

A Probabilistic Model for Reducing Medication Errors

Phung Anh Nguyen^{1,2}, Shabbir Syed-Abdul^{1,2}, Usman Iqbal², Min-Huei Hsu⁴, Chen-Ling Huang⁶, Hsien-Chang Li³, Daniel Livius Cliniciu⁵, Wen-Shan Jian^{3*}, Yu-Chuan Jack Li^{2,7*}

1 Institute of Biomedical Informatics, National Yang-Ming University, Taipei, Taiwan, **2** College of Medicine Science and Technology, Graduate Institute of Biomedical Informatics, Taipei Medical University, Taipei, Taiwan, **3** School of Health Care Administration, Taipei Medical University, Taipei, Taiwan, **4** Bureau of International Cooperation, Department of Health, Taipei, Taiwan, **5** Institute of Translational Medicine, Taipei Medical University, Taipei, Taiwan, **6** Department of Internal Medicine, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan, **7** Department of Dermatology, Taipei Medical University - Wan Fang Hospital, Taipei, Taiwan

Abstract

Background: Medication errors are common, life threatening, costly but preventable. Information technology and automated systems are highly efficient for preventing medication errors and therefore widely employed in hospital settings. The aim of this study was to construct a probabilistic model that can reduce medication errors by identifying uncommon or rare associations between medications and diseases.

Methods and Finding(s): Association rules of mining techniques are utilized for 103.5 million prescriptions from Taiwan's National Health Insurance database. The dataset included 204.5 million diagnoses with ICD9-CM codes and 347.7 million medications by using ATC codes. Disease-Medication (DM) and Medication-Medication (MM) associations were computed by their co-occurrence and associations' strength were measured by the interestingness or lift values which were being referred as Q values. The DMQs and MMQs were used to develop the AOP model to predict the appropriateness of a given prescription. Validation of this model was done by comparing the results of evaluation performed by the AOP model and verified by human experts. The results showed 96% accuracy for appropriate and 45% accuracy for inappropriate prescriptions, with a sensitivity and specificity of 75.9% and 89.5%, respectively.

Conclusions: We successfully developed the AOP model as an efficient tool for automatic identification of uncommon or rare associations between disease-medication and medication-medication in prescriptions. The AOP model helps to reduce medication errors by alerting physicians, improving the patients' safety and the overall quality of care.

Citation: Nguyen PA, Syed-Abdul S, Iqbal U, Hsu M-H, Huang C-L, et al. (2013) A Probabilistic Model for Reducing Medication Errors. PLoS ONE 8(12): e82401. doi:10.1371/journal.pone.0082401

Editor: Francesco Pappalardo, University of Catania, Italy

Received: May 23, 2013; **Accepted:** October 24, 2013; **Published:** December 3, 2013

Copyright: © 2013 Nguyen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This research was funded by Taipei Medical University Hospital and Shuang-Ho Hospital, Taipei, Taiwan, as a part of "A Randomized Control Trial to improve patient's safety by reducing medication errors in computerized physician order entry (CPOE) systems" project, no. 101TMU-SHH-21. The data provided by the Bureau of National Health Insurance, Department of Health, Taiwan. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: jttmutw@gmail.com (WSJ); jaak88@gmail.com (YCJL)

Introduction

Medications are one of the most powerful tools in modern medicine used for the treatment of diseases. Unfortunately, sometimes instead of providing treatment, they can cause considerable harm and even death, especially if prescribing physicians fail to consider relevant patient data and characteristics [1–3]. Barach *et al.* reported that nearly 100,000 individuals die per year in the United States due to preventable medical errors, most of which are medication errors [4,5]. Studies on medication errors during drug administration in surgical units [3,6], the incidence of medication errors in intensive care units [7,8], and in pediatric units [2,9] reveal that in most cases, preventable medication errors cross barriers and reach to patients [2,7,9].

Reducing medication errors to increase patient's safety is crucial to evaluating hospital performance and improving patient outcomes. Wyatt J.C. *et al.* reveals that information technology (IT) boosts clinical leadership for development and procurement in healthcare [10,11] by improving reliability, quality, medication

safety [12,13], and most importantly reducing prescription errors [3,5,13–20]. Information Technology also saves hundreds of billions in annual costs by providing automation for ordering, a key process in modern health care. Such processes include CPOE [15] with clinical decision support (CDS), bar-coded medication administration [21], automated dispensing systems (ADS) [8], and dose drug distribution [22]. Bates D.W. *et al.* claims that the key tools for reducing this gap would be information systems which provide decision support to users during decision making, resulting in improved quality of care [23]. Thus, the knowledge-based CDS review can assure that those orders are safe and comply with guidelines [20,23].

Most knowledge-based systems were implemented for automated methods, statistically developed by experts at a significant cost to maintain assurance and evidence [24–26]. In our study, in order to improve the efficiency of detecting medication errors, we used a set of data mining techniques such as frequent item set mining and association rule mining [27]. Being used in a variety of other fields [28–30], data mining techniques are also successfully

employed in medical informatics applications such as screening tests for preventive care decision support, to locate potentially unknown adverse effects of drugs [31], as well as identifying disease-drug associations in biomedical literature and clinical texts [32]. Wright A. *et al.* used association rule mining to identify clinically accurate associations between medications, laboratory problems [33], and other related studies [34–38].

The aim of this study focuses on medication-disease relationships by applying the association rule mining, and using statistical methods to detect medication errors in the computerized physician order entry (CPOE) systems in order to improve patient’s safety.

Methods

In this study, we developed a model to detect uncommon or rare medication for a given disease when ordering prescriptions based on disease-medication associations. The steps involved in methods are described below (see Figure 1):

- (I). Gathering and cleaning data from the Taiwan National Health Insurance (NHI) database.
- (II). Quantifying Disease-Medication (DM), Medication-Medication (MM) co-occurrences and computing associations’ strength (Q values) of all the associations present in 103,485,067 prescriptions.
- (III). Developing a model that can automatically evaluate the **Appropriateness of a Prescription (AOP)**.
- (IV). Testing the model for 100,000 randomly selected prescriptions, and then validating the results using seven human experts.

Phase (I): Data Gathering and Cleaning

The Taiwan’s NHI claims data from Jan 1, 2002 to Dec 31, 2002 with a total of 263.6 million prescriptions from outpatient visits was employed. Each prescription consists of one to three diagnostic codes and one to fifteen medication codes. We excluded 160.1 million prescriptions due to the following reasons: a) missing or invalid disease codes or medication codes, and b) the use of traditional Chinese medicine prescriptions. Thus, the remaining 103.5 million prescriptions with 204.5 million diagnosis ICD9-CM (International Classification of Disease v.9-Clinical Modification) codes and 347.7 million medications with the Taiwan NHI codes were used in the analysis. These medication codes were mapped to the ATC (Anatomical Therapeutic Chemical) classification code system resulting in 13,070 unique ICD9-CM codes and 1,548 unique ATC codes.

Phase (II): Quantifying Disease-medication (DM), Medication-medication (MM) Co-occurrences and Computing Associations’ Strength (Q Values)

The co-occurrences of disease-medication and medication-medication association were used in our analysis. The data mining techniques used were: 1) Frequent (large) item-set mining and 2) Association rule mining; both are closely related and complementary [33].

- 1. **Frequent item-set mining** is a technique for locating common items’ co-occurrence in a transaction database [39] and determining possible associations among them.
- 2. **Association rule mining** is an extension of frequent item-set mining [28,40], which directs the association between two items in addition to a simple co-occurrence.

Frequent item-set and association rule mining are practical techniques for inferring relationships between disease and medication; thus, many potential rules are often produced through these techniques for filtering, such as a variety of measures for “interestingness” [41,42]. We termed interestingness or lift value as Q value.

Q value is the ratio between the joint probability of disease-medication and medication-medication with respect to their expected probability under the independent assumption known as lift (interest) and relative risk (RR) in similar studies dealing with associations [33,41,43].

Furthermore, the disease and medication are considered to be co-occurring if they appeared in the same prescription (see Appendix S1). Based on this definition, each DM pair and MM pair association’s strength was computed using a 2x2 table (Figure 2). Figure 2 shows the equation used to compute the Q value in this study. Q is defined as [0, +∞]; Q=1 indicates no association between disease and medication, Q <1 indicates that disease and medication are negatively associated (i.e. negative DMQs), and Q >1 indicates that disease and medication are positively associated (i.e. positive DMQs - the prescriptions with disease X containing medication Y occur more often than other medications).

A total of 1.34 million DM and 0.65 million MM pairs with their Q values were computed from 103.5 million prescriptions. The DM and MM associations with less than 5 co-occurrences, by default were considered as “uncommon or rare associations” and were not included in developing the model.

Phase (III): Developing the Appropriateness of a Prescription (AOP) Model

In this phase, the values of DMQs and MMQs computed in phase (II) are used in developing the AOP model. The Appropriateness of a Prescription (AOP) model is developed based upon following rules:

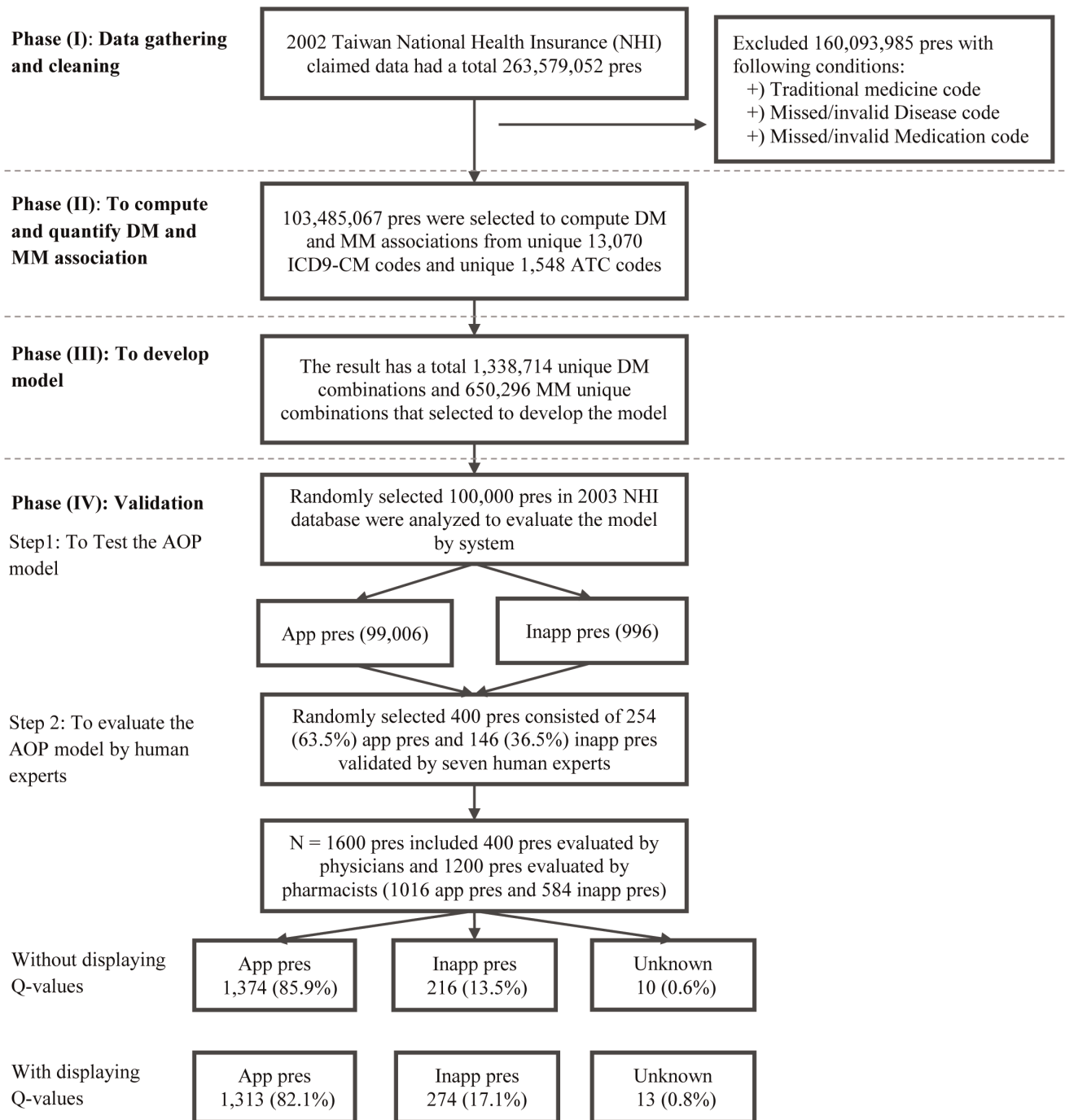
- 1. The number of positive DMQs and positive MMQs should be greater than or equal to the number of medications.
- 2. All diagnoses should have at least one positive DMQ.
- 3. Each medication should have at least one positive DMQ or positive MMQ.

The AOP model is expressed mathematically as:

$$\left\{ \begin{array}{l} \{Q_{D_i M_j}\} + \{Q_{M_j M_k}\} \geq m \\ \exists Q_{D_i M_j} : \{Q_{D_i M_j}\} \geq 1 \\ \forall M : \left[\begin{array}{l} \exists Q_{D_i M_j} : \{Q_{D_i M_j}\} \geq 1 \\ \exists Q_{M_j M_k} : \{Q_{M_j M_k}\} \geq 1 \end{array} \right. \end{array} \right. \quad (i \in \{1, 2, \dots, n\}; j, k \in \{1, 2, \dots, m\})$$

where n is the number of diagnoses; m is the number of medications; Q_{D_iM_j} is a DMQ; Q_{M_jM_k} is a MMQ in the same prescription.

The AOP model was built to evaluate associations (DM and MM) present in each prescription and to make a decision whether the overall prescription is appropriate or not. In other words, the model will consider a prescription as appropriate if and only if it has at least one positive DM association present.

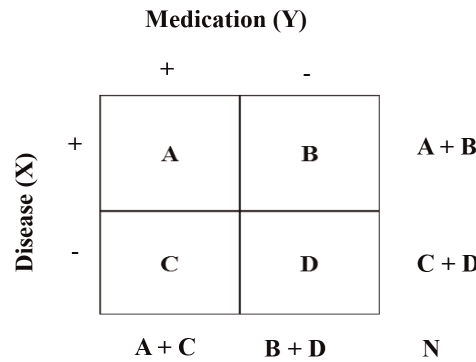


Abbreviation: App pres: appropriate prescriptions; Inapp pres: inappropriate prescriptions; Pres: prescriptions; AOP: appropriate of prescription

Figure 1. Overall of study design.
 doi:10.1371/journal.pone.0082401.g001

Phase (IV): To Test and Evaluate the AOP Model
 Two steps are involved in this phase. First, to test the model based on the verifying dataset and subsequently, evaluating the

results by the human experts including four physicians and three clinical pharmacists to measure the accuracy of the AOP model.
Step 1: To test AOP model. In the first step, 100,000 prescriptions were randomly selected from the 2003 NHI claims



Formula:

$$Q_{XY} = \frac{P(X,Y)}{P(X)*P(Y)} = \frac{P(X|Y)}{P(X)} = \frac{\frac{A}{N}}{\frac{(A+B) * (A+C)}{N}}$$

Note: Generate DM association (MM association) and compute Q in contingency table

Figure 2. Formulation of Q used in this study.

doi:10.1371/journal.pone.0082401.g002

database. Then the AOP model was used to test the selected prescriptions for appropriateness.

Step 2: The evaluation of AOP model by human experts. In the second step, 400 prescriptions were randomly selected out of the 100,000 prescriptions and were tested by our AOP model. The 400 prescriptions selected to be evaluated by experts contained 254 (63.5%) appropriate and 146 (36.5%) inappropriate prescriptions. All experts were explained the purpose of the study and were asked to mark whether they agree, disagree or are unsure regarding the overall prescription data provided to them (see Appendix S2). Next, the same prescriptions were re-evaluated with and without the Q values for each DM association present in the prescriptions. We administered the questionnaires (see Appendix S3) to four physicians at their clinic sites (two hundred prescriptions per each physician) and to three clinical pharmacists (eight hundred prescriptions per each pharmacist) at the hospital pharmacies. Overall, we administered 3,200 prescriptions (1,600 prescriptions without Q values and 1,600 with Q values). The average time spent to fill out both questionnaires was about 45 minutes by each physician, and 150 minutes by each pharmacist.

The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were computed from the results

obtained in order to compare the differences and the consensus between the system and the experts.

Results

Results of Step 1

From a total 100,000 prescriptions 99,004 prescriptions (99.004%) were evaluated as appropriate and 996 prescriptions (0.996%) were evaluated as inappropriate by the AOP model.

Results of Step 2

When the Q values were not disclosed the experts responded to 1590 (99.3%) prescriptions of which 1,374 (85.9%) were appropriate and 216 (13.5%) were inappropriate prescriptions, leaving 10 prescriptions as “unknown”. However, when the Q values were shown in the prescriptions the experts responded only to 1,587 (99.2%) prescriptions of which 1,313 (82.1%) were appropriate, 274 (17.1%) were inappropriate prescriptions, and 13 prescriptions were classified as “unknown” (see Figure 1).

The AOP model results were verified by the experts and the sensitivity, specificity, positive (PPV), and negative predictive values (NPV) were computed (see Table 1). While evaluating prescriptions without Q values, we found that the average sensitivity, specificity, PPV, and NPV of the experts were 71.5%,

Table 1. Performance analysis of the results AOP model Vs. human experts (Without and with Q values).

	Without Q values				With Q values			
	Sens	Spec	PPV	NPV	Sens	Spec	PPV	NPV
Physicians	74.3	82.7	94.8	43.1	76.7	84.9	94.8	50.3
Pharmacists	68.8	75.0	95.6	24.6	74.3	94.2	98.7	40.6
Overall	71.5	78.9	95.2	33.9	75.5	89.5	96.7	45.5

Abbreviation: Sens, sensitivity; Spec, specificity; PPV, positive predictive value; NPV, negative predictive value. Note: Confidence intervals (CIs) were small for each parameter and are thus omitted from the reported results.

doi:10.1371/journal.pone.0082401.t001

Part A: Appropriate prescriptions by both system and human experts

Prescription 1:

No.	ICD9CM codes	Disease name
1.	536.8	Dyspepsia & other specified disorders of function of stomach
2.	564.0	Constipation
3.	465.9	Acute upper respiratory infections of unspecified site

No.	ATC codes	Medication name
1.	A02AF02	ordinary salt combinations and antflatulents
2.	A03FA02	cisapride
3.	A06AB06	senna glycosides
4.	J01FA02	spiramycin
5.	R01BA51	phenylpropanolamine, combinations
6.	V03AX	other therapeutic products

Prescription 2:

No.	ICD9CM codes	Disease name
1.	306.9	Unspecified psychophysiological malfunction
2.	564.9	Unspecified functional disorder of intestine

No.	ATC codes	Medication name
1.	A02AX	antacids, other combinations
2.	A03AX13	silicones
3.	A03FA01	metoclopramide
4.	A03FA01	alprazolam

Prescription 3:

No.	ICD9CM codes	Disease name
1.	465.9	Acute upper respiratory infections of unspecified site
2.	784.1	Throat pain
3.	535.40	Other specified gastritis, without mention of hemorrhage

No.	ATC codes	Medication name
1.	A02AF02	ordinary salt combinations and antflatulents
2.	N02BE01	paracetamol
3.	R01BA52	pseudoephedrine, combinations

Part B: Inappropriate prescriptions by both system and human experts

Prescription 1:

No.	ICD9CM codes	Disease name
1.	401.9	Essential hypertension, unspecified

No.	ATC codes	Medication name
1.	A03BB01	butylscopolamine
2.	N02BE01	paracetamol
3.	R01BA52	pseudoephedrine, combinations
4.	R05CB04	eprazinone
5.	V03AX	other therapeutic products

Prescription 2:

No.	ICD9CM codes	Disease name
1.	708.9	Urticaria, unspecified

No.	ATC codes	Medication name
1.	A02AF02	ordinary salt combinations and antflatulents
2.	D06AX04	neomycin
3.	J01DB01	cefalexin
4.	M01AC01	piroxicam
5.	M01AE09	flurbiprofen

Prescription 3:

No.	ICD9CM codes	Disease name
1.	386.9	Unspecified vertiginous syndromes and labyrinthine disorders

No.	ATC codes	Medication name
1.	A02AX	Antacids, other combinations
2.	J01DB01	cefalexin
3.	M01AE01	ibuprofen
4.	N02BE01	paracetamol
5.	R01BA52	pseudoephedrine, combinations
6.	R05CB03	carbocisteine
7.	R05CB06	ambroxol

Figure 3. Few examples of the appropriate and inappropriate prescriptions.

doi:10.1371/journal.pone.0082401.g003

78.9%, 95.2%, and 33.9%, respectively. However, when Q values were disclosed the average sensitivity, specificity, PPV, and NPV, they were 75.5%, 89.5%, 96.7%, and 45.5%, respectively. A detailed analysis of the results for physicians and clinical pharmacologists versus the AOP model are presented in Table 1.

A few examples of the appropriate and inappropriate prescriptions evaluated by both the AOP model and human experts are shown in Figure 3.

Discussion

In this study, we successfully developed, tested, and validated the AOP model, which is able to predict and identify the appropriateness of the prescriptions. The model reveals a high sensitivity and specificity as well as a high positive predictive value, with a negative predictive value around 50%. This AOP model is built on a simple assumption that commonly prescribed medications to a given diseases (DM) are appropriate, and uncommon or rare combinations of disease and medication that might not be appropriate. For example, Sodium Valproate is one of the most effective antiepileptic drugs for suppressing seizure activity; however, it is now well established that Valproate has major teratogenic properties, and when taken during pregnancy can result in more than a 15-fold increased risk of having children with birth defects [44,45]. Despite Valproate is regarded as one of the commonly used antiepileptic drugs, whenever it is prescribed to the pregnant woman, the AOP will show alert stating it is uncommonly prescribed drug. In addition, we could manually update the software if any new or novel drugs have been added to the hospital formulary. Therefore, the novel therapies that had little or no prior prescribing history would not be further flagged as rare associations.

In this study, we could observe some unexpected strong associations between apparently unrelated disease-medications that we believe were attributable to co-occurrence. For example, let us consider Hypertension as a disease and Insulin as a medication, although Insulin is not directly related with Hypertension; in most prescriptions both hypertension and insulin happen to be present. This is because hypertension has strong association with diabetes. Therefore, Fisher’s exact test with p-value was used to find the significant associations [32,33]. This additional rule was applied to the AOP model to re-evaluate the 400 prescriptions. An increase in the sensitivity and PPV was observed, however, NPV remained the same. This suggests there is no clear “best” statistical method; instead, selection of the statistical method is dependent on researchers’ preferences and the parameters used for analysis [42,46–48]. In addition, in order to improve the NPV, the cut-off value in the model has to be selected carefully. In this study, the cut-off value for Q is 1. For any association (DM or MM) having Q value less than 1 is regarded as a negative association or uncommon association. The cut-off value can be adjusted to improve NPV.

The AOP model is based only on DMQs and MMQs without having any references or guidelines such as drug-drug interactions, adverse drug events and dose guidance. The validation results of both physicians and pharmacists (PPV) were nearly 96% accurate for appropriate prescriptions and only 45% accurate for inappropriate prescriptions. Van.D.S.H et al. and Taylor.L.K et al. studied the overridden rate to alerts reported from 49% to 96% causing alert fatigue to the Electronic Medical Records (EMR) users [49–51]. Compared to their findings, we observed our results show an improved accuracy with a low alert rate given by AOP model, in which only 50% were inappropriate. Thus, if the results were extrapolated to the 0.996% (996) inappropriate prescriptions predicted by this model, the human experts would have considered half of them inaccurate. In this study, we noticed that showing Q valve had no effect on the physicians’ decisions (see Figure S1).

The utilization of association rule mining was reported in several studies such as for relating chief complains and lab results

with disorders [33,52,53]. To our knowledge, this is the first study using association rule mining techniques to identify the uncommon DM and MM combinations. Based upon it, we developed a model that would be able to detect and alert the gaps or medication errors in prescriptions. Thus, our model can significantly help to improve patients' safety and quality of care in the hospitals.

In summary, the AOP model has a variety of applications. It can be used to alert physicians if medication errors are detected while prescribing medications using the CPOE system. Additionally, the model could be used to reduce the size of medication list in the CPOE for a given diagnosis. An automated medication listing system and clinical decision support system (CDSS) can also be developed by using the AOP model.

Limitations

This study has some limitations. We used only 103 million prescriptions to construct our model. It is possible if we analyze 300 or 500 million prescriptions, however, it might affect the association strengths which we obtained by using the original data. Second, only two variables are used, diagnostic code and medication codes in the analysis, however, in the real world the prescription of medications depend on various factors like physician behavior, chief-complaints, lab results, and age and gender of patient etc. Third, the seven human experts analyzed only 400 prescriptions in order to evaluate the results we got from our AOP model.

Conclusion

The AOP model developed in this study is able to detect accurately the inappropriate medications prescribed via COPE system. Thus, the PPV of the validation results from both physicians and pharmacists were accurate for the appropriate prescriptions. Moreover, this model could be applied in clinical practice to aid in improving prescription appropriateness, accuracy, patient safety, and patient care.

References

- Gandhi TK, Weingart SN, Borus J, Seger AC, Peterson J, et al. (2003) Adverse drug events in ambulatory care. *N Engl J Med* 348: 1556–1564.
- Kaushal R, Bates DW, Landrigan C, McKenna KJ, Clapp MD, et al. (2001) Medication errors and adverse drug events in pediatric inpatients. *JAMA* 285: 2114–2120.
- Bates DW, Cullen DJ, Laird N, Petersen LA, Small SD, et al. (1995) Incidence of adverse drug events and potential adverse drug events. Implications for prevention. ADE Prevention Study Group. *JAMA* 274: 29–34.
- Barach P, Small SD (2000) Reporting and preventing medical mishaps: lessons from non-medical near miss reporting systems. *BMJ* 320: 759–763.
- Bates DW, Cohen M, Leape LL, Overhage JM, Shabot MM, et al. (2001) Reducing the frequency of errors in medicine using information technology. *J Am Med Inform Assoc* 8: 299–308.
- Anselmi ML, Peduzzi M, Dos Santos CB (2007) Errors in the administration of intravenous medication in Brazilian hospitals. *J Clin Nurs* 16: 1839–1847.
- Valentin A, Capuzzo M, Guidet B, Moreno R, Metnitz B, et al. (2009) Errors in administration of parenteral drugs in intensive care units: multinational prospective study. *BMJ* 338: b814.
- Chapuis C, Roustit M, Bal G, Schwebel C, Pansu P, et al. (2010) Automated drug dispensing system reduces medication errors in an intensive care setting. *Crit Care Med* 38: 2275–2281.
- Hughes RG, Edgerton EA (2005) Reducing pediatric medication errors: children are especially at risk for medication errors. *Am J Nurs* 105: 79–80, 82, 85 passim.
- Wyatt JC (1995) Hospital information management: the need for clinical leadership. *BMJ* 311: 175–178.
- Heathfield H, Pitty D, Hanka R (1998) Evaluating information technology in health care: barriers and challenges. *BMJ* 316: 1959.
- Fernandopulle R, Ferris T, Epstein A, McNeil B, Newhouse J, et al. (2003) A research agenda for bridging the 'quality chasm.'. *Health Aff (Millwood)* 22: 178–190.
- Bates DW, Gawande AA (2003) Improving safety with information technology. *N Engl J Med* 348: 2526–2534.

Supporting Information

Figure S1 Performance statistics for evaluating system by all experts.

(TIF)

Appendix S1 An example of prescriptions in raw data.

(DOCX)

Appendix S2 Brief Introduction before administering questionnaires.

(DOCX)

Appendix S3 Description of questionnaires that used to evaluate the AOP model.

(DOCX)

Acknowledgments

We would like to thank Dr. Hsiu-Chen Lin, M.D, Ph.D from Pediatrics Department, Taipei Medical University Hospital (TMUH); Dr. Hsiu-Li Lin, M.D, Ph.D from Internal Medicine Department, Cathay Hospital, Taipei City; Dr. Ing-Ruey Chang, M.D, Ph.D from New Taipei Institute of Dermatology; Dr. Chung-Ho Hsieh, M.D, Ph.D from Department of General Surgery, Zhong-Xing Branch of Taipei City Hospital; Han-Yi Tsai, PharmD from Department of Pharmacy, Taipei Veterans General Hospital, Taipei City; Hsuan Chia, Yang, PharmD, MS from Department of Pharmacy, Buddhist Dalin Tzu Chi General Hospital, Chiayi; Dr. Yo-Cheng Chang, Ph.D from College of Medical Science and Technology, Taipei Medical University.

Ethical approval: Ethical approval was not required.

Author Contributions

Conceived and designed the experiments: YCJL WSJ PAN SSA CLH MHH HCL. Performed the experiments: PAN CLH. Analyzed the data: PAN HCL. Contributed reagents/materials/analysis tools: MHH WSJ. Wrote the paper: PAN UI DLC. Final approval of the version to be published: UI SSA DLC. Had full access to all of the data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis: YCJL.

- Kaushal R, Shojania KG, Bates DW (2003) Effects of computerized physician order entry and clinical decision support systems on medication safety: a systematic review. *Arch Intern Med* 163: 1409–1416.
- Bates DW, Leape LL, Cullen DJ, Laird N, Petersen LA, et al. (1998) Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. *JAMA* 280: 1311–1316.
- Bates DW, Kuperman G, Teich JM (1994) Computerized physician order entry and quality of care. *Qual Manag Health Care* 2: 18–27.
- Schiff GD, Rucker TD (1998) Computerized prescribing: building the electronic infrastructure for better medication usage. *JAMA* 279: 1024–1029.
- Blendon RJ, DesRoches CM, Brodie M, Benson JM, Rosen AB, et al. (2002) Views of practicing physicians and the public on medical errors. *N Engl J Med* 347: 1933–1940.
- Teich JM, Merchia PR, Schmitz JL, Kuperman GJ, Spurr CD, et al. (2000) Effects of computerized physician order entry on prescribing practices. *Arch Intern Med* 160: 2741–2747.
- Kuperman GJ, Teich JM, Gandhi TK, Bates DW (2001) Patient safety and computerized medication ordering at Brigham and Women's Hospital. *Jt Comm J Qual Improv* 27: 509–521.
- Paoletti RD, Suess TM, Lesko MG, Feroli AA, Kennel JA, et al. (2007) Using bar-code technology and medication observation methodology for safer medication administration. *Am J Health Syst Pharm* 64: 536–543.
- Taxis K, Dean B, Barber N (1999) Hospital drug distribution systems in the UK and Germany—a study of medication errors. *Pharm World Sci* 21: 25–31.
- Bates DW, Kuperman GJ, Wang S, Gandhi T, Kittler A, et al. (2003) Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *J Am Med Inform Assoc* 10: 523–530.
- Chen CC, Chen K, Hsu CY, Li YC (2011) Developing guideline-based decision support systems using protege and jess. *Comput Methods Programs Biomed* 102: 288–294.
- Liu CC, Chang CH, Su MC, Chu HT, Hung SH, et al. (2011) RFID-initiated workflow control to facilitate patient safety and utilization efficiency in operation theater. *Comput Methods Programs Biomed* 104: 435–442.

26. Meyer J, Ostrzinski S, Fredrich D, Havemann C, Krafczyk J, et al. (2012) Efficient data management in a large-scale epidemiology research project. *Comput Methods Programs Biomed* 107: 425–435.
27. Goethals B (2003) Survey on frequent pattern mining. Helsinki, Finland: Univ. of Helsinki.
28. Sarawagi S, Thomas S, Agrawal R (2000) Integrating Association Rule Mining with Relational Database Systems: Alternatives and Implications. *Data Mining and Knowledge Discovery* 4: 89–125.
29. Iskander J, Pool V, Zhou W, English-Bullard R (2006) Data mining in the US using the Vaccine Adverse Event Reporting System. *Drug Saf* 29: 375–384.
30. Carrino JA, Ohno-Machado L (2005) Development of radiology prediction models using feature analysis. *Acad Radiol* 12: 415–421.
31. Cao H, Markatou M, Melton GB, Chiang MF, Hripcsak G (2005) Mining a clinical data warehouse to discover disease-finding associations using co-occurrence statistics. *AMIA Annu Symp Proc. 2006/06/17 ed.* 106–110.
32. Chen ES, Hripcsak G, Xu H, Markatou M, Friedman C (2008) Automated acquisition of disease drug knowledge from biomedical and clinical documents: an initial study. *J Am Med Inform Assoc* 15: 87–98.
33. Wright A, Chen ES, Maloney FL (2010) An automated technique for identifying associations between medications, laboratory results and problems. *J Biomed Inform* 43: 891–901.
34. Brossette SE, Sprague AP, Hardin JM, Waites KB, Jones WT, et al. (1998) Association rules and data mining in hospital infection control and public health surveillance. *J Am Med Inform Assoc* 5: 373–381.
35. Wright A, Sittig DF (2006) Automated development of order sets and corollary orders by data mining in an ambulatory computerized physician order entry system. *AMIA Annu Symp Proc. 2007/01/24 ed.* 819–823.
36. Chen ES, Cimino JJ (2003) Automated discovery of patient-specific clinician information needs using clinical information system log files. *AMIA Annu Symp Proc. 2004/01/20 ed.* 145–149.
37. Mullins IM, Siadaty MS, Lyman J, Scully K, Garrett CT, et al. (2006) Data mining and clinical data repositories: Insights from a 667,000 patient data set. *Comput Biol Med* 36: 1351–1377.
38. Doddi S, Marathe A, Ravi SS, Torney DC (2001) Discovery of association rules in medical data. *Med Inform Internet Med* 26: 25–33.
39. Maimon O, Rokach L (2010) *Data Mining and Knowledge Discovery Handbook*. Springer New York Dordrecht Heidelberg London: Springer Science, Business Media. 1285 p.
40. Agrawal R, Imielinski T, Swami A (1993) Mining Association Rules Between Sets of Items in Large Databases. *ACM SIGMOD Conference* Washington DC, USA.
41. Geng L, Hamilton HJ (2006) Interestingness measures for data mining: A survey. *ACM Comput Surv* 38: 9.
42. Tan P-N, Kumar V, Srivastava J (2002) Selecting the right interestingness measure for association patterns. *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*. Edmonton, Alberta, Canada: ACM. pp. 32–41.
43. Hidalgo CA, Blumm N, Barabasi AL, Christakis NA (2009) A dynamic network approach for the study of human phenotypes. *PLoS Comput Biol* 5: e1000353.
44. Christensen J, Grønberg T, Sørensen M, Schendel D, Parner ET, et al. (2013) Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *JAMA* 309: 1696–1703.
45. Christensen J, Grønberg T, Sørensen M, Schendel D, Parner ET, et al. (2013) Valproate in pregnancy linked to autism in children. *BMJ* 346: f2602.
46. Guillet F, Hamilton HJ (Feb 1st, 2007) *Quality Measures in Data Mining. Studies in Computational Intelligence*: Springer. 314 p.
47. Kodratoff Y (2001) Comparing machine learning and knowledge discovery in databases: an application to knowledge discovery in texts. In: Georgios P, Vangelis K, Constantine DS, editors. *Machine Learning and Its Applications*: Springer-Verlag New York, Inc. pp. 1–21.
48. Smyth P, Goodman RM (1991) *Rule Induction Using Information Theory. Knowledge Discovery in Databases*. pp. 159–176.
49. Taylor LK, Kawasumi Y, Bartlett G, Tamblin R (2005) Inappropriate prescribing practices: the challenge and opportunity for patient safety. *Healthc Q* 8 Spec No: 81–85.
50. Van Der Sijs H, Aarts J, Vulto A, Berg M (2006) Overriding of drug safety alerts in computerized physician order entry. *Journal of the American Medical Informatics Association* 13: 138–147.
51. Van der Sijs H, Mulder A, van Gelder T, Aarts J, Berg M, et al. (2009) Drug safety alert generation and overriding in a large Dutch university medical centre. *Pharmacoepidemiol Drug Saf* 18: 941–947.
52. Lee C-H, Chen J-C-Y, Tseng VS (2011) A novel data mining mechanism considering bio-signal and environmental data with applications on asthma monitoring. *Computer Methods and Programs in Biomedicine* 101: 44–61.
53. Yang DH, Kang JH, Park YB, Park YJ, Oh HS, et al. (2013) Association rule mining and network analysis in oriental medicine. *PLoS One* 8: e59241.