View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Procedia Technology 9 (2013) 1252 – 1264

Procedia Technology

CENTERIS 2013 - Conference on ENTERprise Information Systems / ProjMAN 2013 - International Conference on Project MANagement / HCIST 2013 - International Conference on Health and Social Care Information Systems and Technologies

Predicting Healthcare Fraud in Medicaid: A Multidimensional Data Model and Analysis Techniques for Fraud Detection

Dallas Thornton[a]*, Roland M. Mueller[b], Paulus Schoutsen[a], Jos van Hillegersberg[c]

[a] *San Diego Supercomputer Center, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA, 92093-0505, USA*
[b] *Department of Business and Economics, Berlin School of Economics and Law, Badensche Straβe 52, 10825 Berlin, Germany*
[c] *Department of Information Systems and Change Management, University of Twente, Drienerlolaan 5, 7522 NB Enschede, Netherlands*

Abstract

It is estimated that approximately $700 billion is lost due to fraud, waste, and abuse in the US healthcare system. Medicaid has been particularly susceptible target for fraud in recent years, with a distributed management model, limited cross-program communications, and a difficult-to-track patient population of low-income adults, their children, and people with certain disabilities. For effective fraud detection, one has to look at the data beyond the transaction-level. This paper builds upon Sparrow's fraud type classifications and the Medicaid environment and to develop a Medicaid multidimensional schema and provide a set of multidimensional data models and analysis techniques that help to predict the likelihood of fraudulent activities. These data views address the most prevalent known fraud types and should prove useful in discovering the unknown unknowns. The model is evaluated by functionally testing against known fraud cases.

© 2013 The Authors Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
Selection and/or peer-review under responsibility of SCIKA – Association for Promotion and Dissemination of Scientific Knowledge

Keywords: Fraud Detection; Predictive Analytics; Healthcare; Medicaid; Data Modeling

* Corresponding author. Tel.: +1-858-534-8364; fax: +1-858-225-3661.
E-mail address: dallas@sdsc.edu.

2212-0173 © 2013 The Authors Published by Elsevier Ltd. Open access under CC BY-NC-ND license.
Selection and/or peer-review under responsibility of SCIKA – Association for Promotion and Dissemination of Scientific Knowledge
doi:10.1016/j.protcy.2013.12.140

1. Introduction

Roughly one-third of the $2.7T spent on healthcare in the US is attributable to fraud, waste, and abuse [1]. Payers for healthcare services must deal with fraudulent practitioners, organized criminal schemes, and honest providers who make unintended mistakes while billing for their legitimate services. The US Medicaid system is particularly susceptible fraud and abuse, since it is harder to exclude problematic providers and is managed separately and with limited coordination across the states. Each state has program sovereignty and maintains its own eligibility and benefits criterion, making nationwide initiatives challenging.

While there is undoubtedly an abundance of fraud in the system, states and the federal government lack sophisticated fraud control systems. Current systems are static, lack real time detection, and focus on detection in specific claim transactions without looking for patterns of suspicious behavior over time and the relationships and interactions between relevant entities.

Fraud control is a risk management activity akin to others but with its unique challenges. In June 2002, Donald Rumsfeld, then United States Secretary of Defense, succinctly addressed the challenge [2]: “The message is that there are no ‘knowns’. There are thing we know that we know. There are known unknowns. That is to say there are things that we now know we don't know. But there are also unknown unknowns. There are things we don't know we don't know. So when we do the best we can and we pull all this information together, and we then say well that's basically what we see as the situation, that is really only the known knowns and the known unknowns. And each year, we discover a few more of those unknown unknowns.”

This challenge of unknown unknowns faces program administrators looking to root out fraud in healthcare. New and existing data must be integrated, mastered, and utilized in new ways to discover the unknown unknowns. While we can be assured that this cat and mouse game will continue as fraudsters adapt to our new knowns and their countermeasures, we can undoubtedly make significant strides in detecting potential patterns of fraud, waste, and abuse in the system, ending these vulnerabilities and removing known bad actors.

In this paper, we applied Hevner et al. [3] to help us develop a framework for fraud detection in Medicaid that provides specific data models and techniques that identify the most prevalent fraud schemes and should help identify the unknown unknowns. The environment, discussed in the first section, includes the payers, providers, and patients. The knowledge base, covered in section two, is represented by fraud detection literature and the state of the industry. Based on this analysis, in the third section we develop a multidimensional schema based on Medicaid data and describe a set of multidimensional models and techniques to detect fraud in large sets of claim transactions. In the fourth section, these artifacts are evaluated through functionally testing against known fraud schemes. As in the domain of risk management as a whole, healthcare fraud control must address the unknown unknowns. This paper offers a set of multidimensional data models and analysis techniques that can be used to detect the most prevalent known fraud types and should prove useful in detecting the unknown unknowns.

2. Environment

The following definition of fraud from the US Department of Health and Human Services [4] will be used for the purposes of this paper: “Fraud is the intentional deception or misrepresentation that an individual knows to be false or does not believe to be true and makes, knowing that the deception could result in some unauthorized benefit to himself/herself or some other person.” Within the healthcare system three main parties commit fraud [5]: healthcare providers, beneficiaries (patients), and insurance carriers. Providers are the initiating actor for billing insurers, and, as such, quickly become the nexus for fraud schemes. When a provider participates in Medicaid, the provider agrees to the reimbursement rates set by the state and submits claims for payment directly to the state or managed care entity. If the provider is not participating in Medicaid, the provider sends the patient the bill, which he or she pays before requesting Medicaid reimbursement. The agency or insurer processes the claim and sends an explanation of benefits to the patient that describes the services paid for along with their codes and costs.

States operate claims processing systems that perform various prepayment checks and edits to inspect the claim's legitimacy. Edits and audits verify information with honest providers in mind, but they are not designed to detect fraud schemes of any depth [6]. These systems simply cannot verify that the service was provided as claimed, that the diagnosis is correct, or whether the patient is even aware of the claimed services.

Table 1 provides an analysis of the primary Medicaid actors that guides our antifraud framework design.

Table 1. Medicaid Environment Overview

Actor	Patient	Provider	State Medicaid Agency / Insurer	CMS (Federal)
Roles	Enroll in Medicaid. Receive care. Receive EOB.	Enroll with Medicaid. Provide care. Submit claims. Receive payment.	Sets (comparatively low) reimbursement rates. Enroll beneficiaries and providers. Pay legitimate claims. Prosecute fraudulent claims. Provide EOB to patient.	Pay state matching funds on claims. Ensure state matching funds are well-spent.
Capabilities	Could: commit, conspire to commit, or report fraud.	Could: commit, conspire to commit, or report fraud.	Can analyze patients, providers, and claims within its jurisdiction.	Could simplify data sharing across states. Could provide common tools for states to use for detecting fraud.
Characteristics	Wants to receive quality care at low out-of-pocket costs. Insurance identity theft has little direct impact on patient.	Desire quick reimbursement for services rendered. Unhappy with low reimbursement rates. Will opt-out of participation with significant burdens. Bad actors can sap millions quickly.	Want to reduce fraud / waste / abuse. Tight state budgets limit operational dollars available to combat fraud. Afraid of discovering fraud that is unrecoverable, as state is responsible for both loss and to reimburse federal matching funds. Afraid of impositions on providers pushing them out of the system, reducing access to care.	Want to reduce fraud / waste / abuse. Afraid of impositions on providers that may push them out of the system, reducing access to care. Afraid of impositions on states that may reduce cooperation with federal initiatives such as ACA.
Fraud / Anti-Fraud Strategies	Could: receive kick-backs, sell credentials, receive free services, or simply look the other way.	Could: phantom bill, up code, unbundle, miscode, bribe patients, perform unnecessary services, or refer patients to collusive providers. Could also sell credentials for billing and/or be extorted by organized crime.	Checks claims against known 'edits'. Performs some data analysis of state claims paid as a source of audits. Audits providers for reasonableness and accuracy. Prosecutes blatant fraud through the courts system. Excludes proven fraudulent providers.	Aggregates data at a national level for analysis. Supports focused state and interstate collaborations, auditing providers with the state and providing funding for specific anti-fraud collaboration efforts. Provides training to state staff. Prosecutes blatant fraud legally. Excludes fraudulent providers.
Structure and Culture	Millions of independent actors.	Millions of independent actors.	Struggle between pleasing providers and finding fraud. Program integrity usually in separate silo away from payment and	Struggle between pleasing providers and finding fraud. Program integrity usually in separate silo away from payment

			enrollment operations.	and enrollment operations.
Processes	Receive services. EOBs sent to patient by insurer.	Enroll in Medicaid. Provide care. Bill for care. Respond to audits. Maintain records.	Provider and beneficiary enrollment. Claims payment process. Claims data extract process for CMS. Provider audit processes. Audit findings extrapolation process.	Medicare provider and beneficiary enrollment. State data quality processes. State audit support and collaboration processes.

3. Knowledge Base

3.1. Classifying fraud

According to Sparrow [6] there are two different types of fraud: “hit-and-run” and “steal a little, all the time”. “Hit-and-run” perpetrators simply submit many fraudulent claims, receive payment, and disappear. “Steal a little, all the time” perpetrators work to ensure fraud goes unnoticed and bill fraudulently over a long period of time. The provider may hide false claims within large batches of valid claims and, when caught, will claim it an error, repay the money, and continue the behavior. The FBI [7] highlights and categorizes some of the most prevalent known Medicaid fraud schemes:

- Phantom Billing – Submitting claims for services not provided.
- Duplicate Billing – Submitting similar claims more than once.
- Bill Padding – Submitting claims for unneeded ancillary services to Medicaid.
- Upcoding – Billing for a service with a higher reimbursement rate than the service provided.
- Unbundling – Submitting several claims for various services that should only be billed as one service.
- Excessive or Unnecessary Services – Provides medically excessive or unnecessary services to a patient.
- Kickbacks – A kickback is a form of negotiated bribery in which a commission is paid to the bribe-taker (provider or patient) as a quid pro quo for services rendered [8].

Sparrow [6] proposes that for effective fraud detection one has to look at the data beyond the transaction level, defining seven levels of healthcare fraud control (see Table 2).

Table 2. Levels of healthcare fraud control, adapted, from Sparrow [6]

		Level Focus
Level 1	Single Claim, or Transaction	The claim itself and the related provider and the patient.
Level 2	Patient / Provider	One patient, one provider, and all of their claims.
Level 3	a. Patient	One patient and all of its claims and related providers.
	b. Provider	One provider and all of its claims and related patients.
Level 4	a. Insurer Policy / Provider	Patients that are covered by the same insurance policy and are targeted by one provider.
	b. Patient / Provider Group	One patient being targeted by multiple providers within a practice.
Level 5	Insurer Policy / Provider Group	Patients with the same policy being targeted by multiple providers within a practice.
Level 6	a. Defined Patient Group	Groups of patients being targeted by providers. (e.g. patients living in the same location)
	b. Provider Group	Groups of providers targeting their patients. Groups can be providers within the same practice, clinics, hospitals, or other arrangements.
Level 7	Multiparty, Criminal Conspiracies	Multiparty conspiracies that could involve many relationships.

Each higher level involves larger fraud schemes with more people involved and an increased difficulty of being detected. According to Sparrow [6] the bulk of the industry's detection toolkit is focused at level 1 and 3. Before payment, the transaction (level 1) and patient level (level 3a) may be evaluated. For example, are there claims for multiple childbirths within 9 months? Post payment analysis may focus on the provider level (level 3b). For example, is a doctor billing more hours of office visits than possible?

3.2. Healthcare fraud detection literature

The literature about fraud within the healthcare can be divided into three categories. The first category provides an overview of the field. It focuses on what kind of statistical methods can be used. For example, Travaille et al. [9] created an overview of statistical methods used by fraud detection within other industries and how they can be applied within the healthcare industry. Li et al. [5] surveyed healthcare industry methods and found combinations of unsupervised and supervised methods used together with profiling. The second category provides results on actual applications of the methods to find its usefulness in detecting fraud. For example Copeland et al. [10] discussed unsupervised methods to find Medicaid fraud within Nevada. Yang and Hwang [11] looked at using the order of which services are performed for fraud detection. This category helps in choosing a method for fraud detection by being able to compare the results of individual methods. The third category is focused on general methods and models to improve fraud detection. For example Morris [12] describes five key components how the current health system has to change to better battle fraud. Major and Riedinger [13] describe a workflow and system to setup fraud detection departments with results of its use in the real world. Similar work was done by Ortega et al. [14], who introduced a data mining based system that decreased the time it takes to detect fraud by 76% from an average of 8.6 months to 2 months. Because Major and Riedinger and Ortega et al., describe real systems that are used to find fraud they cannot go into details of the exact working of the systems. Doing this would give fraud perpetrators an advantage on penetrating the fraud defense. This paper belongs in this third category and focuses on building data views and applied techniques for predictive analytics based on Sparrow's seven levels of fraud control.

4. A Multidimensional Data Model and Analysis Techniques for Fraud Detection

4.1. A Medicaid multidimensional schema

In this section, we present the design of a multidimensional schema that based on Medicaid data will underpin our analysis and allow for the creation of different views of that data that address Sparrow's classification of fraud types. Medicaid providers use four different claim forms to submit claims to the source system: CMS1500 for outpatient professional services, J400 for dental services, UB-04 for institutional claims, and the Drug Claim Form for pharmacy claims. These claims vary slightly in the information collected by purpose, but the general data structure is similar, defining who did what to whom when and why. To maintain the data granularity and specificity, four different claim types: inpatient, long term, pharmacy and professional will be introduced. It should be noted that, as the data is specific to the type of service provided, most commercial insurance claims follow a similar template.

A general core that each claim exists of can be extracted among the different claim forms: patient, provider, diagnoses, procedures, and amounts charged. In our model, the fact table represents a single line from a claim to offer the most flexibility to the user. For each claim-line, a type field links to type-specific detailed information. Based on the desired views from the last section, the following dimensions are included: date (claim filed, service, paid), provider (executing, referring, billing), patient, insurer policy, treatment, diagnosis, claim type, drug, outcome, location. The following numeric facts can be distinguished, some computed by the other facts: Covered charges ($), Non-covered charges ($), Total charges ($), Units of service, Number of days between claim filled and paid, Number of days between service and claim paid, Distance between provider and patient, Number of days between service and claim filled, Covered price per unit, Total price per unit, and Treatment duration. Figure 1 shows the resulting multidimensional schema.

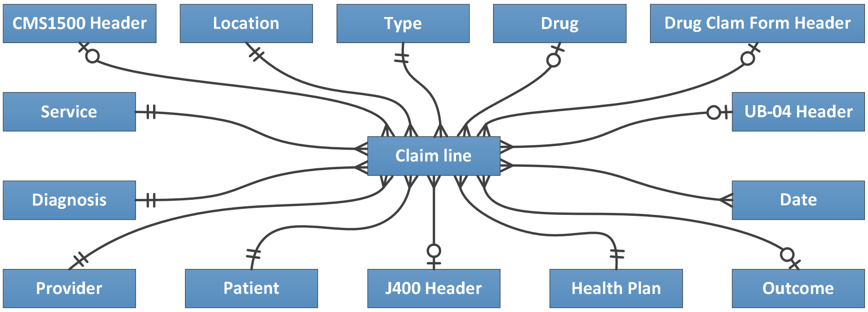

Fig. 1. Medicaid Multidimensional Schema

4.2. Data models addressing levels of fraud

Based on the Medicaid environment and available knowledge base, we developed multidimensional data models representative of Sparrow's fraud classifications and accompanying analysis techniques for detecting the most prevalent fraud types at each level.

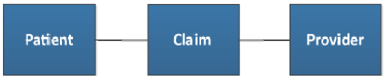

Fig. 2. Level 1 Entities -- Single Claim, or Transaction

Level 1 depicts what today's claims processing systems see: one claim, with its relevant patient and provider. Typically decisions possible at this level are programmed as edits in claims processing system to prevent fraud. Using this level, for example, one can reject duplicate services on a claim and check to see that services are consistent with diagnosis code(s).

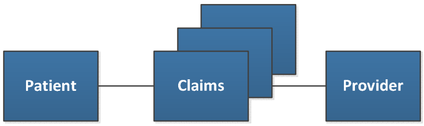

Fig. 3. Level 2 Entities -- Patient / Provider

Level 2 focuses on the relationship between a patient and a provider, including all claims billed. Duplicate billing can be flagged by checking all claims for duplicate provider, patient, and service date. Unbundling could be discovered looking for multiple services from the same provider across claims that should have been grouped. Excessive or unnecessary services could surface when care patterns do not match diagnoses.

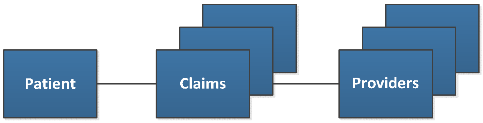

Fig 4. Level 3a Entities -- Patient

Level 3a shows all claims and providers treating a single patient. Phantom billing could be discovered looking at the patient claims vs. prior medical history. How does the patient's temporal claims pattern compare with other patients? One could search for claims that seem unreasonable, such as medically impossible services given known history or services on the same day at two locations far apart. This is the best place to see duplicate billing, checking for all claims for duplicate service performed on the same date across all providers. Upcoding could be discovered looking at claims across providers for consistency. For example, a cardiologist billing for a complicated open heart surgery with an anesthesiologist billing for a simple procedure on the same date of service is suspicious. Unbundling schemes could also surface analyzing multiple providers providing components of a bundled service to one patient. Excessive or unnecessary services across providers can be found by comparing the patient’s service pattern with others with similar diagnosis codes. One can also identify groups of providers through utilization coincidence across patient profiles at this level. These groups are important in detecting more complex fraud schemes in 4b and 6b.

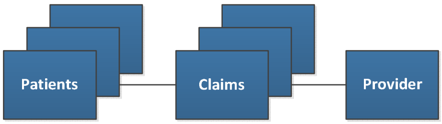

Fig. 5. Level 3b Entities -- Provider

Level 3b exposes the provider. A wealth of knowledge can be gained by simply analyzing the provider’s service distribution and frequency against peers. Clustering analysis of these profiles shows clusters of specialists. Distribution outliers and frequency outliers should be evaluated by medical subject matter experts for legitimacy. Geospatial analysis of patient distance to provider in this model adds additional detail. This is one of the best views to spot phantom billing, upcoding, unbundling, and excessive or unnecessary services.

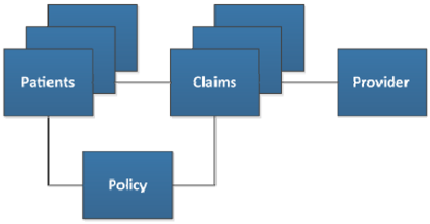

Fig. 6. Level 4a Entities -- Insurer Policy / Provider

Level 4a focuses on analyzing claim pattern differences across different insurance policies or insurers. This could expose providers targeting specific insurers. Patient distributions across insurers are a telling story, as most providers have a somewhat diverse patient base. Providers with high proportions of patients and claims billing specific programs, especially government ones, should be evaluated closely. Phantom billing, upcoding, unbundling, excessive or unnecessary services, and kickbacks could surface with this analysis. Unfortunately, multiple insurer data is rarely available for this analysis.

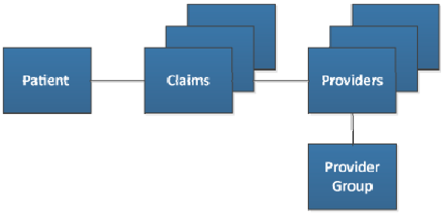

Fig. 7. Level 4b Entities -- Patient / Provider Group

Level 4b looks at all claims for one patient across a known group of providers such as a common clinic. Here, fraud schemes directed within the group and spread amongst providers may stand out. Level 6b is a more effective model for analyzing more complex schemes, and 3a and 3b cover simpler methods.

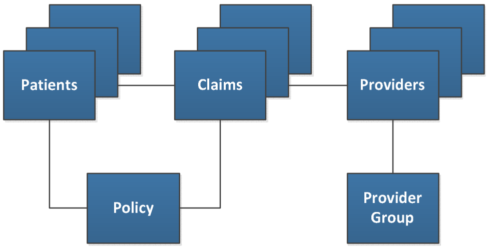

Fig. 8. Level 5 Entities -- Insurer Policy / Provider Group

Level 5 is a combination of level 4a and 4b, showing policy-based variations in a provider group's services. This should be used to evaluate the provider group's insurance billing distribution compared with his/her peers' distribution with a similar patient demographic sampling. Referral patterns may also show policy-specific variations that could be explainable or not. Similar or identical service patterns from the same providers in the group to many patients varying by insurance should be evaluated for reasonableness.

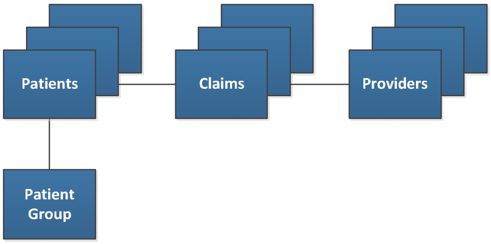

Fig. 9. Level 6a Entities -- Defined Patient Group

Level 6a focuses on patient groups, such as residents of a common nursing home. Here, one would compare claims profiles for patients within the group to similar individuals outside the group. Are the services claimed normal demographically? Are certain providers disproportionately servicing this group with services that would not be commonly needed in the environment? For example, does every patient receive orthotic shoe inserts from a provider? This may seem normal when looking at a DME provider alone in 3b, but overlaying the patient group in 6a connects these patients in a disproportionate way and highlights possible excessive or unnecessary services. Numerous common patient mailing addresses, projected here in a patient group, could point to identity theft where a billing provider has changed the patient's address to phantom bill.

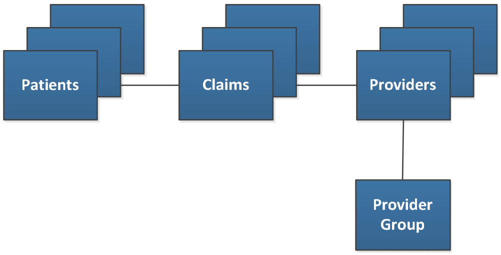

Fig. 10. Level 6b Entities -- Provider Group

Level 6b looks at all claims across a known provider group. Since providers work together and can also bill individually or through clinics or hospitals, this is one of the most useful views of the data. Clustering analysis of the group's service distribution will highlight like groups and identify outlier groups for further analysis. Link analysis of referrals and/or prescriptions along with frequency comparisons with similar provider groups can help detect excessive or unnecessary services that cross provider lines but enrich the group as a whole.

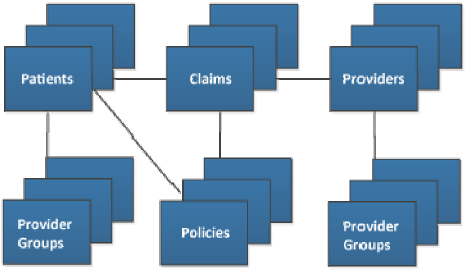

Fig. 11. Level 7 Entities -- Multiparty, Criminal Conspiracies

According to Sparrow [6] the "art of detection at this level involves watching for broad patterns of coincidence or connection between hundreds or thousands of otherwise innocuous transactions". Level 7 combines all previous data views and concerns all fraud that is part of criminal networks which involve many different beneficiaries and/or providers. This much larger data view, spanning billions of claims in the case of Medicaid, is the most rich, delivering the ability to perform complex network analysis that could detect intricate conspiracies. However, performance of analysis here will be much lower than in previous levels. So, it is best for targeted analysis that could not be performed in lower-level views.

4.3. Using the views to detect fraud

As we discussed in each level, it is important for the analyst to know where to look to find specific kinds of fraud. Table 3 maps the six most common types of fraud to the levels that they will most likely be found.

Table 3. Level Usefulness in Detecting Prevalent Fraud Types

		Phantom Billing	Duplicate Billing	Upcoding	Unbundling	Excessive or Unnecessary Services	Kickbacks
Level 1	Single Claim, or Transaction				*	*	
Level 2	Patient / Provider		*		*	*	
Level 3	a. Patient	*	***	*	***	*	
	b. Provider	**		***	*	***	
Level 4	a. Insurer Policy / Provider	**		*	**	**	*
	b. Patient / Provider Group	*	*	*	*	*	
Level 5	Insurer Policy / Provider Group	**		**	**	**	*
Level 6	a. Defined Patient Group	**		*	*	**	**
	b. Provider Group	**		***	**	***	*
Level 7	Multiparty, Criminal Conspiracies	**		**	*	**	***

Usefulness: * Low ** Medium *** High

5. Evaluation

We will evaluate this model and process by subjecting it to a variety of recent healthcare fraud cases. These real-world cases will be mapped against the data models and methods suggested in this paper to test whether the fraud could be detected.

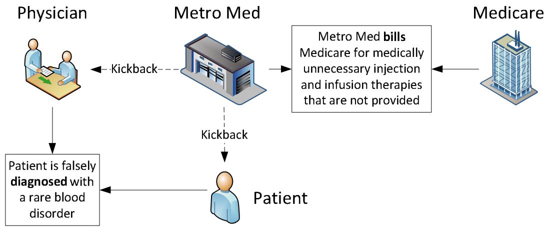
Physician
Metro Med
Medicare
Kickback
Metro Med bills Medicare for medically unnecessary injection and infusion therapies that are not provided
Kickback
Patient is falsely diagnosed with a rare blood disorder
Patient

Fig. 12. Case 1: HIV injection and infusion Medicare fraud scheme

In a case published by the United States Department of Justice [15], a physician, Rene De Los Rios, was convicted of five felony fraud counts for her part in a Medicare fraud scheme, defrauding the government of $23M in collusion with an HIV infusion clinic, Metro Med. In 2003, Metro Med began operating as an HIV infusion clinic that purportedly provided injection and infusion therapies to HIV positive Medicare beneficiaries. In fact, these services were medically unnecessary and not provided. Metro Med paid cash kickback payments to patients for their collusion. De Los Rios was paid to be the licensed physician that would to order tests, sign medical forms and charts (often never seeing a patient), and make it appear that legitimate medical services were being provided. He diagnosed almost all of the patients with the same rare blood disorders to maximize Medicare reimbursements and prescribed expensive medications such as Winrho, Procrit, and Neupogen, to further bill Medicare. Metro Med paid the defendant $3,000 per week for his involvement in the scheme. The Figure 12 visualizes that was happening.

This scheme includes multiple types of fraud, including phantom billing, medically unnecessary services, and kickbacks. To detect this and similar schemes, one could compare physician service profiles with their peers using the level 3b data model. Dr. De Los Rios's excessive diagnosis of a rare blood disorder and/or abnormal, expensive prescriptions would have been an outlier that could have pointed to this problem. Identifying these patients as a suspect patient group, level 6a would show the other providers (the clinic and possibly more) that could be involved in this scheme.

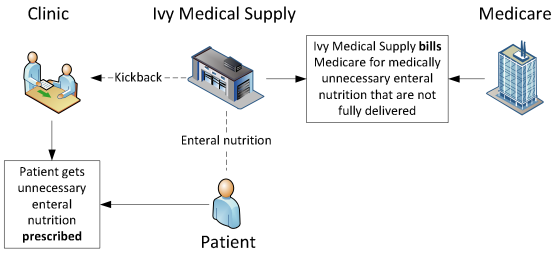
Clinic
Ivy Medical Supply
Medicare
Kickback
Ivy Medical Supply **bills** Medicare for medically unnecessary enteral nutrition that are not fully delivered
Enteral nutrition
Patient gets unnecessary enteral nutrition **prescribed**
Patient

Fig. 13. Case 2: False claims to Medicare for durable medical equipment

In another case [16], California doctors and colluding durable medical equipment (DME) suppliers allegedly submitted over $5M in false claims to Medicare. The defendants prescribed and billed for enteral nutrition, a liquid nutritional supplement provided via feeding tube directly into the stomach, duodenum or jejunum. The doctors, Dr. Augustus Ohemeng and Dr. George Tarryk, wrote fraudulent prescriptions for patients who did not have feeding tubes. George Laing, who managed the clinic where Tarryk and Ohemeng practiced, allegedly received kickbacks in exchange for referring the prescriptions to Ivy Medical Supply. Ivy then fraudulently billed Medicare for the enteral nutrition, even though it was not medically necessary and was not delivered to patients in the quantities billed to Medicare.

This scheme also includes multiple types of fraud, including phantom billing, medically unnecessary services, and kickbacks. Using level 3a, comparing the service pattern of these patients with their peers would highlight the fact that their patients were not previously billed for surgically inserting a feeding tube. This could be explained if the feeding tube was inserted while enrolled in a different health insurance, but the cluster of these patients would stand out nonetheless. Linking these patients together to look at their servicing providers would highlight the actors involved in this scheme. Alternately, abnormally high prescribing patterns for enteral nutrition could be seen in level 3b when comparing the prescribing pattern of the two doctors to their peers. Identifying these patients as a suspect patient group, level 6a would show the other providers (the clinic and possibly more) that could be involved in this scheme.

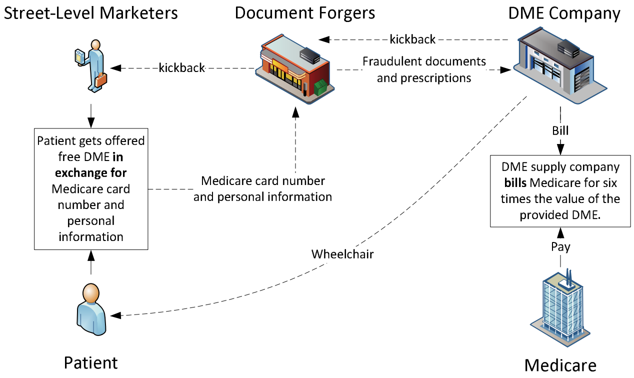
Street-Level Marketers
Document Forgers
DME Company
kickback
kickback
Fraudulent documents and prescriptions
Patient gets offered free DME **in exchange for** Medicare card number and personal information
Medicare card number and personal information
Bill
DME supply company **bills** Medicare for six times the value of the provided DME.
Pay
Wheelchair
Patient
Medicare

Fig. 14. Case 3: Billing Medicare for unnecessary expensive, high-end power wheelchairs and orthotics

A 2012 case involving $14.2M in Medicare fraud [17] showed a DME company purchased both fraudulent prescriptions and patient information to fraudulently bill Medicare for expensive, high-end power wheelchairs and orthotics that were medically unnecessary or never provided. Company owners hid the money they used to pay for these fraudulent prescriptions by writing checks to a third shell company called “Direct Supply.” The checks would then be cashed and the money used to pay kickbacks to street-level marketers offering free power wheelchairs and other DME in exchange for Medicare IDs and personal information. This information would then be used to create fraudulent prescriptions. Depicted in Figure 14, this scheme includes multiple types of fraud, including phantom billing, upcoding, medically unnecessary services, and kickbacks.

This fraud scheme is difficult to detect due to the number of actors and likely multiple (faked) prescribing physicians. However, the model may uncover this fraud over time through various analyses:

- Level 3b: Is the DME provider’s service distribution abnormal compared to peers? Only billing for high-end wheelchairs may stick out, though it could be explained if the company only sold high-end models.
- Level 3b: Is the company showing unnatural growth patterns? Business are built to grow and can do so quickly, but does this agree with the multitude of prescribing physicians? Put another way, a business usually grows quickly by taking on a large customer/referrer, such as becoming the preferred supplier for a hospital vs. tens or hundreds of independent prescribing physicians. This aberrance may show up.
- Level 3a: A third, and likely more telling, indicator could come from analysis of the patient temporal claims patterns compared with other patients. That these patients did not have prior mobility-related claims before should both set them apart from peers. Then analyzing the suspect patient group using level 6a would show the other providers (the clinic and possibly more) that could be involved (the common DME supplier) or compromised (the supposed referring physicians) in this scheme.

6. Conclusion

We structure our design science contribution according to the guidelines of Hevner et al. [3]. The research addresses a relevant and important problem in Medicaid healthcare fraud detection. This paper offers artifacts including a set of multidimensional data models and analysis techniques for healthcare fraud detection along with a projection of these models to a Medicaid-specific schema that would accommodate this analysis. The artifacts are evaluated by discussing their potential in detecting three cases of healthcare fraud. The paper contributes to the literature by mapping the different levels of Sparrow [6] to a set of multidimensional data models and analysis techniques useful at each level for fraud detection. The representation of the artifact used data modeling as a construction method and was evaluated to a list of the most prominent healthcare fraud types. We used the domain context of Medicaid and discussed different design alternatives. We communicated the model to stakeholders in Medicaid, including applying the multidimensional schema in practice.

Through this research, we learned many lessons about antifraud efforts. Significant healthcare subject matter expertise is required to design analysis techniques and interpret their results. Potential entity relationships, medical necessity, and legitimacy in the context of finding the unknown unknowns are extremely difficult to comprehensively model. Our artifacts provide a roadmap that can guide an analyst that evaluates the detected patterns. Lack of training data (marked fraudulent claims) today complicates the application of supervised techniques in the space. It is envisioned that, as models such as these are applied and used by fraud analysts in a structured environment, we can develop training data based on analyst decisions, both around likely fraudulent and likely appropriate claims, further enriching the data and opening doors to supervised techniques. The quest to identify the unknown unknowns in healthcare fraud will never end, but with structured data models, analysis techniques, and continual feedback, we can advance the state of the art in fraud detection and make inroads into an important societal challenge.

References

[1] Kelley RR. Where can $700 Billion in Waste be cut annually from the US Healthcare System? Thomson Reuters 2009; TR-7261 10/09 LW.

[2] Rumsfeld D. Defense.gov News Transcript: Secretary Rumsfeld Press Conference at NATO Headquarters; 2002. Retrieved 21.2.2013, from http://www.defense.gov/transcripts/transcript.aspx?transcriptid=3490

[3] Hevner AR, March ST, Park J, Ram S. Design science in information systems research. MIS Quarterly 2004; 28, 1: 75-105.

[4] Department of Health and Human Services. Medicare A/B Reference Manual - Chapter 21 - Benefit Integrity and Program Safeguard Contractors; 1998. Retrieved 21.2.2013, from https://http://www.novitas-solutions.com/refman/chapter-21.html

[5] Li J, Huang K-Y, Jin J, Shi J. A survey on statistical methods for health care fraud detection. Health Care Management Science 2008; 11, 3: 275-287.

[6] Sparrow MK. License To Steal: How Fraud bleeds America's health care system. Boulder: Westview Press; 2000.

[7] FBI. 2009 Financial Crimes Report; 2009. Retrieved 21.2.2013, from http://www.fbi.gov/stats-services/publications/financial-crimes-report-2009

[8] Albrecht WS, Albrecht CC, Albrecht CO, Zimbelman MF. Fraud Examination. Mason, Ohio: Cengage Learning; 2012.

[9] Travaille P, Müller RM, Thornton D, Hillegersberg J van. Electronic Fraud Detection in the U.S. Medicaid Healthcare Program: Lessons Learned from other Industries. In: Proceedings of the 17th Americas Conference on Information Systems (AMCIS), Detroit, USA, AIS; 2011.

[10] Copeland L, Edberg D, Panorska AK, Wendel J. Applying Business Intelligence Concepts to Medicaid Claim Fraud Detection. Journal of Information Systems Applied Research 2012; 5, 1: 51-61.

[11] Yang W-S, Hwang S-Y. A process-mining framework for the detection of healthcare fraud and abuse. Expert Systems with Applications 2006; 31, 1, 56-68.

[12] Morris L. Combating Fraud In Health Care: An Essential Component Of Any Cost Containment Strategy. Health Affairs 2009; 28, 5, 1351-1356.

[13] Major JA, Riedinger DR. EFD: A Hybrid Knowledge/Statistical-Based System for the Detection of Fraud. The Journal of Risk and Insurance 2002; 69, 3, 309-324.

[14] Ortega PA, Figueroa CJ, Ruz GA. A medical claim fraud/abuse detection system based on data mining: a case study in Chile. DMIN 2006; 6, 26-29.

[15] United States Department of Justice. Miami Doctor Convicted in $23 Million Medicare Fraud Scheme; 2011. Retrieved 21.2.2013, from http://www.justice.gov/opa/pr/2011/April/11-crm-482.html

[16] United States Department of Justice. 8 Los Angeles-Area Residents Charged In Nationwide Medicare Fraud Strike Force Takedown; 2012. Retrieved 21.2.2013, from http://www.justice.gov/usao/cac/Pressroom/2012/055.html

[17] United States Department of Justice. Los Angeles Church Pastor Sentenced to Serve 36 Months in Prison for $14.2 Million Medicare Fraud Scheme; 2012. Retrieved 21.2.2013, from http://www.justice.gov/opa/pr/2012/February/12-crm-256.html