A Personalized and Learning Approach for Identifying Drugs with Adverse Events

Sug Kyun Shin¹, Ho Hur², Eun Kyung Cheon³, Ock Hee Oh¹, Jeong Seon Lee⁵, Woo Jin Ko⁶, Beom Seok Kim⁷, and YoungOk Kwon⁵

¹Department of Internal Medicine, Nephrology Division, National Health Insurance Service Ilsan Hospital, Goyang; Departments of Surgery, Pharmacy, and Urology, National Health Insurance Service Ilsan Hospital, Goyang; ⁴FirstDIS Ltd., Seoul; ⁵Division of Business Administration, Sookmyung Women’s University, Seoul; ⁶Department of Internal Medicine, Nephrology Division, College of Medicine, Yonsei University, Seoul, Korea.

Purpose: Adverse drug events (ADEs) are associated with high health and financial costs and have increased as more elderly patients treated with multiple medications emerge in an aging society. It has thus become challenging for physicians to identify drugs causing adverse events. This study proposes a novel approach that can improve clinical decision making with recommendations on ADE causative drugs based on patient information, drug information, and previous ADE cases.

Materials and Methods: We introduce a personalized and learning approach for detecting drugs with a specific adverse event, where recommendations tailored to each patient are generated using data mining techniques. Recommendations could be improved by learning the associations of patients and ADEs as more ADE cases are accumulated through iterations. After consulting the system-generated recommendations, a physician can alter prescriptions accordingly and report feedback, enabling the system to evolve with actual causal relationships.

Results: A prototype system is developed using ADE cases reported over 1.5 years and recommendations obtained from decision tree analysis are validated by physicians. Two representative cases demonstrate that the personalized recommendations could contribute to more prompt and accurate responses to ADEs.

Conclusion: The current system where the information of individual drugs exists but is not organized in such a way that facilitates the extraction of relevant information together can be complemented with the proposed approach to enhance the treatment of patients with ADEs. Our illustrative results show the promise of the proposed system and further studies are expected to validate its performance with quantitative measures.

Key Words: Adverse drug event, data mining, clinical decision making, learning

INTRODUCTION

Adverse drug events (ADEs), as an important cause of morbidity and mortality, have been one of the most critical issues in healthcare, leading to high health and financial costs.¹⁻⁵ ADE is defined as "any injury resulting from medical intervention related to a drug." and, according to the Food and Drug Administration, it can arise with "any use of the drug and with any route of administration, formulation, or dose, including an overdose." While ADEs refer to an injury that occurs during treatment with medications but do not necessarily have a causal relationship with the treatment, adverse drug reactions (ADRs), as a subset of ADEs, include only injury that occurs with an appropriate use of medication.⁶,⁷ According to National Center for Health Statistics, drug-related admissions to hospitals are significantly increasing and ADEs are one of the key reasons for admission to hospital.¹⁰⁻¹¹ The costs of ADEs,
therefore, represent a considerable burden in overall medical expenses.\textsuperscript{5,14,15}

However, it has been proven that many ADEs were preventable because they were predictable from the known pharmacology of the drugs and their interactions. Some studies show that preventable ADEs were caused mostly by errors at the ordering and administration stages.\textsuperscript{6,16} Healthcare providers have accordingly attempted to use computerized information systems designed to detect errors in time to prevent ADEs, which can ultimately reduce the soaring cost of healthcare.\textsuperscript{17-21} For example, retrospective medical chart reviews, voluntary incident reporting, and prospective ADE surveillance have been used as a method for identifying ADEs but they are known to be expensive and time-consuming.

To overcome these problems, computerized methods based on electronic health record data are developed, which allow for prospective detection and prompt interventions. In particular, many efforts are made to generate ADE detection rules using data mining techniques.\textsuperscript{22,23} Although the accuracy or efficacy of the detection rules is difficult to validate in general, most studies demonstrate their outperformance over traditional voluntary reporting or manual chart review.\textsuperscript{24} With the increase in digital medical data and recent technological advances, healthcare providers try to extract insights from available healthcare big data, which can be one example of evidence-based medicine.\textsuperscript{25} To preemptively predict ADEs, recent work has started to use data from multiple sources to identify signals of potential adverse events, including electronic health records and narrative documents posted on social media site or medical message boards.\textsuperscript{26-28}

Although earlier work has been focused on the detection and prediction of ADEs using various computerized methods, a few research has been conducted to identify the most likely culprit drugs causing an adverse event. In an aging society, there are more elderly patients who tend to take more drugs at the same time and also over longer periods,\textsuperscript{29,30} expecting that patients treated with multiple medications are more likely to suffer an ADE. Thus, it becomes challenging for physicians to identify the potential drugs that can cause an adverse event among the drugs that a patient is already taking.

This study addresses this problem by proposing a personalized and learning approach that can improve the treatment of patients who already suffer from ADEs. In other words, the focus of this study is on helping clinical decisions of physicians by recommending ADE causative drugs based on healthcare big data, mostly in situations where patient prescriptions should be altered to relieve adverse events. The novelty of the proposed approach comes from the fact that causative drugs are recommended to enhance clinical decision to adverse events. The recommendations are personalized to each patient and evolve over time based on historical ADEs and outcome data as well as patient information.

**MATERIALS AND METHODS**

We first explored the current issues and challenges present in National Health Insurance Service (NHIS) Ilsan Hospital. When a patient is presented with general side effects, such as nausea and vomiting, after taking their medication, it is not easy for physicians to immediately identify culprit drugs involved in adverse events, especially in polypharmacy. In the current system, physicians can find a list of drugs that the patient is currently taking on one screen but more detailed information for each drug appears on a separate page. Therefore, it is not possible for physicians to read through all the pages and select drugs causing a particular symptom in a short period of time. Furthermore, the problem is exacerbated if the patient is taking multiple drugs that present with the same side effect symptom(s).

To solve this problem, we proposed a new system that can provide recommendations on the most likely ADE culprit drugs to facilitate more prompt and effective clinical decisions. The system-generated recommendations are personalized for each patient by using three different categories of data -patient information, drug information, and ADE reports. Then, using data mining techniques, the patterns of ADEs were discovered for each group of patients with similar demographics and medical records, and the most likely culprit drugs were recommended for each group. More importantly, the proposed system is designed to evolve using an iterative learning approach where more ADE cases and physicians’ feedback, i.e., causality assessment between the recommended drug and adverse events, can be used for better recommendations.

The study protocol was approved by the Institutional Review Board of the authors’ institution (NHIMC 2015-04-026).

**Overview of the proposed approach**

The data we require to generate recommendations includes patient data that is based on electronic medical records, general drug information obtained from pharmaceutical R&D data, and historical ADE data reported by hospital staffs. The ADE database is built by integrating all the data, and data mining technique is applied to identify and recommend the most likely culprit drugs causing a particular patient’s ADEs. While any recommendation techniques used in e-commerce applications can also be applied, a decision tree model, as an example, has been chosen and tested.\textsuperscript{29-31} The recommendations are then accessed by physicians, and their feedback on the actual causal relationships between the recommended drug and adverse events is also stored in the ADE database, which will be used for subsequent recommendations. An overview of the proposed approach is described in Fig. 1.

We developed a prototype system using open source programs including Apache web server, Mysql database, and PHP programs. To generate recommendations, RapidMiner is used as a data mining tool.\textsuperscript{32} Two representative cases are illustrated to demonstrate how this proposed approach can overcome
the pitfalls of current systems in one hospital setting. Although further research is needed for validation, we believe this work could provide a direction to help the healthcare industry to solve the ADE-related problems by enabling more prompt and efficient clinical decision to the ADEs.

**Data sources and descriptive statistics**

We collected the data from the following three sources. First, patient information was obtained from Electronic Medical Record/Order Communication System (EMR/OCS) in the NHIS Ilsan Hospital in Korea. By considering the information of each patient, e.g., in terms of gender, age, patient type (out-patient, inpatient, and emergency), medical department, treatment date, medical order (name of drug, drug code, dose, etc.), different drugs can be recommended as the most likely causative drugs. Sample patients for testing the proposed system were chosen among 16 patients treated at 23 departments from January to June 2015.

Second, general drug information was obtained from FirstDIS Ltd., the Drug Information Research Institute at Sookmyung Women’s University (http://www.firstdis.co.kr, Seoul, Korea). The collected drug information was being used by the NHIS Ilsan Hospital but the information of each drug appeared on a separate page in the current system. Moreover, it has become even more difficult to retrieve relevant information since a patient takes multiple drugs and a drug can have multiple side effects. In addition to patient information, frequency and severity of an adverse event were used in determining the possibility of each candidate drug as being the cause of an adverse event.

Lastly, we used historical data on ADEs at the chosen hospital. The NHIS Ilsan Hospital tracks ADEs by hospital staffs. The seriousness and the causality between an adverse event and the drug whose prescription was altered to relieve the adverse event were used in ranking the candidate drugs and determining the most likely causative drugs. We collected 1147 observations reported between January 2014 and June 2015. Initially, the ADE cases that were manually reported by staffs are used to create rules. However, with the proposed approach, ADE cases can be automatically stored into the ADE database without the aid of staffs. Furthermore, physicians provide their prescription change and feedback on its outcome, which are also directly stored in the database. Details on the three types of data used for the proposed system are summarized in Table 1.

We next present the descriptive statistics for the ADE reports in Tables 2 and 3. We had more female patients than male patients, and the majority of the patients were above 40 years of age. Most patients were from the departments of general surgery and orthopedics. The 5 most frequently found drugs and side effects are listed. Sixteen percent of the ADEs were attributable to Tramadol HCI and nausea was the most frequently occurring ADE, appearing in 22% of cases. Note that the sample data does not necessarily present all the ADE cases in the chosen hospital and the rules used to generate recommendations would be different with different sets of data. This work is geared to show the promise of the proposed approach with exemplar cases.

**Application of recommendation techniques**

Among the various data mining techniques, we applied a decision tree model to find the recommendations for each patient. The reason we chose the decision tree model is that the output of a decision tree can be easily interpreted as rules, and any assumptions on the data are not required. The rule with the highest confidence is taken to treat adverse events, which explains why a particular drug can be the cause of the chosen adverse events. Furthermore, physician feedback on the system-generated recommendations, i.e., causality assessment,
was periodically stored into the ADE database, allowing the updating of the decision model. Therefore, the model was able to evolve over time as more ADE cases were accumulated, along with causality information assessed by physicians. This personalized and iterative learning approach helped develop a more accurate decision tree model with newly added ADE cases, which could eventually improve the performance of recommendations. In addition to the decision tree model, other data mining techniques can be used for recommendations and, furthermore, several techniques can be combined to improve the output. For example, as more patient data is collected, we can use a clustering method to divide patients into a number of groups and then a decision tree can be constructed for each group. In this way, patients within the same group are likely to have similar but more accurate recommendations, as a result enhancing the performance of the system.

RESULTS

Development of a prototype system in one hospital setting
As described in Fig. 2, we built an ADE database from the above-mentioned three data sources, and then a prototype system with a general flow. When a patient presented with a particular ADE, a physician could enter information such as patient number (coded), name of symptom, and the date when the patient started feeling the symptom. A list of drugs taken by the patient at the date of symptom commencement was then extracted from the ADE database, and the most likely culprit drugs were recommended based on previous ADE reports. Physicians were asked to make a clinical decision regarding the symptom by choosing a drug, which may or may not be the most highly ranked in a list of recommendations, and by changing the prescription of the chosen drug (e.g., reduction of dosage or discontinuation). More importantly, the physician would assess the causality between the chosen drug and the ADE after a reasonable time period has elapsed, and record the outcome in the database, which could then be used for future recommendations. Following WHO-Uppsala Monitoring Center criteria, causality assessment of ADEs is categorized into four classes: certain, probable, possible, unlikely.23,34

We now illustrate two cases to show how the proposed system works to recommend ADE causative drugs for each patient and how the recommendations can be used in supporting a physician’s clinical decision making. In the two cases, the most highly recommended drug was validated as a key cause by several physicians.

| Table 1. Three Types of Data—Description and Source |
|-----------------------------------------------------|
| **Type of data** | **Description** | **Source** |
| Patient information | Patient no., gender, age, patient type (outpatient, emergency, inpatient), department, treatment date, medication order (name of drug, drug code, dose, etc.) | EMR/OCS |
| Drug information | Drug code, drug name, side effect name, frequency (below 1%, 1–5%, above 5%), severity (severe-life threatening, less severe, non-life threatening) | FirstDIS |
| Historical ADEs data | Patient no., drug code, side effect name, seriousness (serious, not serious), causality (certain 100%, probable/likely 75%, possible 50%, unlikely 25%) | ADEs reports |

ADEs, adverse drug events; EMR, Electronic Medical Record; OCS, Order Communication System.

| Table 2. Patient Variables |
|---------------------------|
| **Gender** | **Frequency** | **%** |
| Female | 722 | 62.9 |
| Male | 425 | 37.1 |
| Total | 1147 | 100.0 |
| **Age** | | |
| ≤20 | 53 | 4.6 |
| 21–40 | 213 | 18.6 |
| 41–60 | 465 | 40.5 |
| ≥61 | 416 | 36.3 |
| Total | 1147 | 100.0 |
| **Medical department (Top 5)** | | |
| General surgery | 172 | 15.0 |
| Orthopedics | 168 | 14.6 |
| Neurosurgery | 116 | 10.1 |
| Gastroenterology | 101 | 8.8 |
| Comprehensive medical testing center | 80 | 7.0 |

| Table 3. Drugs and Adverse Events in Our Dataset (Top 5) |
|--------------------------------------------------------|
| **Drug name** | **Frequency** | **%** |
| Tramadol HCI | 184 | 16.0 |
| Tridol | 87 | 7.6 |
| Iomeron | 53 | 4.6 |
| Paramacet | 51 | 4.4 |
| Acupan | 40 | 3.5 |
| **Adverse event** | | |
| Nausea | 253 | 22.1 |
| Nausea-vomiting | 167 | 14.8 |
| Dizziness | 97 | 8.5 |
| Rash, urticaria, pruritus | 71 | 6.2 |
| Urticaria | 67 | 5.8 |
an adverse event. According to the decision tree, medical de-

Tramadol HCI was recommended as the most likely cause of

were ranked as listed in step 2 (Fig. 3), more specifically why

causality between vomiting and Tramadol HCI as probable/

Finally, the physician would have reported the outcome of the

This case features a 73-year-old male patient from the depart-

ment was the most important factor to identify ADE caus-

ative drugs. If we consider the case of the patients from the de-

partment of neurosurgery, there were only two patients in the

same group as the patient of case 1 (i.e., age between 71 and 80

and gender=male) due to a limited dataset. While not enough

data are available, the highlighted rule clearly shows that a pa-

ient similar to both of them may also have had vomiting after

taking Tramadol HCl with 100% confidence.

Case 2: female, 30-year-old, department of surgery, dizziness,

not serious, from 2015-04-18

This case features a 30-year-old female patient from the de-

partment of surgery who experienced dizziness after taking

medication on April 18, 2015. Similar to the case 1, a physician

could have used the proposed system for this patient follow-

ning the previously outlined four steps. In this case, Pethidine

HCI was recommended by the system as the drug that may be

the source of dizziness side effect. Fig. 5 shows the rules applied

to generate the recommendation in this case, as well as the

recommendation itself. This permits the physician fully under-

stand the details of the recommendation process.

**DISCUSSION**

In the current system, the information of individual drugs ex-

ists, however, is not organized in such a way where relevant

information can be collated and extracted together. The sys-

tem can be complemented with the personalized and learning

approach proposed in this work to enhance the treatment of

patients with ADEs. The applications of data mining tech-
niques with patient information, drug information, and ADE

reports would be able to tailor treatment recommendations for
each patient. The performance of recommendations would
improve by learning from more cases and the feedback of phy-
sicians.

As a future work, the proposed system can be extended with
external healthcare data such as social media postings, fo-
rums, and Internet blogs, and it would also be interesting to
see how the system can be applied to mobile applications. In
addition, while the proposed approach works well when there
are many patients with similar demographics, medical history,
and symptoms because more data allows for a more accurate
model, we can extend it to manage rare adverse events with
additional data. For example, ARD Probability Scale (Naranjo
Scale)\(^{35}\) can be used as an additional attribute in creating a
model where the drug with the highest score is more likely to
be recommended. However, it may be considered intrusive to
ask physicians to answer the questionnaire to gain Naranjo
Scale for each drug, especially when a patient is treated with
multiple medications. Thus, it is also important to find a bal-
ance between intrusiveness and rich information. It would be
interesting to find less intrusive but effective methods for rare

**Exemplar cases and evaluations**

**Case 1: male, 73-year-old, department of neurosurgery,
vomiting, not serious, from 2015-03-09**

This case features a 73-year-old male patient from the depart-
ment of neurosurgery who experienced vomiting after taking
medication on March 9, 2015. A physician could use the pro-
posed system for this patient following the four steps outlined
in Fig. 3. The six drugs amongst the list of drugs that the pa-
tient was taking on the symptom start date, were found to be
related to vomiting according to the database. Among the six
drugs, Tramadol HCl was the most highly ranked according
to the decision tree model, and thus would have been rec-
ommended as the main culprit drug. In addition to the name of
candidate drugs/substances, the system displayed more in-
formation on the drug such as the frequency and severity of
the symptom. The physician could have chosen any drug from
the list of recommendations based on past treatments and
medications on the particular patient. In this case, the physi-
cian would have accepted the recommendation, and hence
entered a prescription change for the selected drug in step 3.
Finally, the physician would have reported the outcome of the
prescription change a few days later, i.e., assessment of the
causality between vomiting and Tramadol HCl as probable/likely (50%). This result would have been stored and used later
for patients with symptoms and demographics similar to the
ones of this patient.

Fig. 4 presents the decision tree and the rules that were used
to provide the recommendations. This shows how the six drugs
were ranked as listed in step 2 (Fig. 3), more specifically why
Tramadol HCl was recommended as the most likely cause of
an adverse event. According to the decision tree, medical de-

![Fig. 2. Overall flow of the proposed system.](image-url)
Identifying Drugs with Adverse Events

[Step 1] Enter patient no., symptom name, start date

[Step 2] Select the most highly ranked drug among the recommendations with "Vomiting"

[Step 3] Enter any change(s) in medication of the selected drug

[Step 4] Assess the causality between the selected drug and an adverse event

Fig. 3. Overall flow of the proposed system for case 1 (Steps 1–4).
cases, which we leave as a future work. Furthermore, some symptoms can be caused by a certain disease itself, rather than by a drug taken for treatment. Thus, the current system can be augmented with patients’ clinical information such as diagnosis, and surgical history, which we expect to help distinguish ADEs from disease symptoms. The effects of the additional information will be elucidated in further studies.

While two representative cases demonstrate the potential for the proposed approach in helping clinical decision making, this study should be interpreted in light of several shortcomings. First, the recommendations were obtained by applying a decision tree model, however, any data mining technique can be used to generate a recommendation. The effectiveness of different data mining techniques should be examined and compared. Second, our data is limited to only a few types of adverse events since the ADE cases were collected from a small subset of departments in the hospital. To ensure the robustness of the proposed approach, more diverse cases need to be analyzed. The proposed system is designed to collect ADE cases from all departments across the hospital, and different rules can be created for the patients from different departments, showing the potential and extensibility of the system. Finally, the proposed approach in this study was evaluated with only two cases. Further studies are therefore needed to assess the direct effect of the system on treatment outcomes using quantitative performance metrics. Despite these limitations, we believe that our study provides valuable implications for the importance of the incorporation of patient and drug information in clinical decision making.

ACKNOWLEDGEMENTS

This work was supported by the National Health Insurance Corporation Ilsan Hospital grant (2015-15).
REFERENCES

1. Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, et al. Adverse drug reactions as cause of admission to hospital: prospective analysis of 18,820 patients. BMJ 2004;329:15-9.

2. Budnitz DS, Pollock DA, Weidenbach KN, Mendelsohn AB, Schroeder TJ, Annest JL. National surveillance of emergency department visits for outpatient adverse drug events. JAMA 2006;296:1858-66.

3. Page RL, Ruscin JM. The risk of adverse drug events and hospital-related morbidity and mortality among older adults with potentially inappropriate medication use. Am J Geriatr Pharmacother 2006;4:297-305.

4. Sarkar U, López A, Maselli JH, Gonzales R. Adverse drug events in U.S. adult ambulatory medical care. Health Serv Res 2011;46:1517-33.

5. Poudel DR, Acharya P, Ghimire S, Dhitral R, Bharati R. Burden of hospitalizations related to adverse drug events in the USA: a retrospective analysis from large inpatient database. Pharmacoepidemiol Drug Saf 2017;26:635-41.

6. Leape LL, Bates DW, Cullen DJ, Cooper J, Demonaco HJ, Gallivan J, et al. Systems analysis of adverse drug events. JAMA 1995;274:35-43.

7. Safety Reporting Requirements for INDs and BA/BE Studies, Food and Drug Administration, 2012, https://www.fda.gov/downloads/drugs/guidanceregulatoryinformation/guidances/ucm227351.pdf.

8. Bates DW, Boyle DL, Vander Vliet MB, Schneider J, Leape L. Relationship between medication errors and adverse drug events. J Gen Intern Med 1995;10:199-205.

9. Nebeker JR, Barach P, Samore MH. Clarifying adverse drug events: a clinician’s guide to terminology, documentation, and reporting. Ann Intern Med 2004;140:795-801.

10. Bourgeois FT, Shannon MW, Valim C, Mandl KD. Adverse drug events in the outpatient setting: an 11-year national analysis. Pharmacoepidemiol Drug Saf 2010;19:901-10.

11. Howard RL, Averly AJ, Howard PD, Partridge M. Investigation into the reasons for preventable drug related admissions to a med admissions unit: observational study. Qual Saf Health Care 2003;12:280-5.

12. Kongkaew C, Hann M, Mandal I, Williams SD, Metcalfe D, Noyce PR, et al. Risk factors for hospital admissions associated with adverse drug events. Pharmacotherapy 2013;33:827-37.

13. Lazarou J, Pomeranz BH, Corey PN. Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. JAMA 1998;279:1200-5.

14. Bates DW, Spell N, Cullen DJ, Burdick E, Laird N, Petersen LA, et al. The costs of adverse drug events in hospitalized patients. JAMA 1997;277:307-11.

15. Ernst FR, Grizzle AJ. Drug-related morbidity and mortality: updating the cost-of-illness model. J Am Pharm Assoc 2001;41:192-9.

16. Bates DW, Cullen DJ, Laird N, Petersen LA, Small SD, Servi D, et al. Incidence of adverse drug events and potential adverse drug events. Implications for prevention. JAMA 1995;274:29-34.

17. Jha AK, Kuperman GJ, Ritzenberg E, Teich JM, Bates DW. Identifying hospital admissions due to adverse drug events using a computer-based monitor. Pharmacoepidemiol Drug Saf 2001;10:113-9.

18. Murff HJ, Patel VL, Hripcsak G, Bates DW. Detecting adverse events for patient safety research: a review of current methodologies. J Biomed Inform 2003;36:131-43.

19. Beuscart R, McNair P, Darmoni SJ, Koutkia V, Maglaveras N, Beusc- cart-Zephir MC, et al. Patient safety: detection and prevention of adverse drug events. Stud Health Technol Inform 2009;150:968-71.

20. Lee JH, Park KH, Moon HJ, Lee YW, Park JW, Hong CS. Spontaneous reporting of adverse drug reactions through electronic submission from regional society healthcare professionals in Korea. Yonsei Med J 2012;53:1022-7.

21. Park K, Soukavong M, Kim J, Kwon KE, Jin XM, Lee J, et al. Signal detection of imipenem compared to other drugs from Korea adverse event reporting system database. Yonsei Med J 2017;58:564-9.

22. Chazard E, Ficheur G, Bernonville S, Luyckx M, Beuscart R. Data mining to generate adverse drug events detection rules. IEEE Trans Inf Technol Biomed 2011;15:823-30.

23. Chazard E, Merlin B, Ficheur G, Sarfati JC; PSIP Consortium, Beusc- cart R. Detection of adverse drug events: proposal of a data model. Stud Health Technol Inform 2009;148:63-74.

24. Forster AJ, Jennings A, Chow C, Leeder C, van Walraven C. A systematic review to evaluate the accuracy of electronic adverse drug event detection. J Am Med Inform Assoc 2012;19:31-8.

25. Groves P, Kayayli B, Knott D, Van Kuiken S. The ‘big data’ revolution in healthcare. McKinsey Quarterly. 2013.

26. Leaman R, Wojtulewicz L, Sullivan R, Skariah A, Yang J, Gonzalez G. Towards internet-age pharmacovigilance: extracting adverse drug reactions from user posts to health-related social networks. Proc. 2010 Works BioNLP 2010;117-25.

27. Benton A, Ungar L, Hill S, Hennessey S, Mao J, Chung A, et al. Identifying potential adverse effects using the web: a new approach to medical hypothesis generation. J Biomed Inform 2011;44:989-96.

28. HARP R, Vilar S, Dumouchel W, Salmasian H, Haerian K, Shah NH, et al. Combing signals from spontaneous reports and electronic health records for detection of adverse drug reactions. J Am Med Inform Assoc 2013;20:413-9.

29. Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. IEEE Trans Knowl Data Eng 2005;17:734-49.

30. Linoff GS, Berry MJ. Data mining techniques: for marketing, sales, and customer relationship management. 3rd ed. John Wiley & Sons; 2011, p.888.

31. Provest F, Fawcett T. Data science for business: what you need to know about data mining and data-analytic thinking. 1st ed. Sebastopol CA, O'Reilly Media; 2013.

32. Kotu V, Deshpande B. Predictive analytics and data mining: concepts and practice with rapidminer. Waltham MA, Morgan Kaufmann; 2014.

33. Edwards IR, Aronson JK. Adverse drug reactions: definitions, di- agnosis, and management. Lancet 2000;356:1255-9.

34. Son MK, Lee YW, Jung HY, Yi SW, Lee KH, Kim SU, et al. Comparison of the Naranjo and WHO-Uppsala Monitoring Centre criteria for causality assessment of adverse drug reactions. Korean J Med 2008;74:181-7.

35. Naranjo CA, Busto U, Sellers EM, Sandor P, Ruiz L, Roberts EA, et al. A method for estimating the probability of adverse drug re- actions. Clin Pharmacol Ther 1981;30:239-45.