On Knowledge Discovery in Open Medical Data on the Example of the FDA Drug Adverse Event Reporting System for Alendronate (Fosamax)

Pinar Yildirim[1], Ilyas Ozgur Ekmekci[1], and Andreas Holzinger[2]

[1] Department of Computer Engineering, Faculty of Engineering & Architecture, Okan University, Istanbul, Turkey
{pinar.yildirim,oekmekci}@okan.edu.tr
[2] Institute for Medical Informatics, Statistics & Documentation Medical University Graz, A-8036 Graz, Austria
andreas.holzinger@meduni-graz.at

Abstract. In this paper, we present a study to discover hidden patterns in the reports of the public release of the Food and Drug Administration (FDA)'s Adverse Event Reporting System (AERS) for alendronate (fosamax) drug. Alendronate (fosamax) is a widely used medication for the treatment of osteoporosis disease. Osteoporosis is recognised as an important public health problem because of the significant morbidity, mortality and costs of treatment. We consider the importance of alendronate (fosamax) for medical research and explore the relationship between patient demographics information, the adverse event outcomes and drug's adverse events. We analyze the FDA's AERS which cover the period from the third quarter of 2005 through the second quarter of 2012 and create a dataset for association analysis. Both Apriori and Predictive Apriori algorithms are used for implementation which generates rules and the results are interpreted and evaluated. According to the results, some interesting rules and associations are obtained from the dataset. We believe that our results can be useful for medical researchers and decision making at pharmaceutical companies.

Keywords: Open medical data, knowledge discovery, biomedical data mining, osteoporosis, drug adverse event, alendronate (fosamax), apriori algorithm, co-occurrence analysis.

1 Introduction

Open data is generally a big issue in every scientific community. Much has changed within the last years and Science has become part of our modern civilization and should be, and be seen to be, a public enterprise, not a private enterprise done behind closed laboratory doors [1]. Consequently, Open Data is considered an important issue in several scientific communities for some time now [2] and very recently it is in debate in the biomedical area [3], [4].

A. Holzinger and G. Pasi (Eds.): HCI-KDD 2013, LNCS 7947, pp. 195–206, 2013.

A big asset of open data in research is that it allows to build on the work of others more efficiently and helps to speed the progress of science [5]; because to build on previous discoveries, there must be trust in the validity of prior research – but most important, it facilitates trust between researchers and with the public, due to the fact that it allows secondary analyses that expand the usefulness of datasets and the resulting knowledge gained [6].

Postmarket surveillance for adverse events is an essential component of every national and regional health system for assuring drug safety [7]. The Food and Drug Administration (FDA) in USA is responsible not only for approving drugs for marketing but also for monitoring their safety after they reach the market. This function is carried out by the FDA's Office of Drug Safety, which maintains a spontaneous reporting database called the Adverse Event Reporting System (AERS). AERS receives adverse events information from two principal sources: mandatory reports from pharmaceutical companies on adverse events that had been spontaneously communicated to the firms, primarily by physicians and pharmacists: and adverse event reports that physicians, pharmacists, nurses, dentists and consumers submit directly to the FDA's MedWatch program. Since 1969, more than 2 million adverse event reports have been submitted to the FDA. In the United States, the estimated cost of morbidity and mortality related to adverse drug reactions (ADRs) is more than $75 million annually, and ADRs are among the top 10 leading causes of death [8].

The FDA can restrict distribution of a drug, and on rare occasions, it may request a drug's withdrawal from the market, or the manufacturer may voluntarily withdraw the drug [8]. In September 2004, Merck&Co, Inc, voluntarily withdrew rofecoxib (vioxx) from the global market because of an increased risk of cardiovascular events. Two months later, the FDA announced that manufacturers of isotretinoin will obtain registration of prescribers of isotretinoin, dispencing pharmacies, and patients who are prescribed the drug. The agency also announced the requirement of documentation of a negative pregnancy test result before isotretinoin is given to women capable of becoming pregnant. In April 2005, valdecoxib was withdrawn from the market because of serious dermatological conditions and an unfavorable risk vs benefit profile [9].

The aim of this study is to explore hidden relationship between alendronate (fosamax) and adverse events. Alendronate(fosamax) is the main medication for osteoporosis disease. Osteoporosis is widely recognised as an important public health problem because of the significant morbidity, mortality and costs associated with its complications, namely fractures of the hip, spine, forearm and other skeletal sites[10]. Across the whole of Europe, an estimated 3.1million fragility fractures occur each year in men and women age 50 years or over, including 620,000 cases of hip fracture, 490,000 clinical vertebral fractures and 574,000 forearm fractures. The incidence of fractures is highest amongst elderly white women, with one in every two women suffering an osteoporosis related fracture in their lifetime [11], [12], [13].

It is also known that there are some adverse events of drugs that are associated with a specific gender. Hence, we also considered that patient demographics such as age, gender may be associated with differential risk to alendronate(fosamax) and we present an approach for knowledge discovery on the adverse events of this drug. We analyzed the data from the FDA's AERS to discover associations between patient

demographics, the adverse event outcomes (death, hospitalization, disability etc.) and adverse events of alendronate(fosamax). We aimed to capture some serious and useful information buried in event reports that are not easily recognized in clinical trials by researchers, clinicians and pharmaceutical companies.

2 Related Work

Novel data mining algorithms were developed and several studies were performed for knowledge discovery on drug adverse events associations. Tatoneti et al., mined the FDA's AERS for side-effect profiles involving glucose homeostatis and found a surprisingly strong signal for comedication with pravastatin and paroxetine[14].

Kadoyama et al., searched the FDA's AERS and carried out a study to confirm whether the database could suggest the hypersensitivity reactions caused by anticancer agents, paclitaxel, docetaxel, procarbazine, asparaginase, teniposide and etoposide. They used some data mining algorithms such as proportional reporting ratio (PRR), the reporting odds ratio(ROR) and the empirical Bayes geometric mean (EBGM) to identify drug-associated adverse events and consequently, they detected some associations[15].

Tamura et al., reviewed FDA AERS to assess the bleeding complications induced by the administration of antiplatelets and to attempt to determine the rank-order of the association. According to their results, both aspirin and clopidogrel were associated with haemorrhage, but the association was more noteworthy for clopidogrel; however, for gastrointestinal bleeding complications, the statistical metrics suggested a stronger association for aspirin than clopidogrel[16].

Gandhi et al., investigated the FDA's AERS to detect cardiovascular thromboembolic events associated with febuxosat. They implemented Bayesian statistics within the neural network architecture to identify potential risks of febuxosat. Their study indicated continuing combination cases of cardiovascular thrombotic events associated with the use of febuxosat in gout patients [17].

3 Methods

Data Sources

Input data for our study are taken from the public release of the FDA's AERS database, which covers the period from the third quarter of 2005 through the second of 2012. The data structure of AERS consists of 7 datasets: patient demographic and administrative information (DEMO), drug/biologic information (DRUG), adverse events (REAC), patient outcomes (OUTC), Report Sources (RPSR), drug therapy start and end dates (THER), and indications for use/diagnosis (INDI).The adverse events in REAC are coded using preferred terms (PTs) in the Medical Dictionary for Regulatory Activities (MedDRA) terminology. All ASCII data files are linked using ISR, a unique number for identifying an AER. Three of 7 files are linked using DRUG_SEQ, a unique number for identifying a drug for an ISR [18]. The outline of 7 datasets is shown in Fig.1 [19].

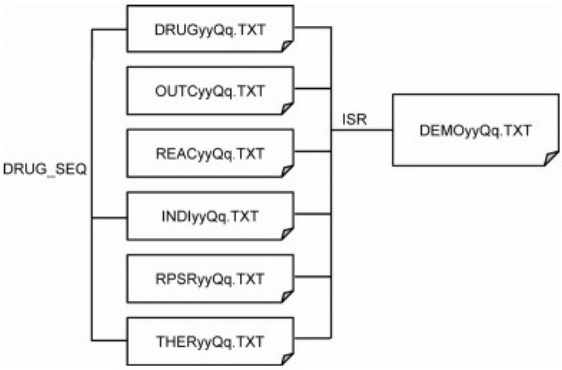

Fig. 1. Data structure of AERS

Association Rule Mining and Apriori Algorithm

Frequent sets play an essential role in studies on data mining to find interesting patterns from databases, such as association rules, correlations and sequences. Association rule mining explores hidden association or correlation relationships among a large set of data items. With massive amounts of data continuously being collected and stored, many industries are becoming interested in mining association rules from their database. The discovery of interesting association relationships among huge amounts of business transaction records can help in many business decision making processes[20].

Association rule mining methods try to find frequent items in databases. Given a set of transactions T(each transaction is a set of items), an association rule can be expressed in the form $X \Rightarrow Y$, where X and Y are mutually exclusive sets of items[21].

The rule's statistical significance is measured by *support*, and the rule's strength by *confidence*. The *support* of the rule is defined as the percentage of transactions in T that contain both X and Y, and may be regarded as , the probability of $X \cup Y$ (X union Y). The confidence of the association rule is the ratio of the support of the itemset $X \cup Y$ to the support of the itemset X, which roughly corresponds to the conditional probability *P(Y/X)*.

Association rule generation is usually split up into two steps:

- Find all combinations of items whose supports are greater than a user-specified minimum support(threshold). The combinations are called frequent itemsets.
- Use the items from frequent itemsets to generate the desired rules. More specifically, the confidence of each rule is computed, and if it is above the confidence threshold, the rule is satisfied [21],[22].

The Apriori algorithm is the most efficient algorithm for discovering association rules in large database. The pseudo code for Apriori algorithm is as follows:

C_k = Candidate itemset of size k
L_k = Frequent itemset of size k
L_1 ={Frequent items};
for (k = *1; L_k! = Ø; k++*) **do begin**
C_{k+1} = candidates generated from L_k

for each trancation *t* in database **do**
#increment the count of all candidates in C_{k+1} that are contained in *t*
L_{k+1} = candidates in C_{k+1} with minSupport
end

Return $U_k L_k$;

4 Experimental Results

The FDA's AERS datasets are taken from FDA's web site. The datasets in different time periods were combined into a single dataset and imported into Microsoft SQL Server 2012 database management system as database tables. Then, alendronate(fosamax) related records were selected to create a dataset for association analysis. In total, 9229 alendronate(fosamax) associated adverse event reports were collected from the whole database. Considering mostly seen adverse events, most frequent adverse events with alendronate were searched. Table 1 list the number of co-occurrences of adverse events with alendronate (fosamax) and Figure 2 shows the graphical representation of them. Based on the number of co-occurrences, fall has the highest frequency with alendronate, followed by osteonecrosis, pain, femur fracture, pneumonia, dyspnoea and anaemia in this order.

Table 2 also lists and Figure 3 shows the graphical representation of the number of co-occurrences of adverse event outcomes with alendronate(fosamax). Based on the number of co-occurrences, mostly seen outcome is OT(Other), followed by HO(Hospitalization), DS(Disability), LF(Life-Threatening), RI(Required Intervention to Prevent Permanent Impairment/Damage) and CA(Congenital Anomaly) in this order.

Normalisation is an important concern for studies based on data mining because some entities may have some variations such as synonyms and brand names. These names should be detected and normalized. In our study, alendroante has also some variations. For example, fosamax and adronat are the variations of alendronate. These variations were found by searching Drugbank database and mapped to one specific name. Drugbank is one of the biggest resource for drugs and currently contains >4100 drug entries, corresponding to >12000 different trade names and synonyms [23].

Table 1. The number of co-occurrences of adverse events with alendronate(fosamax)

Adverse event	N (the number of co-occurrences)
Fall	593
Osteonecrosis	556
Pain	553
Femur fracture	504
Pneumonia	498
Dyspnoea	473
Anaemia	463
Anxiety	457
Hypertension	448
Depression	446

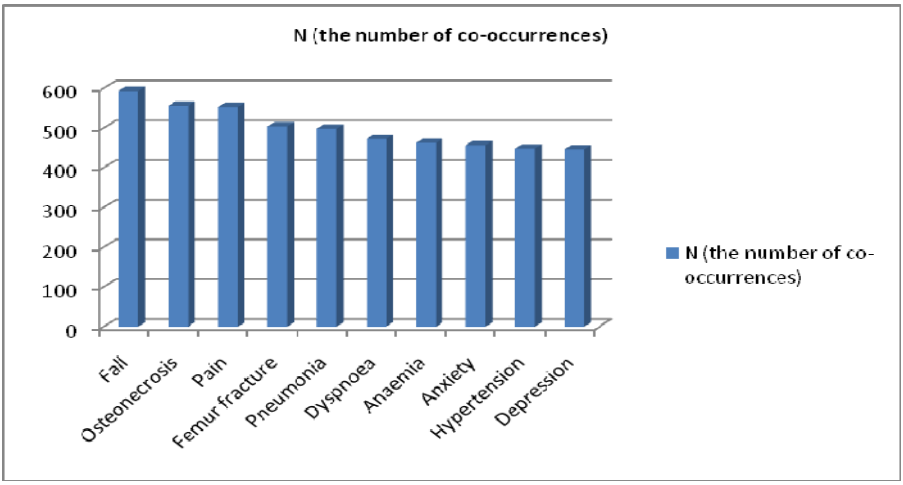

Fig. 2. The number of co-occurrences of adverse events with alendronate(fosamax)

Table 2. The number of co-occurrences of adverse event outcomes

Outcome	N(the number of co-occurrences)
OT(Other)	2821
HO(Hospitalization)	2745
DS(Disability)	1700
DE(Death)	1130
LF(Life-Threatening)	507
RI(Required Intervention to Prevent Permanent Impairment/Damage)	291
CA(Congenital Anomaly)	35

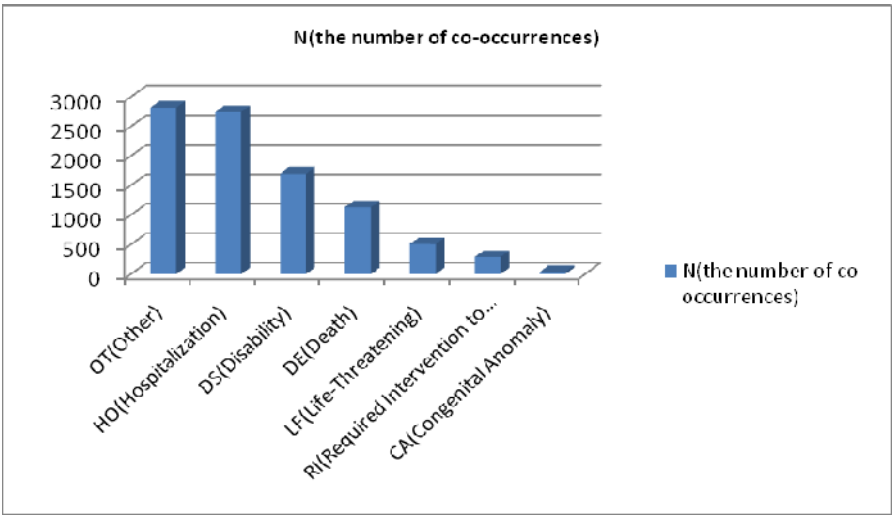

Fig. 3. The number of co-occurrences of adverse event outcomes

Table 3. Age Categories

Age	Category
0-6	Preschool child
7-12	Child
13-24	Young
25-43	Adult
44-64	Middle aged
65≤	Aged

Table 4. The number of reports by gender

Gender	The number of reports
Female	6149
Male	2821
NULL	249
Unknown	8
NS	2

After normalizing, the dataset was created and it contains patient demographics such as age, gender and adverse events. The attributes of the dataset were directly collected from database and age was categorized (Table 3). The dataset consists of three attributes: age, gender and the adverse events and 9229 instances.

The Apriori algorithm was used to perform association analysis on the characteristics of patient demographics and adverse events for alendronate(fosamax). Table 4 shows the number of reports by gender. WEKA 3.6.6 software was used. WEKA is a collection of machine learning algorithms for data mining tasks and is open source software. The software contains tools for data pre-processing, classification, regression, clustering, association rules and visualization [24], [25]. The application of the Apriori algorithm on the dataset generated many rules. Some rules are;

1. Age=44-64 Adverse Event=Arthralgia==> Gender=Female conf:(0.75)
2. Age≥65 Adverse Event = Femur fracture ==> Gender=Female conf:(0.74)
3. Age≥ 65 Adverse Event =Anaemia ==> Gender=Female conf:(0.73)
4. Adverse event=Impaired healing ==> Gender=Female conf:(0.71)
5. Age=44-64 Adverse Event=Diarrhoea ==> Gender=Female conf:(0.7)
6. Age≥65 Adverse Event=Hypertension ==> Gender=Female conf:(0.68)
7. Age≥ 65 Adverse Event =Nausea ==> Gender=Female conf:(0.72)
8. Age ≥65 Adverse Event=Depression ==> Gender=Female conf:(0.64)
9. Age=44-64 Adverse Event=Anxiety ==> Gender=Female conf:(0.63)
10. Age ≥65 Adverse Event=Osteonecrosis ==> Gender=Female conf:(0.63)

We also used the Predictive Apriori algorithm to compare the results of the Apriori algorithm. Similarly to the Apriori algorithm, the Predictive Apriori algorithm generates frequent item sets, but it uses a dynamically increasing minimum support threshold for the best rules concerning a support-based corrected confidence value. A rule is added is: the expected predictive accuracy of this rule is among the "n" best and it is not subsumed by a rule with at least the same expected predictive accuracy [26].The application of the Predictive Apriori algorithm on the dataset generated many rules. Some rules are;

1. Gender=Null Adverse Event=Depression 8 ==> Age ≥65 8 acc:(0.96017)
2. Age= 13-24 Adverse Event =Femur fracture 3 ==> Gender=Female 3 acc:(0.7783)
3. Age= 7-12 Adverse Event=Femur fracture 3 ==> Gender=Female 3 acc:(0.7783)
4. Age= 25-43 Adverse Event=Fall 30 ==> Gender=Female 26 acc:(0.71601)
5. Gender=Unkown Adverse Event=Pain 2 ==> Age= 44-64 2 acc:(0.69808)
6. Age= 13-24 Adverse Event=Depression 2 ==> Gender=Female 2 acc:(0.69808)
7. Age= 13-24 Adverse Event=Pneumonia 2 ==> Gender=Female 2 acc:(0.69808)
8. Age= 13-24 Adverse Event=Pain 2 ==> Gender=Male 2 acc:(0.69808)
9. Age ≥65 Adverse Event=Femur fracture 241 ==> Gender=Female 179 acc:(0.69727)
10. Age= 44-64 Adverse Event=Anaemia 229 ==> Gender=Female 167 acc:(0.68664)

The results of both Apriori and Predictive Apriori algoritms revealed some patterns in the dataset. According to the results of Apriori algorithm, some adverse events such as arthralgia, anaemia and diarrhoea have strong associations with middle aged patients (between 44 and 64). For example, rule 1 means that the possibility of arthralgia with middle aged and female patients is 75% (confidence). On the other hand, hypertension, nausea and depression are seen in patients over 65 years old.

Alendronate (fosamax) has some well known adverse events such as pain, femur fractures and nausea. The worst side effect is osteonecrosis of the jaw (ONJ), which is rare. Based on our results, osteonecrosis associated with patients over 65 years old. In addition, some events such as femur fracture, depression and anaemia are seen in the the rules which both Apriori and Predictive Apriori algorithms generated.

Medical researchers and clinicians can investigate our patterns and conduct clinical trials to discover new ideas for the assessment of drug safety of alendronate.

5 Discussion

The FDA's AERS database is considered an important resource, but some limitations were pointed out. First, the reports in the FDA's AERS contain errors, duplicate entries and missing data resulting in misclassifications. For example, patient age was not reported in many reports. Second, the structure of some datasets belonging to specific time periods are not compatible to the others. This means that reporting patterns and database structure changed over time. To overcome problems with data quality, we used some preprocessing methods. We omitted or corrected some records containing errors and missing data. For example, we omitted some records containing missing age or adverse events. Despite some limitations, the FDA's AERS database is a rich source to identify some important associations between drugs and adverse events [18].

Our study has both advantages and disadvantages. There are some web based tools to analyze the FDA's AERS database. The Drugcite processes the datasets between 2004 and 2012 and generates similar results with our study. We used this tool to find most ranked adverse events with alendronate(fosamax) and compared with our results. According to Drugcite, three most frequent adverse events are femur fracture, osteonecrosis and fall. On the other hand, our study highlighted that fall, osteonecrosis and pain are the top events. Because of some data quality problems with the datasets in different time periods above, we integrated the datasets which covers the third quarter of 2005 through the second of 2012. We think that this can result in some differences in both studies. Comparing our study and Drugcite tool, we found some statistical results for adverse events and outcomes with alendronate(fosamax) and implemented apriori algorithm to discover some association rules between patient demographics and adverse events. Drugcite is an useful tool which provides detail information about drugs and analyzes the FDA's AERS database, but it has no data mining functions. Therefore, we can conclude that our results are more informative and better than Drugcite's outputs.

Apart from the FDA's AERS database, medical records created in hospital information systems may be an important resource to determine drug adverse events and their outcomes. Therefore, a way to approach which includes both issues in the FDA's AERS and medical records can be used to reveal serious risks of a drug.

In this study, we found most ranked adverse events of alendronate(fosamax) and some relationships with patient demographics. Then, we searched some websites providing drug information such as http://www.nlm.nih.gov/medlineplus to compare our results. This comparison highlighted that some adverse events of alendronate(fosamax) are well known but some of them such as anxiety and depression have not been recognized. We also implemented both apriori and predictive apriori algorithms and discovered that anxiety has some associations with middle aged patients (between 44 and 64) and depression is seen over patients over 65 years old. This highlights that alendronate(fosamax) may cause psychological problems. We think that medical researchers should consider the results of our study and do clinical studies to determine the risks of the drug.

6 Conclusion

The aim of biomedical research is to discover new and useful knowledge to make contributions to better diagnosis, decision making and treatment [27],[28],[29]. The FDA's AERS is one of the big resources to reach this goal. In this study, we focused on knowledge discovery for alendronate(fosamax) drug and carried out a study based on patient demographics and adverse events relationships in the AERS reports. Alendronate(fosamax) is widely used for the treatment of osteoporosis and this disease is one of the major health problem in the world. Therefore, there is a big scientific interest in treatment and prevention of this disease.

We utilized database and computational techniques and then applied Apriori and Predictive Apriori algorithms to our dataset to obtain some rules. Our results show that some adverse events of alendronate(fosamax) have not known and patient demographics can have relationships with these events of the drug. Medical experts, researchers and pharmaceutical companies can explore and interprete these relationships. In conclusion, we believe that our study can make important contributions to postmarketing information in the assessment of drug safety of alendronate(fosamax).

Acknowledgements. We cordially thank Dr. Cinar Ceken for her help and medical expert advice. We also thank the SouthCHI'13 reviewers for their thorough review and helpful comments to further improve our paper.

References

[1] Boulton, G., Rawlins, M., Vallance, P., Walport, M.: Science as a public enterprise: the case for open data. The Lancet 377, 1633–1635 (2011)

[2] Rowen, L., Wong, G.K.S., Lane, R.P., Hood, L.: Intellectual property - Publication rights in the era of open data release policies. Science 289, 1881–1881 (2000)

[3] Thompson, M., Heneghan, C.: BMJ OPEN DATA CAMPAIGN We need to move the debate on open clinical trial data forward. British Medical Journal 345 (2012)
[4] Hersey, A., Senger, S., Overington, J.P.: Open data for drug discovery: learning from the biological community. Future Medicinal Chemistry 4, 1865–1867 (2012)
[5] Walport, M., Brest, P.: Sharing research data to improve public health. The Lancet 377, 537–539 (2011)
[6] El Emam, K., Arbuckle, L., Koru, G., Eze, B., Gaudette, L., Neri, E., Rose, S., Howard, J., Gluck, J.: De-identification methods for open health data: the case of the Heritage Health Prize claims dataset. J. Med. Internet Res. 14(1), 1–24 (2012)
[7] Hochberg, A.M., Reisinger, S.J., Pearson, R.K., O'Hara, D.J., Hall, K.: Using Data Mining to Predict Safety Actions From FDA Adverse Event Reporting System Data. Drug Information Journal 41, 633–643 (2007)
[8] Ahmad, S.R.: Adverse Drug Event Monitoring at the Food and Drug Administration. J. Gen. Intern. Med. 18(1), 57–60 (2003)
[9] Wysowski, D.K., Swartz, L.: Adverse Drug Event Surveillance and Drug Withdrawals in the United States, 1969-2002. Arch. Intern. Med. 165, 1363–1369 (2005)
[10] Hoblitzell, A., Mukhopadhyay, S., You, Q., et al.: Text mining for bone biology. In: HPDC 2010 (2010)
[11] Honeywell, M., Philips, S., Vo, K., et al.: Teriparatide for Osteoporosis: A Clinical Review. Drug Forecast 28(11), 713–716 (2003)
[12] Cummings, S.R., Melton, L.J.: Epidemiology and outcomes of osteoporotic fractures. Lancet 2002 359(9319), 1761–1767 (2002)
[13] Carmona, R.H.: Bone Health and Osteoporosis: A Report of the U.S. Surgeon General (2004)
[14] Tatonetti, N.P., Denny, J.C., Murphy, S.N., Fernald, G.H., Krishnan, G., Castro, V., Yue, P., Tsau, P.S., Kohane, I., Roden, D.M., Altman, R.B.: Detecting Drug Interactions From Adverse –Event Reports:Interaction Between Paroxetine and Pravastatin Increases Blood Glucose Levels. Clin. Pharmacol. Ther. 90(1), 133–142 (2011)
[15] Kadoyama, K., Kuwahara, A., Yamamori, M., Brown, J.B., Sakaeda, T., Okuno, Y.: Hypersensitivity Reactions to Anticancer Agents:Data Mining of the Public Version of the FDA Adverse Event Reporting System, AERS. Journal of Experimental & Clinical Cancer Research 30(93), 1–6 (2011)
[16] Tamura, T., Sakaeda, T., Kadoyama, K., Okuno, Y.: Aspirin and Clopidogrel-associated Bleeding Complications: Data Mining of the Public Version of the FDA Adverse Event Reporting System, AERS. International Journal of Medical Sciences 9(6), 441–446 (2012)
[17] Gandhi, P.K., Gentry, W.M., Bottorff, M.B.: Cardiovascular thromboembolic events associated with febuxostat:Investigation of cases from the FDA adverse event reporting system database. Seminars in Arthritis and Rheumatism 12, 1–5 (2012)
[18] Kadoyama, K., Miki, I., Tamura, T., Brown, J., Sakaeda, T., Okuno, Y.: Adverse Event Profiles of 5-Fluorouracil and Capecitabine: Data Mining of the Public Version of the FDA Adverse Event Reporting System, AERS, and Reproducibility of Clinical Observations. Int.J.Med.Sci. 9(1), 33–39 (2012)
[19] Kadoyama, K., Sakaeda, T., Tamon, A., Okuno, Y.: Adverse event Profile of Tigecycline:Data Mining of the Public Version of the U.S. Food and Drug Administraion Adverse Event Reporting System. Biol. Pharm. Bull. 35(6), 967–970 (2012)
[20] Han, J., Micheline, K.: Data mining: concepts and techniques. Morgan Kaufmann, San Francisco (2001)

[21] Zhu, A., Li, J., Leong, T.: Automated Knowledge Extraction for Decision Model Construction: A Data Mining Approach. In: AMIA 2003 Symposium Proceedings, pp. 758–762 (2003)
[22] Kim, J., Washio, T., Yamagishi, M., Yasumura, Y., Nakatani, S., Hashimura, K., Hanatani, A., Komamura, K., Miyatake, K., Kitamura, S., Tomoike, H., Kitakaze, M.: A Novel Data Mining Approach to the Identification of Effective Drugs or Combinations for Targeted Endpoints:Application to Chronic Heart Failure as a New Form of Evidence-based Medicine. Cardiovascular Drugs and Therapy 18, 483–489 (2004)
[23] Drugbank, http://www.drugbank.ca (last access: January 02, 2013)
[24] Hall, M., Frank, E., Holmes, G., Pfahringe, B., Reutemann, P., Witten, I.E.: The WEKA data mining software: an update. ACM SIGKDD Explorations Newsletter 11(1) (2009)
[25] WEKA: Weka 3: Data Mining Software in Java (last access: May 02, 2012), http://www.cs.waikato.ac.nz/ml/weka
[26] Scheffer, T.: Finding Association Rules That Trade Support Optimally against Confidence. In: Siebes, A., De Raedt, L. (eds.) PKDD 2001. LNCS (LNAI), vol. 2168, pp. 424–435. Springer, Heidelberg (2001)
[27] Holzinger, A., Simonic, K.M., Yildirim, P.: Disease-disease relationships for rheumatic diseases Web-based biomedical textmining and knowledge discovery to assist medical decision making. In: IEEE COMPSAC 36th International Conference on Computer Software and Application, pp. 573–580. IEEE, New York (2012)
[28] Holzinger, A., Yildirim, P., Geier, M., Simonic, K.-M.: Quality-based knowledge discovery from medical text on the Web Example of computational methods in Web intelligence. In: Pasi, G., Bordogna, G., Jain, L.C. (eds.) Qual. Issues in the Management of Web Information. ISRL, vol. 50, pp. 145–158. Springer, Heidelberg (2013)
[29] Holzinger, A., Scherer, R., Seeber, M., Wagner, J., Müller-Putz, G.: Computational Sensemaking on Examples of Knowledge Discovery from Neuroscience Data: Towards Enhancing Stroke Rehabilitation. In: Böhm, C., Khuri, S., Lhotská, L., Renda, M.E. (eds.) ITBAM 2012. LNCS, vol. 7451, pp. 166–168. Springer, Heidelberg (2012)