medical necessity (CMN) with their unique physician identification number (UPIN). Routine or random verification of CMN with the signing physician assists in verifying authentic DME claims. Fraudulent DME suppliers may use a broad range of referring physicians to thwart provider-level check efforts when they only have a limited number of beneficiaries on Medicare/Medicaid. Because verification is laborious and time consuming, a potential variable to check for in detecting fraud is how many referring physicians a DME supplier uses.

1.3 Literature Review

The literature review on fraud detection reveals three main strategies: auditing, supervised, and unsupervised methods. Auditing strategies, at best, use random stratification sampling methods to obtain samples from the spectrum of different claim types (Buddhakulsomsiri and Parthanadee, 2008). Auditing strategies cannot pinpoint suspicious claims from millions of claims in a data set. Because medical experts must be hired to review random claims case-by-case, this strategy is costly, time consuming, and inaccurate. As Ortega, Figueroa, and Ruz (2006) point out, finding the source of fraud (insured, provider, etc.) using statistical fraud detection methods is a far more efficient strategy than analyzing individual medical claims. However, statistical fraud detection is not a panacea. Each suspect claim needs a human interface to determine if it is indeed fraudulent (Sokol et al., 2001).

Statistical fraud detection methods can be classified as either supervised or unsupervised (Bolton and Hand, 2002). Supervised methods require samples from both known fraudulent and non-fraudulent records where characteristics for each type are modeled. The model yields a probability or suspicion score. Based on these models, new observations are assigned into one of the classes. One weakness associated with supervised methods is the needed guarantee of truly fraudulent claims and truly legitimate claims. Legitimate transactions far outweigh fraudulent ones. Creating models from these unbalanced classes can cause misspecification (Bolton and Hand, 2002). Another weakness is that supervised models cannot detect new types of fraud because the fraud models are created from past fraud strategies. Despite these weaknesses, supervised methods are widely used (Bolton and Hand, 2002). Neural networks (NNs), decision trees, fuzzy logic, and Bayesian networks enumerate some examples of supervised methods (Li et al., 2008).

Unsupervised methods tag outliers in a data set, and the outliers are then marked for potential investigation. An advantage of unsupervised methods is that they can detect new types of fraud. A disadvantage is that they detect accidental errors made by innocent providers, which is a different problem than fraud (Bolton and Hand, 2002). Clustering prevails as a popular tool for unsupervised fraud detection (Bolton and Hand, 2002). Hodge and Austin reviewed many common outlier detection techniques. Programs such as Electronic Fraud Detection and SmartSifter utilize unsupervised methods (Li et al., 2008). Prescription of DME varies significantly with patients' preferences and physicians' regimen (Iezzoni, 2002), so some detected outliers will actually be unobjectionable.

1.3.a Papers using auditing strategies

In an Office of Inspector General (OIG) Health and Human Services (HHS) 2002 investigation, auditors found that 21 percent (1,030) of claims failed to comply with Medicare rules and regulations (Improper Fiscal Year 2002 Medicare Fee-For-Service). Auditors examined 4,985 claims from 610 beneficiaries valued at $6.2 million. Compliance was based on (1) services rendered by certified Medicare providers to eligible beneficiaries; (2) proper reimbursement according to Medicare rules and regulations and; (3) medically necessary, correctly coded, and sufficiently documented. A medical review staff compared each claim to the medical records supporting services billed. Many OIG HHS investigations rely on audits.

Buddhakulsomsiri and Parthanadee (2008) presented a stratified random sampling plan to estimate billing accuracy in health care systems. Percent population accuracy and total dollar accuracy were the two measures of interest and were based on counts of binary responses (correctly or incorrectly processed). The paper compared various stratification methods including cluster analysis. The researchers found that selection of the optimal sampling method depends on what is being measured (percent accuracy or dollar accuracy) and what is known about the population structure.

Wickizer (1995) analyzed DME claims from two multi-county areas within a region served by an insurance carrier. Using a difference-in-difference method the researcher compared an intervention site to a control site from two multi-county areas. Nurse analysts reviewed DME claims to verify accuracy. As testament to the time factor in using audits, this study has a sample size of just 231 observations. Twenty-one months (January 1990 – September 1991) of claims data provided the variables used to measure DME utilization for four different types of DME. The number of order requests per month, submitted charges per month, Medicare-allowed payments per month, and the percentage of DME requests denied per month were the variables of interest. Data on other covariates including the number of hospital discharges per 1,000 Medicare beneficiaries, the number of primary care and specialist physicians per 1,000 in the population, the number of DME suppliers per 1,000 Medicare beneficiaries, and the number of nursing homes per 10,000 Medicare beneficiaries were gathered to control for external factors that could influence DME utilization. The findings show that DME utilization management programs reduced the number of requests, submitted charges, and Medicare payments in three out of the four targeted DME items.

1.3.b Papers Discussing Supervised Methods

Artìs et al. (2002) estimated the probability of fraud given different characteristics using a discrete choice model. They applied a multinomial logit model using utility maximization theory as the basis for consumer choice. Claims were classified into three categories – legitimate, fraud for personal profit, and fraud for third party benefit. The model estimated the probability for each category and the amount of influence the exogenous variables (individual characteristics) had on that probability. The data contained an oversampling of fraudulent claims in order to get a good representation for the group, making the sample non-random. Although the model was applied to automobile insurance fraud, it could be useful to study fraud in other fields.

He et al. (1998) used a multi-layer perceptron network to classify general practitioner (GP) profiles into categories ranging from normal to abnormal. Physicians, hired as expert consultants, indentified 28 features, which summarized a GPs 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 consultants classified all 1,500 profiles. Then the sample was divided into two groups with 750 profiles for the training set and 750 profiles for the test set. The one example of a feature they used was the proportion of initial-to-subsequent consultations. An appropriately practicing physician should have a low proportion of follow-up consultations. The researchers concluded that a two-class neural network classification system was a viable method for detecting fraud.

Ortega, Figueroa and Ruz (2006) held meetings with medical experts to assist in developing a set of variables used to discriminate between fraudulent and honest claims. After checking for correlation among features and consistency with empirical knowledge, 125 features were selected. To avoid training biases, two percent of the outliers were removed, and the features underwent linear normalization. Neural network models were applied to four entities: medical claims, the insured, the medical professional, and the employer. Feedback from the models was used as additional input in the other three sub-models.

Bolton and Hand (2002) defined links analysis, a different form of supervised methods. Links analysis uses the social network contacts of known fraudsters and searches for similar records in the database. For example, in telecommunications networks, security investigators have found that fraudsters seldom work in isolation from each other. Also, after an account has been disconnected for fraud, the fraudster will often call the same numbers from another account. Telephone calls from an account can thus be linked to fraudulent accounts to indicate intrusion.

1.3.c Papers Using Unsupervised Methods

Major and Riedinger (2002) conducted a large review (22,000 providers) of medical insurance claims to test an electronic fraud detection (EFD) program. The technique compared individual provider characteristics to their peers. Provider comparison groups should be similar, 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). The process creates a database with suspicious claims. Then, claims are grouped according to the provider. Providers who have many anomalies were referred to an investigative team.

Bolton and Hand (2002) used an unsupervised detection tool, called Peer Group Analysis for observing behavior over time. Objects were grouped according to common characteristics and then compared to their peers. If objects were reclassified into another group, or if they deviated too far away from the peer group, they were flagged as meritorious for investigation. Changing the number of objects in the peer group, affects the sensitivity of the analysis. Choice of a suitable distance (from peers) metric depends on the data – continuous, discrete, categorical or a mix of these variables. The method was applied to credit card accounts where the average amount spent over a four-week period of a target account was compared to its peer group.

Another method used for credit card fraud detection is named Break Point Analysis. It operates at the account level in points of observation or time. Window parameters were set to include a certain number of past observations. New observations were compared to past behavior. Sudden increases in frequency or amount of credit card transactions can be indicators of fraudulent behavior. An advantage of Break Point Analysis is that it does not require a summarization of the data at fixed time points, e.g. weekly.

Geyer and Williamson (2004) reviewed three Benford's Law methods for fraud detection. 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 (1992) first applied Benford's Law to fraud detection (for tax data) by developing the Distortion Factor (DF) Model. The DF model compares the first digit frequencies of observations in a data set to the expected frequencies of Benford's Law. The data set is then classified as overstated or understated, and the DF model gives the extent of over or under statement. Ley used a Bayesian approach on one-day returns on the Dow Jones Industrial Average and the Standards and Poor's Index. The results indicated that the DJIA and the S&P follow Benford's Law. Table 1 shows the probabilities of the first significant digits of Benford's Law. Benford's law is scale and base invariant.

Table 1: Benford's Law Distribution of first digits

Table 1: The percentages of the first significant digits of Benford's Law									
First Digit	1	2	3	4	5	6	7	8	9
Probability	30.1%	17.6%	12.5%	9.7%	7.9%	6.7%	5.8%	5.1%	4.6%

Chapter 2: Data and Method

Both Li et al. (2008) and Sokol et al. (2001) estimate that data preparation consumes 80 percent of the time in a fraud detection project. Dispite the importance of this step, only three papers formally addressed it (Li et al., 2008; Sokol et al., 2001; Lin and Haug, 2006). Database structures of raw claims data and electronic health records are designed to support financial transactions and health care delivery and must be reshaped to support data analysis operations. In 'flattening' a dataset, one transforms the desired party of examination (patient, supplier, etc.) into rows and the variables into columns. Dana Edburg, Phd., University of Nevada, Reno and her IT students cleaned and processed the data for this project.

A state Medicaid agency provided claims data that linked provider, facility, and prescription claims over a five-year period from January 2005 to December 2009. During this time, Medicare/Medicaid reimbursed 693 DME supply companies for $87,340,766 worth of equipment. The data of interest for this study was the characteristics of DME suppliers enrolled to provide services to Medicaid patients.

A profile consisting of 12 features was created for each of the 321 DME suppliers providing incontinence briefs. If a DME supplier fell into the outlier range of any of the features, the company was assigned a point. The total points were tallied with a maximum of 12 possible points and a minimum of zero. The more points a provider has, the more suspicious the company looks. Below are the 12 features examined:

1. **Diapers per claim –** Providers in the upper echelons were marked as suspicious. Medicaid rules limit patients to 300 diapers per month. The more diapers supplied per claim, the more money a fraudulent company can make.
2. **Percent of extra-large diapers**– Providers in the upper echelons were marked as suspicious. A fraudulent organization would be keen on making a quick profit, thereby, submitting orders for the most expensive items.
3. **Percent of patients below 60 –** Companies that had a high ratio of younger patients were marked as suspicious. Older patients are more likely to need diapers. A phony DME supplier may have access to a limited number of patients; thereby, they do not have the luxury of selecting older patients.
4. **Number of different supplies provided –** A small number of items billed earned the company a point. Fictitious organizations may not know the in-and-outs of medical terminology. They may prefer to stick with only a couple of items to bill and limit the number of denied claims to avoid examination.
5. **Number of supplies per patient –** A high ratio was regarded as suspicious. It would be unlikely for a person to order many different sizes of a diaper. Sure, a person can be trying to find the right size, but too much variation is suspicious. Also, suppliers may be trying to add to their bill by supplying unneeded additional items.
6. **Number of operating months –** Companies that only operated for a limited time were assigned a point. Fraudulent companies wishing to make a large amount of money in a short amount of time will likely operate on a short-term basis. Additionally, a company may be testing the waters by submitting claims to see if they

are denied. Once they figure out a combination that works, they may start a different company for mass billing to avoid payment spikes that may alert an investigation.

7. **Charge submit amount minus charge allow amount –** A small difference between the two earned the company a point. Phony suppliers will have to fabricate a charge submit price because they won't have that information from the DMEPOS wholesale invoice. A scamming company may not be original in this field and choose something close to the charge allow amount.
8. **Third party amount per patient –** Companies that have a low number were given a point. This variable is based on a per-patient level to normalize DME supply companies based on their size. The billing process requires that the third-party pay first and the DME supplier needs to submit an explanation of benefits from the third party before Medicaid pays. Therefore, a fraudulent company will not want to hassle with the extra paperwork and time needed when a third party is involved.
9. **Denied procedure codes per patient –** A high number of denied claims per patient earned the DME supplier a point. Fraudulent suppliers likely do not have access to a patient's current status. The patient may have recently entered a hospital, nursing home, or may have died making the patient ineligible to receive DME. Procedure codes were chosen over denied claims because many items can be submitted on one claim. Suppliers testing to see if items and related information pass the payment test, may submit multiple items per claim to avoid having a high number of denied claims.
10. **Number of distinct primary DX per person.** Companies that did not use a variety of diagnosis codes (a small ratio) earned a point. Fraudulent companies may prefer to stick with what they know works. Once they find a diagnosis code that passes the

internal audit, they may be reluctant to try a different one for fear of denied claims and an ensuing audit.

11. **Number of preauthorization requests -** A small number of preauthorization requests earned the company a point. Fictitious organizations may prefer to limit their exposure to the system; whereas, legitimate companies will likely require pre-authorizations at some point.
12. **Percent of dollar total that was related to incontinence diapers –** A large ratio earned the company a point. Fraudulent companies may pick a certain category to specialize in; whereas a legitimate company may not be able to profitably survive by supplying just one category of items, especially if the item is not a high priced one like incontinence briefs.

Chapter 3: Results

Table 2 displays the cutoff amount for each of the variables according to either the upper or lower tails of the distribution.

Table 2: Cutoff amount for the variables

Variable	Upper 5th percentile	Lower 5th Percentile
Variable 1	272.5	
Variable 2	68.4	
Variable 3	77.8	
Variable 4		2.0
Variable 5	1.7	
Variable 6		1.0
Variable 7		8.7
Variable 8		0.0*
Variable 9	3.0	
Variable 10		0.03
Variable 11		0.0**
Variable 12	72.5	

*75 percent of the companies had zero.
**90 percent of the companies had zero

Variable 1: Diapers per claim – The first variables is perhaps the most interesting. If a supplier consistently orders 300 diapers per claim for multiple patients, this means 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. Teresa Serratt, PhD. in University of Nevada's Orvis School of Nursing explained that initially a patient will require more briefs because they need to stock up around the house and other places they might be. 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.

Anything over 272.5 was in the upper 95 percentile. Using this metric, I flagged 16 companies listed in the table below.

Table 3: Companies flagged under variable 1

Company ID	Diapers Per Claim	Total Claims
3302100	300	27
3302448	300	576
100503886	300	3
100505813	300	3
100508690	300	826
100509994	300	130
3302827	298	192
100510601	297	1
3302790	297	9
100500017	296	62
3302995	295	4
3302044	294	7
3302133	289	13
100501150	288	2
100509795	288	1
100503791	274	2076

Variable 2: Percent of extra-large diapers– Next, I examined the outliers in the percentage of extra-large diapers ordered. To avoid penalizing a company that has a long-running relationship with a client, the percentage of patients ordering expensive items was calculated instead of the percentage of extra large diapers ordered. If the combination of a supplier, product code, and patient id matched, than all orders with that same combination were grouped together. A problem with grouping order this way, is that if a patient switches brief size, they would be counted as two separate patient numbers. Most companies (221) served patients that either switched brief sizes or ordered more than one product. 99 companies consistently had one product for each

patient. Table four displays prices for incontinence briefs to illustrate the motivation a company would have in ordering extra large briefs.

Table 4: Product pricing for incontinence briefs

Procedure Code	Description	Price
T4521	Adult size brief/ diaper sm	$0.56
T4522	Adult size brief/diaper med	$0.60
T4523	Adult size brief/diaper lg	$0.81
T4524	Adult size brief/diaper xl	$0.97
T4525	Adult size pull-on sm	$0.84
T4526	Adult size pull-on med	$0.84
T4527	Adult size pull-on lg	$0.88
T4528	Adult size pull-on xl	$1.01
T4529	Ped size brief/diaper sm/med	$0.43
T4530	Ped size brief/diaper lg	$0.47
T4531	Ped size pull-on sm/med	$0.54
T4532	Ped size pull-on lg	$0.54
T4533	Youth size brief/diaper	$0.56
T4534	Youth size pull-on	$0.80
T4535	Disposable liner/shield/pad	$0.37
T4541	Large disposable underpad	$0.41
T4542	Small disposable underpad	$0.41
T4543	Disp bariatric brief/diaper	$1.96

The top two most expensive adult briefs are the adult size pull-on extra large and the bariatric brief. Companies with less than five patients were deleted. DME suppliers with more than 68.4 percent of their patients needing the most expensive items are given a point. Using these results, nine companies were tagged as shown in Table 5.

Table 5: Companies flagged under variable 2

Company ID	% extra large	# of patients	# of claims
3302973	100	7	15
100508690	100	138	826
100509994	100	45	130
100510770	100	7	7
3302448	83.6	152	576
3307840	83.3	5	29

3301853	81.8	10	75
3316122	73.9	18	109
3316859	68.4	14	120

Variable 3: Percent of patients below 60 – This variable was also prone to penalizing DMEPOS companies that had a long-run relationship with patients under age 60. To avoid this, the data was grouped the same way as explained in the previous variable. Over the five year period, a person could have transitioned to a new age group. Out of the 321 companies, 145 had multiple age groups per person because patients switched age groups as they get older or the claim form filed incorrectly. Companies that supplied items for less than ten people were excluded before the summary was run. Companies with more than 77.8 percent of their patients below 60 are given a point. Using this metric, 28 companies were flagged as shown in Table 6.

Table 6: Companies flagged under variable 3

Company ID	% below 60	Count of patients
100504115	100.0	10
100503857	100.0	14
3302080	100.0	15
3302140	95.2	22
3316170	95.2	37
3302243	95.0	16
3316175	93.8	25
3302683	93.3	14
3302102	92.3	261
3302620	92.3	12
100511657	91.4	203
100509417	91.3	21
3302335	90.9	10
3316160	90.0	10
3302030	89.7	149

3316122	87.5	18
3316001	87.5	26
3316157	85.3	25
3316158	84.6	12
100509159	84.6	12
3313006	81.3	12
100513241	80.0	10
3302125	80.0	18
3316002	79.3	28
100505908	78.6	11
3316005	78.4	79

Variable 4: Number of different supplies provided –Companies that supplied items for less than 5 people were excluded before the summary was run. Companies that supplied only two items were given a point. Using this metric, eight companies were flagged.

Table 7: Companies flagged under variable 4

Company ID	# of patients	Supplies per patient
100502592	6	1.5
100509476	8	1.4
3302129	6	1.3
100504397	8	1.3
100509512	8	1.3
100513239	6	1.2
100505551	12	1.2
3302401	7	1.1

Variable 5: Number of supplies per patient – Companies that supplied items for less than five people were excluded before the summary was run. Anything over 1.14 would be in the upper 95 percentile. Using this metric, 16 companies were flagged and listed in the table below.

Table 8: Companies flagged under variable 5

Company ID	Supplies per patient	# items supplied
3302933	5	5
3302783	4	4
100504450	3	3
3302019	3	3
100510709	3	3
100511485	3	3
100500837	3	6
3316500	2	2
100503886	2	2
100509421	2	2
100514638	2	2
3302532	2	2
100509472	2	4
100509413	2	4
100509494	1.8	9
3302029	1.75	7

Variable 6: Number of operating months – Only 17, or about five percent of companies had more than 40 months of operation months. This suggests a high turnover rate in the industry. Anything below one is in the lower 5th percentile. Using this metric, I flagged 27 companies listed in the table below.

Table 9: Companies flagged under variable 6

Company ID	# of operating months	1st Payment Month	Last Payment month	No. of claims
3302529	1	200505	200505	1
3302505	1	200507	200507	1
100505393	1	200602	200602	1
3313150	1	200606	200606	2
3302191	1	200609	200609	1
3304009	1	200611	200612	3
3302530	1	200612	200612	1
100510770	1	200705	200708	11
100511913	1	200707	200707	1
100510601	1	200710	200710	1
3388078	1	200712	200712	1
100509441	1	200801	200801	1

100509504	1	200802	200802	2
100509456	1	200810	200810	1
3302990	1	200810	200810	1
100509419	1	200811	200811	1
3312060	1	200811	200811	1
100514786	1	200907	200908	2
100515391	1	200908	200908	1
100510205	1	200909	200909	1
100509795	1	200910	200910	1
100518252	1	201001	201001	3
100517752	1	201001	201001	6
100517934	1	201001	201001	2
100509721	1	201001	201001	1
100506835	1	201003	201003	1
100517590	1	201005	201005	1

Variable 7: Charge submit amount minus charge allow amount –Companies that had less than 20 claims were eliminated because there may not be enough variation in those cases. Anything below 8.68 would be in the lower 5 percentile. Using this metric, I flagged nine companies listed in the table below.

Table 10: Companies flagged under variable 7

Company ID	Charge Submit – Charge	# of unique claims
3302448	0	576
3302556	0	34
3302580	0	48
100500017	0	40
100508690	0	826
100509994	0	130
100515628	0	248
100518238	0	23
100503124	8.34	27

Variable 8: Third party amount per patient – Companies that supplied items for less than five people were excluded before the summary was run. Only 18 of the remaining companies had payment from either a third party or Medicare. All companies that had

payment from Medicare also had third party payments. Because such a large number of companies do not seek third party payment, the cutoff at the zero dollar amount captured 75 percent of the sample. Each of the remaining 148 companies was given a point. The table displaying the tagged companies was skipped due to the large number selected.

Variable 9: Denied procedure codes per patient – Anything over three was in the upper 95 percentile. Using this metric, five companies were flagged and listed in the table below.

Table 11: Companies flagged under variable 9

Company ID	# Denied PROC_CD	Denied claim per patient
3302389	15	5
3302935	4	4
3313095	4	4
3316008	7	3.5
3302913	108	3

Variable 10: Number of distinct primary DX per person. Anything under 0.028629 was in the lower fifth percentile. Using this metric, 16 companies were flagged and listed in the table below.

Table 12: Companies flagged under variable 10

Company ID	Primary DX/Patients	Number of patients	Number of distinct primary DX
100500017	0	8	0
3316927	0	66	0
3302851	0	2	0
100503909	0	37	0
100508690	0	138	0
100509994	0	45	0
100510770	0	7	0
100509795	0	1	0
100509084	0	21	0

3302448	0.006579	152	1
100513556	0.007246	414	3
100503791	0.010033	299	3
3302030	0.013423	149	2
3302662	0.01548	323	5
3316080	0.022222	45	1
3302827	0.025	40	1

Variable 11: Number of pre-authorization the company needed. **–** Of the 320 DME companies that supplied briefs, 19 obtained preauthorization for specialized orders. The remaining 301 companies were given a point. Because of the large number of companies without preauthorization numbers, this table was skipped.

Variable 12: Percent of dollar total amount that was related to incontinence diapers

Companies supplying less than or equal to five patients were dropped before the summary was run. Any company whose percent total net pay in incontinence briefs was 72.5 percent of their total net pay for all DMEPOS items was upper fifth percentile. Using this metric, 8 companies were flagged and listed in the table below.

Table 13: Companies flagged under variable 12

Company ID	Total Net Pay	Total net pay related to diapers	% net pay related to diapers	Total patients
3302081	847.05	782.71	0.92	7
3302140	8965.82	6508.77	0.73	22
3302448	179760	146160	0.81	152
3302973	46.98	46.98	1.0	7
3313006	7459.99	5406.74	0.72	12
3316122	10222.76	8846.56	0.87	18
3316154	24725.08	19283.42	0.78	31
3316160	12673.18	9478.54	0.75	10

Tallying the Points - Finally, points across all variables were totaled to reveal the most suspicious of the companies. Six with the highest scores were selected and displayed in table 14.

Table 14: The most suspicious companies

Company ID	1	2	3	4	5	6	7	8	9	10	11	12	Total	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

The total amount of money spent of suspicious claims totals to $449,100, or 5.9 percent of the total amount spent on incontinence briefs during this five year period.

3.1 Future Refinements

Judging the results, some variables failed to categorize the supplier into useful groups. For example, in the third party amount per patient, variable 8, only 19 companies received payments from third parties. This essentially grouped the suppliers into two categories where either they did or they did not have payment from third parties and most did not. Another example where this occurred is the number of preauthorization requests, variable 11. The vast majority of companies never obtained a preauthorization. On the other hand, features one and two (diapers per claim and percent extra large) seem to be more important and had definite right tail outliers as determined by a histogram.

Total billed charges of some flagged companies were less than $1,000; whether these billings were fraudulent is insignificant. For example, in feature 1 (diapers per claim), a few companies had a high average of diapers per claim, but when looking at the total claims submitted, they would be regarded as insignificant. Six out of 16 companies had less than five submitted claims.

To improve this model, a weighting system can be applied in two ways. (1) Give the variables higher weight if they are potentially more likely to flag fraudulent companies. Variables 1 and 2 should carry more weight than variables 7, 8, and 11. (2) Variables results should be weighted by the size of patients served or claims submitted. Referring to variable 4 (number of unique items supplied), some companies one supplied one item, but the number of patients served is also very small. If this variable is weighted by number of patients, it would refine the point system to emphasize established companies, not companies whose have netted insignificant amounts. This recommendation also holds for most of the variables examined.

3.2 Tips for implementation

If a fraud task force wished to implement this method, there are three important considerations. (1) When taking the upper or lower fifth percentile, the cutoffs are assigned based on statistics, not logic. Thought should be given to whether the cutoff divides the groups into questionable and likely benign categories. If not, the weight for that variable should be lowered. (2) One concern 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 with some minor adjustment to catch criminals before they go out of business.

Chapter 4: Conclusion

A popular accepted figure for health care fraud is ten percent of expenditures. Fraud in durable medical equipment is especially attractive because one does not need any medical training and the transactions are faceless. Once a company passes the initial inspection by the fraud control unit, they can theoretically bill as much as they want without ever having face-to-face interactions with the state.

Recent health care legislation expects to control rising health care expenditures by combating fraud, waste and abuse. In order for this ideal to become a reality, a statistical approach needs to be taken. Of the three methods found in the literature (auditing, supervised, and unsupervised), only the unsupervised method was practical in this case. Auditing methods are time consuming and inefficient. Modeling fraudulent claims to identify similarities was not an option either due to a lack of known fraudulent claims that the State could provide. Thus, identifying outliers in variables, or a unsupervised method, was chosen as the preferred method.

Twelve variables were created and the outliers in each case were assigned a point. Next, the companies' points were tallied, and the ones with the most points were marked as suspicious. Of course, this does not necessarily mean they are fraudulent. Only an investigation can determine that. This statistical method can be used as a tool to alert compliance officers to anomalous observations. The method used in this study flagged six suspect companies whose total payments exceeded $449,000. More importantly, after presenting the results to the State, three of the flagged companies were confirmed by the fraud control unit as fraudulent.

Before applying this strategy to real-time claim data, some important revisions should be made to enhance its effectiveness. Most notably, a weighting system should be incorporated that will account for a company's size and how reliable a variable was at pinpointing suspicious tactics.

References

Artìs, Manuel; Ayuso, Mercedes; Guillèn, Montserrat. "Detection of Automobile Insurance Fraud with Discrete Choice Models and Misclassified Claims." *The Journal of Risk and Insurance*, vol. 69, no. 3 (2002) 325-340.

Bolton, Richard J. and Hand, David J. "Statistical Fraud Detection: A Review." *Statistical Science*, vol. 17, no. 3, August 2002, pp 235-255.

Buddhakulsomsiri, Jirachai and Parthanadee, Parthana. "Stratified random sampling for estimating billing accuracy in health care systems." *Health Care Manage Sci*, vol. 11, 2008, pp 41–54.

Centers for Medicare and Medicaid Services. *National Health Expenditure Data-Historical*. Updated 01/05/2010. US Department of Health and Human Services. Date accessed: 01/11/2010.

Centers for Medicare and Medicaid Services. *Medicare Claims Processing Manual: Chapter 1 – General Billing Requirement.* Revised 09/04/2009. US Department of Health and Human Services. Accessed 01/24/2010. http://www.cms.hhs.gov/manuals/downloads/clm104c01.pdf

"FAQ's." *GM Associates, Inc.* Date accessed 01/24/2010. http://gmassocinc.com/7-faq.htm

GAO: US Government Accountability Office. *Improper Payments: Progress Made but Challenges Remain in Estimating and Reducing Improper Payments*. April 22, 2009. GAO-09-628T

Geyer, Christina Lynn and Williamson, Patricia Pepple. "Detecting Fraud in Data Sets Using Benford's Law." *Communications in Statistics*, vol. 33, no. 1, 2004, pp. 229-246.

He, Hongxing; Wang, Jincheng; Graco, Warwick; Hawkins, Simon. "Application of Neural Networks to Detection of Medical Fraud." *Expert Systems With Applications*, vol. 13, no. 4, 1997, pp. 329-336.

Hodge, V.J. and Austin, J. "A survey of outlier detection methodologies." *Artificial Intelligence Review*, vol. 22, no 2, pp. 85-126, 2004.

Iezzoni, Lisa I. "Using Administrative Data to Study Persons with Disabilities." *The Milbank Quarterly*, Vol. 80, No. 2, pp.347-379, 2002.

Li, Jing; Keui-Ying, Huang; Jionghua, Jin; Jianjun, Shi. "A survey on statistical methods for health care fraud detection." *Health Care Management Science*, vol. 11, no. 3, pp. 275-87, September 2008.

Lin, Jau-Huei and Haug, Peter J. “Data Preparation Framework for Preprocessing Clinical Data in Data Mining.” Department of Biomedical Informatics, University of Utah and Intermountain Healthcare. AMIA 2006 Symposium Proceedings pp. 489-493.

Major, John A. and Riedinger, Dan R. “EFD: A Hybrid Knowledge/Statistical-Based System For The Detection of Fraud. *The Journal of Risk and Insurance*, vol. 69, no. 3, pp. 309-324, 2002.

Nigrini, MJ. “The Detection of Income Tax Evasion Through an Analysis of Digital Frequencies.” PhD thesis, University of Cincinnati, OH, USA. 1992.

Ortega, Pedro A., Figueroa, Christian J., and Ruz, Gonzalo A. “A Medical Claim Fraud/Abuse Detection System based on Data Mining: A Case Study in Chile.” (2006) http://www.mec.cf.ac.uk/~scegr2/pub/DMI5560.pdf. Accessed 06/15/2010.

SHICK: Senior Health Insurance Counseling for Kansas. Definition and Examples of Fraud and Abuse. http://www.aging.ks.gov/SHICK/Fraud_Abuse/Examples.htm. Accessed 06/12/2010.

Sparrow, Malcolm K and Travis, Jeremy. “Fraud Control in the Health Care Industry: Assessing the State of the Art.” *National Institute of Justice: Research in Brief* (1998) http://www.ncjrs.gov/pdffiles1/172841.pdf. Accessed 22 December 2009.

Sokol, Lisa; Garcia, Bob; Rodriguez, Justin; West, Mike; Johnson, Kirk. “Using Data Mining to Find Fraud in HCFA Health Care Claims.” *Top Health Information Management*, vol. 22, no. 1, pp. 1-13, 2001.

United States Attorney's Office: Western District of Virginia. Health Care Fraud. http://www.justice.gov/usao/vaw/health_care_fraud/index.html . Accessed 06/15/2010.

US Department of Health and Human Services: Office of Inspector General. “Improper Fiscal Year 2002 Medicare Fee-For-Service Payment.” A-17-02-02202. January 2003.

Wickizer, Thomas M. “Controlling Outpatient Medical Equipment Costs through Utilization Management.” *Medical Care*, Volume 33, Number 4, 1995: pp 383-391.