Mining medical specialist billing patterns for health service management

Yin Shan[1], David Jeacocke, D. Wayne Murray, Alison Sutinen
Program Review Division, Medicare Australia
134 Reed St. North, Tuggeranong
ACT 2900, Australia

{Yin.Shan, David.Jeacocke, Wayne.Murray, Alison.Sutinen}@medicareaustralia.gov.au

Abstract

This paper presents an application of association rule mining in compliance in the context of health service management. There are approximately 500 million transactions processed by Medicare Australia each year. Among these transactions, there exist a small proportion of suspicious claims. This study applied association rule mining to examine billing patterns within a particular specialist group to detect these suspicious claims and potential fraudulent individuals. This work identified both positive and negative association rules from specialist billing records. All the rules identified were examined by a subject matter expert, a practicing clinician, to classify them into two groups, those representing compliant patterns and non-compliant patterns. The rules representing compliant patterns were then used to detect potentially fraudulent claims by examining whether claims are consistent with these rules. The individuals whose claims frequently break these rules are identified as potentially high risk. Due to the difficulty of direct assessment on high risk individuals, the relevance of this method to fraud detection is validated by analysis of the individual specialist's compliance history. The results clearly demonstrate that association rule mining is an effective method of identifying suspicious billing patterns by medical specialists and therefore is a valuable tool in fraud detection for health service management.

Keywords: association rule, negative association rule, health data mining, fraud detection, open source data mining.

1 Introduction

There has been an increasing interest in mining health service management data (Becker, Kessler and McClellan, 2005, Lin *et. al.*, 2008, Yang and Hwang, 2006). This is partially due to the fact that public health systems in many countries have consumed a significant portion of governments' expenditure and can be subject to abuse. At the same time, it provides an extremely rich dataset and many challenging research questions, such as detecting fraudulent practice, or inappropriate billing, to facilitate more efficient use of the resources In Australia, a government agency, Medicare Australia, administers Medicare, a fee for service national health funding system for Australians. It is also responsible for undertaking reviews to ensure the integrity of associated health programs administered under Medicare. There have been a series of studies that have applied a range of data mining techniques to the Medicare Australia data for various compliance purposes, such as applying Neural Networks and Boosted Regression Trees to detect fraudulent behaviour by general practitioners (Pearson, Murray and Mettenmeyer 2005), K-nearest neighbour method for fraud detection (He, Graco and Yao 1999) and positive association rule mining to better understand medical practice patterns (Semenova, 2004).

Although there have been an increasing number of applications of association rule mining, they usually focus on positive rules and discovering common patterns. There is very limited research on association rule mining in detecting anomalous patterns for compliance purposes in the medical service domain. This study applied association rule mining for fraud detection in a specialist population. In addition to conventional positive rule mining, negative association rule mining was also applied in this study. We were able to show negative association mining rules to be particularly useful in this application and possibly other fraud detection problems.

Rules describe typical patterns of practice, which may reflect compliant or non-complaint patterns. All rules must therefore be evaluated by a subject matter expert to determine their relevance. Rules representing non-compliance may imply some commonly entrenched incorrect specialist billing practices. This is very useful in improving compliance, in particular, it may assist with clarifying billing regulations, for example through targeted educational interventions.

Rules representing compliant practices are valuable to identify specialists who may not be practicing in accordance with their peers and may therefore be billing either inappropriately or fraudulently. Individual specialists, whose claims frequently break these rules, may present a high risk of inappropriate or fraudulent billing patterns. It would be impossible to definitively validate whether specialists identified by association rules were engaged in inappropriate practices, without a comprehensive review of their billing practices being undertaken by a panel of peers. Thus an indirect approach to validation was undertaken in this study. The effectiveness of association rule mining to detect fraud or inappropriate practice was evaluated by analysing specialists' compliance history. We were able to show that with this effectiveness measure, association rule mining was a valid and promising method of identifying potentially fraudulent billing patterns.

The remaining sections of this paper are organised as follows. The problem domain is briefly introduced in Section 2. Section 3 provides a short description of positive and negative association rule mining. Experiments are reported in Section 4. The evaluations of the results numerically and by subject matter experts are covered in Section 5. Section 6 presents the conclusions and future research.

2 Specialist Billing Patterns

Specialists claim a significant portion of Medicare benefits. There are dozens of small specialist groups based on their specialties, consisting of tens to hundreds of specialists each. In contrast General Practitioners (GPs) as one professional group are much larger in number (over 25,000 nation-wide) and exhibit relatively much less variation in their practice patterns. Because of their unique practice styles and small group sizes, specialist groups impose interesting challenges to automatic fraud detection approaches.

One of the main compliance tasks in specialist groups is to ensure specialists bill items according to the intent specified under the Medicare system. Ideally, billing patterns may be identified as anomalous through a clear difference in the pattern of services rendered by other specialists in the same specialty group. The discovery of anomalous billing patterns may identify a range of issues from fraud, inappropriate billing to billing arising from new technologies and procedures. In each case, the billing pattern discovered may assist Medicare Australia to determine the effective advice to policy makers and compliance response to ensure the integrity of its programs. If these patterns can be identified in time and correctly, the response can be made to benefit both Medicare in reducing inappropriate payments and sometimes to the profession to determine areas of the Medicare Benefits Schedule (MBS) (Australian Government, 2007) requiring clarification or new Medicare items.

3 Association Rules

Association rule mining (Agrawal, Imielinski, and Swami, 1993, Agrawal and Srikant, 1994) has drawn a lot of attention because of its effectiveness and intuitive representation. This has resulted in many efficient algorithms being proposed. For completeness in this paper, we only present a brief description. More details can be found in other literature sources (Agrawal, Imielinski, and Swami, 1993).

Let $I=\{i_1,i_2,\ldots,i_m\}$ be a set of literals, called items. Let D be a set of transactions, where each transaction T is a set of items such that $T \subseteq I$. A transaction T contains X if $X \subseteq T$. An association rule is in the form $X \rightarrow Y$, where $X \subset I$, $Y \subset I$, and $X \cap Y = \varnothing$. The rule $X \rightarrow Y$ holds in the transaction set D with confidence c if c% of transactions in D that contain X also contain Y. The rule $X \rightarrow Y$ has support s in the transaction set D if s% of transactions in D contain $X \cup Y$.

The association rule defined above, describing the presence of items, can't completely meet our needs. In the compliance domains, it is natural to ask complementary questions. For example, one common non-compliant billing practice is to bill some additional items in conjunction with commonly billed items for the same service. If a rule can tell us what should *not* be billed with other commonly billed items, such non-compliant "add-on" billing can be easily identified. This type of association rules which can describe the absence of items are called negative association rules (Savasere, Omiecinski and Navathe 1998, Zhang and Zhang 2002).

To avoid confusion, the previously defined association rules are called positive association rules. Negative association rules are in the form $X \rightarrow \neg Y$, which can be interpreted as that if X is present, it is unlikely that Y would be present too. Because the formal full definition of negative association rule is similar to that of the positive ones, it is omitted here. Negative association rules also have similar measures of confidence and support as in for the positive rules and therefore they aren't reiterated here.

4 Experiments

The data set used in this study was drawn from Medicare Australia's Enterprise Data Warehouse, covering billing records of a specialist group for the second quarter of 2007 (1 April, 2007 – 30 June, 2007 inclusive). The data was organised in episodes which were defined as all the items claimed or billed for one patient on one day by one specialist. Obviously, an episode corresponds to a transaction in the context of association rule mining. This data set contained 63010 episodes (transactions). Removing those episodes which contained only one item, resulted in 32476 episodes deemed suitable for further study.

Because of the nature of the specialist practice examined, there were as many as 620 procedural MBS items presented in this dataset. This results in as many as 7989 unique billing episodes because of the distinctive needs of each individual patient. After conducting a series of empirical studies, we determined that the settings of 80% confidence and 0.1% support produced optimal results.

In total, 215 association rules, including both positive and negative rules, were identified. These rules were presented to the subject matter expert for evaluation and specialists were checked against these rules to study the effectiveness of the method in identifying potential non-compliant individuals.

The association rule discovery was undertaken using the Christian Borgelt, Artamonova and others' open source implementation (Borgelt and Kruse 2002, Artamonova, Frishman and Frishman, 2007) of Apriori (Agrawal, Imielinski and Swami 1993)

5 Evaluation

There were three components to the evaluation undertaken. The first involved examination of each of the rules identified by the subject matter expert, to determine whether the services reflected by the items in one rule might conceivably be billed together. Thus, all rules were examined to assess their clinical relevance. While only a single subject matter expert was utilised in this preliminary study, consideration was given to whether the items could be appropriately billed under Medicare rather than whether they considered the management approach was consistent with their own practice.

The second aspect of the evaluation involved the comparison of an individual specialist against rules representing compliant patterns, to determine whether billing patterns of the specialists observed were likely to reflect non-compliant billing patterns and the number of occasions on which this occurred was recorded. The specialists who broke rules on a great number of occasions were identified as high risk.

Finally the compliance histories of the specialists identified as high risk were examined. This provides us with a good indication of the effectiveness of association rules in identifying inappropriate or fraudulent claims patterns.

5.1 Rule evaluation

In total, 215 rules were identified, including 192 negative rules and 23 positive rules. It was not surprising that more negative rules were discovered because for positive rule discovery only the presence of items are considered while for the negative rule discovery both presence and the absence of the items are considered. Another observation was that although there were over 20 positive rules identified, there were only a very small number of unique items involved.

The negative rules were much stronger than positive rules in terms of confidence. The minimum confidence of negative rules was 95.95% while it was only 80.25% for positive rules. This fits with the clinical context, based on the description of items in the Medicare Benefit Schedule, where some services were explicitly stated that they should not be billed together on the same day. So the negative rules consistent with these patterns should be very strong. Another example this study highlighted, was that there were several very strong negative rules, indicating procedural items should not be billed on the same day with an initial attendance. Clinical observation tells us it is extremely unlikely that a specialist would perform procedures on the very first consultation with the patient. Thus, it is not surprising these negative rules describing these clinic scenarios are very strong.

On the other hand, the clinic scenarios indicated by positive rules are not as definitive as those by negative rules. For positive rules, close examination reveals that these rules themselves, not violations of them, can represent inappropriate billing. However, it is also possible that these rules may reflect new patterns of billing by specialists, possibly related to a new specialist technology or technique, resulting in a small number of specialists starting billing differently from their peers.

5.1.1 Common compliant patterns

Negative rules represented common patterns that were generally considered consistent with the billing rules covered under the Medicare Benefit Schedule. It is very encouraging that some of negative rules correspond very well to some unusual combinations in this specialist group, which have been alerted to Medicare Australia's clinical experts by other sources.

The subject matter experts found it was more intuitive to interpret the negative rules and the implication of their violation than those of positive rules. The subject matter expert concluded that violations of these rules are likely to be good indications of non-compliant billing. Violations of these negative rules suggest billing additional items not normally billed by the majority of specialists. Those specialists who violate these rules frequently are thus markedly different from their peers. Therefore, concerns may need to be raised regarding the appropriateness of the services provided by these specialists.

Of the 192 negative rules identified, 30 rules had a confidence value of 1.0, which was considered neither numerically interesting, nor in the opinion of the subject matter expert, as being useful for compliance purposes. These rules were thus removed. For the remaining 162 negative rules, the subject matter expert classified them into three groups based on the likelihood inappropriate billing (see Table 1). High rating indicates the rules make strong sense in the domain. If any of these rules was broken, it was almost certain to suggest an incorrect billing to Medicare Australia. The low rating indicates that although breaking these rules might be inappropriate other appropriate billing explanations may also exist. In other words, low rating rules might not be sufficient for detecting inappropriate billing. It is worth mentioning although these low rating rules may not be sufficient for direct identification of inappropriate billing, they still provide valuable information on profiling specialists for related compliance activities. As can been seen from Table 1, more than half of the rules (56.18%) discovered, comprising high and medium rate rules, are regarded as suitable for detecting inappropriate billing.

Rating	Number of Rules
High	53 (32.72%)
Medium	38 (23.46%)
Low	71 (43.83%)

Table 1: The risk rating of the negative rules

5.1.2 Common non-compliant patterns

An unusual finding relating to the positive association rules, was that all of the positive rules were unexpected to some extent, i.e., positive rules can not be fully explained by the subject matter expert. As mentioned, there may be several possibilities to explain the occurrence of these frequent patterns. One possibility is that these rules indicate inappropriate billing practices. It is also possible that these rules reflected new billing patterns by specialists where the service rendered may not yet be reflected within the MBS billing structure. The third possibility is that these rules describe incorrect billing due to uncertainty or misunderstanding among specialists in relation to the correct billing method, other than deliberately taking advantage of MBS benefit.

Although there are 23 positive rules, these related to mainly two sets of items. Further research is proposed to determine the nature of these two sets of items so as to better understand the clinical context. Once confirmed by further research, these positive rules would be very valuable in assisting Medicare's educational intervention or government policy responses.

5.2 Relevance to compliance

To identify the specialists with anomalous, potentially fraudulent, behaviour, the rules were matched against all the episodes each specialist rendered. This allowed the total number of occasions where rules were broken by each specialist to be identified. The number of rule violations provided an indication of how much one specialist deviate from their peers.

As listed in Table 1, 162 negative rules are rated from low to high by the subject matter expert. Rules rated medium or high may be directly related to non-compliant practices. Therefore, all the specialists were checked against these high or medium rating rules. The specialists who broke these rules on the greatest number of occasions were identified as high risk. The best way of validate whether the individual specialists identified by this method were truly non-compliant, would be a review by a panel of peers or investigation possibly followed by legal action, which is prohibitively costly and time consuming. Luckily, there is a database available, called PRISM maintained by Medicare Australia, containing records of medical practitioners who have been approached in relation to previous compliance activities. Therefore, an alternative performance validation method is to match specialists identified by association rule mining against their compliance history in PRISM. This provides a reasonable estimate of the effectiveness of association rule mining in detecting non-compliant practice. Validation was against records within PRISM not necessarily linked to outcomes, however it is known that the majority of these records relate to specialists for compliance related issues.

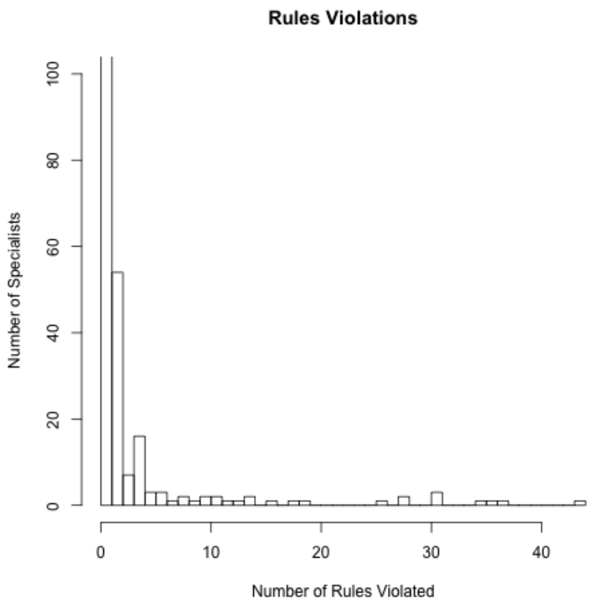

Figure 1: Number of Rule Violations

There were 779 specialists included in this study, based upon their derived specialty. Among them, there are only 129 specialists breaking rules on one or occasions. As can be seen in Figure 1, for those specialists who broke rules, most of them only broke rules on one occasion. The highest number of occasions of breaking rules is 44. The total number specialists who did not break any rules was 650 (83.44% of all specialists identified), which greatly exceeds the upper limit of y-axis in Figure 1.

As demonstrated in Figure 1, amongst those specialists who broke one or more rules, the vast majority of them break rules on less than five occasions and only an extremely small number of specialists broke the rules on more than 20 occasions. Therefore for further analysis, the specialists were divided into three overlapping groups based upon the number of occasions they had broken rules. These three classes were that specialists who broke rules in:

1) 1 or more occasions;
2) 5 or more occasions;
3) more than 20 occasions.

The accuracies of the association rules in detecting likely inappropriate billing are listed in Table 2. There were 10 specialists who broke these rules on more than 20 occasions. Amongst these, 5 had compliance records in PRISM resulting in an estimated accuracy of 50%. For specialists who broke more than 5 rules, the estimated accuracy was 25.81%, compared to an accuracy of 29.46% for those specialists who broke more than 1 rule. In order to put these accuracies into perspective, we constructed a baseline classifier, which randomly samples the data. As 163 of the 779 specialists have more than one compliance record, this results in an accuracy of 20.92%. Therefore, it is clear that the association rules mining method utilised in this study outperforms the random sampler. The fact that breaking even one or more rule can give us an accuracy of 29.46%, better than the baseline classifier (20.92%), suggests that breaking even one negative rule may be a good indication of non-compliant practice.

The compliance data base contains information about practitioners engaged in possible fraud or inappropriate practices. Often practitioners may be identified as having concerns in multiple areas. It was not possible to ensure that all practitioners from the PRISM data base were selected for compliance related issues. Some cases may have reflected past targeting strategies. This may have resulted in misclassification of practitioners. Provided, as may be assumed likely, this misclassification was non-differential it might be expected that the overall accuracy levels related to the number of rules violated would be higher, as this would have a dilutive effect.

No previous compliance activities have been undertaken in relation to the newly identified rules from this analysis.

Rules Violated	Specialists Identified	Specialists with compliance records	Accuracy
≥20	10	5	50.00%
≥5	31	8	25.81%
≥1	129	38	29.46%
Baseline	779	163	20.92%

Table 2: Accuracy of association rule in detecting potentially inappropriate practises, as measured by the percentage of specialists with past compliance records.

We are aware that specialists may have different numbers of records in their compliance history, which suggests some specialists have multiple incidents of non-compliant practice or have been engaged in multiple compliance activities. To measure the severity of a specialist's possible non-compliant practice, we calculated the average number compliance records listed in Table 3. For all the specialists who had compliance records, on average they had 1.47 records per specialist. For the three classes of specialists identified above, they have on average 1.53, 1.63 and 1.80 records per specialist, repectively. It is clear the there is a close relationship between the number of occasions where rules were broken and severity of non-compliance of a specialist, measured by average number of records. In combination, these findings suggest that association rule mining can not only identify potential non-compliant specialists but also give us a good indication of the severity of their non-compliant behaviour.

Rules Violated	Specialists with compliance records	Average No. of records
≥20	5	1.80
≥5	8	1.63
≥1	38	1.53
All records	163	1.47

Table 3: Average number of compliance records per specialist.

We also checked the specialists against all the negative rules, not just the high and medium rating ones. It was unexpected that this gave similar accuracy to only checking against high and medium rating negative rules. This probably suggests that any broken rules may flag possible fraudulent activities. However, we will not be able to explore this further in the paper.

5.3 Negative rules vs. Positive rules

It was reported by the subject matter expert that negative rules may have certain advantages in compliance. Negative rules represent the absence of items being billed. In the compliance context, it may often be the case that more items than necessary are billed for financial gain. Such billing patterns are well captured by negative rules.

6 Conclusions and future research

This paper presents a novel application of association rule mining and demonstrates how both positive and negative association rule mining can be used with the aims of detecting fraud and inappropriate practice in the health service management domain.

The results were validated in several ways. The subject matter experts have confirmed the clinical relevance of the rules discovered. The individual specialists identified have good overlap with specialists who have compliance records. This demonstrates the effectiveness of this method for fraud detection and compliance. For further validation, this method is compared to the baseline classifier. This method significantly outperformed the baseline classifier. It is worth mentioning we have also demonstrated that this method may give a good indication of severity of the potential non-compliant activity as well.

This research clearly demonstrates that methods used were effective and we see immediate potential in the use of these methods to identify other relevant specialist groups to support compliance activities within Medicare Australia. Medicare Australia has run this technique against a limited number of medical specialties at this point and further validation will be undertaken based on compliance intervention feedback and future specialty group analyses. It is envisaged that this technique might be applied to a broad range of specialty and practitioner groups covered the Medicare system.

Further work is proposed as follows to enhance the use and evaluation association rules mining in compliance in the field of health service management.

A more comprehensive validation may be conducted. Information from other sources regarding the high risk specialists identified can be collected and limited cases might be audited to provide a more comprehensive assessment on the effectiveness and the accuracy of the association rule mining in detecting fraud and inappropriate billings.

It is possible that some of the rules identified by this type of analysis maybe false alarms, reflecting appropriate though infrequent practice. For this reason it is important that the outcomes of this form of analysis are further reviewed by a subject matter expert.

This technique is a substantial improvement over random auditing of practitioners engaged in specialised areas of practice.

This research focuses on the analysis of episodes, which only considers all items rendered in one given day. Therefore, there is no particular order recorded among items involved in one episode. However, the chronological order of the items may be crucial in determining the appropriateness of the billing. Therefore, it may be promising if the episode defined in the paper is expanded to "total episodes of care", covering 28 days, where the chronological information is recorded. With this added time dimension, this would be an interesting challenge for temporal data mining and may help further enhance the detection of non-compliant billings by Medicare.

7 Acknowledgements

The authors would like to thank Christian Borgelt and Irena Artamonova for their assistance in providing their implementation of association rule mining packages. The authors also wish to acknowledge the input of the three anonymous referees whose insightful comments improved the final version of this paper.

8 References

Agrawal, R., Imielinski, T. and Swami, A. (1993): Mining association rules between sets of items in large databases. *Proc. ACM SIGMOD International Conference on Management of Data*, Washington DC, USA, **22**:207-216, ACM Press.

Agrawal, R, and Srikant, R. (1994) Fast algorithms for Mining Association Rules. *Proceedings of the International Conference on Very Large Data Bases*, Satiago, Chile, 487 – 499.

Artamonova, I., Frishman. G., Frishman D. (2007) Applying negative rule mining to improve genome annotation. *BMC Bioinformatics*. (Jul 21);**8** (1):261.

Becker, D. and Kessler, D. and McClellan, M. (2005) Detecting Medicare abuse. *Journal of Health Economics*. **24**(1): 189-210.

Borgelt, C. and Kruse, R. (2002). Induction of Association Rules: Apriori Implementation. 15th Conference on Computational Statistics (Compstat 2002, Berlin, Germany) Physica Verlag, Heidelberg, Germany. http://www.borgelt.net/apriori.html

He H., Graco, W. and Yao X. (1998) Application of Genetic Algorithm and k-Nearest Neighbour Method in Medical Fraud Detection. In Second Asia-Pacific Conference on Simulated Evolution and Learning (SEAL '98), Canberra, Australia. pp. 74-81, LNAI 1585 Springer

Lin, C., Lin, C-M., Li, S-T. and Kuo, S-C. (2008) Intelligent physician segmentation and management based on KDD approach. (2008) *Expert Systems with Applications*. **34**(3): 1963—1973. Pergamon Press, Inc. Tarrytown, NY, USA

Medicare Benefit Schedule Book (2007) Australian Government, ISBN 1-74186-363-5, Department of Health and Ageing, Australian Government. Canberra. ISBN 1-74186-363-5

Pearson, R., Murray, W. and Mettenmeyer, T. (2006) Finding Anomalies in Medicare. *electronic Journal of Health Informatics*. **1**(1): e2. (www.ejh.net/ojs/index.php/ejhi/issue/view/1)

Savasere A., Omiecinski E., and Navathe S. B. (1998) Mining for Strong Negative Associations in a Large Database of Customer Transactions. *Proceedings of the International Conference on Data Engineering*, February

Semenova T. (2004) Discovering patterns of medical practice in large administrative health databases. *Data Knowledge Engineering*. **51**(2): 149—160, Elsevier Science Publishers B. V. ISSN 0169-023X Amsterdam, The Netherlands

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

Zhang, C., Zhang, S. (2002) Association Rule Mining: Models and Algorithms. *Lecture Notes in Computer Science*, **2307**. ISBN: 978-3-540-43533-4