Review

Data Mining of the Public Version of the FDA Adverse Event Reporting System

Toshiyuki Sakaeda [1✉], Akiko Tamon [2], Kaori Kadoyama [1], and Yasushi Okuno [3✉]

1. Center for Integrative Education in Pharmacy and Pharmaceutical Sciences, Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto 606-8501, Japan;
2. Kyoto Constella Technologies Co., Ltd., Kyoto 604-8156, Japan;
3. Department of Systems Biosciences for Drug Discovery, Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto 606-8501, Japan.

✉ Corresponding author: Toshiyuki Sakaeda, Ph.D., Center for Integrative Education in Pharmacy and Pharmaceutical Sciences, Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto 606-8501, Japan, Tel: +81-75-753-9560, Fax: +81-75-753-9253, e-mail: sakaedat@pharm.kyoto-u.ac.jp; Yasushi Okuno, Ph.D., Department of Systems Biosciences for Drug Discovery, Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto 606-8501, Japan, Tel&Fax: +81-75-753-4559, e-mail: okuno@pharm.kyoto-u.ac.jp.

Abstract

The US Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS, formerly AERS) is a database that contains information on adverse event and medication error reports submitted to the FDA. Besides those from manufacturers, reports can be submitted from health care professionals and the public. The original system was started in 1969, but since the last major revision in 1997, reporting has markedly increased. Data mining algorithms have been developed for the quantitative detection of signals from such a large database, where a signal means a statistical association between a drug and an adverse event or a drug-associated adverse event, including the proportional reporting ratio (PRR), the reporting odds ratio (ROR), the information component (IC), and the empirical Bayes geometric mean (EBGM). A survey of our previous reports suggested that the ROR provided the highest number of signals, and the EBGM the lowest. Additionally, an analysis of warfarin-, aspirin- and clopidogrel-associated adverse events suggested that all EBGM-based signals were included in the PRR-based signals, and also in the IC- or ROR-based ones, and that the PRR- and IC-based signals were in the ROR-based ones. In this article, the latest information on this area is summarized for future pharmacoepidemiological studies and/or pharmacovigilance analyses.

Key words: adverse event; Adverse Event Reporting System; FAERS; database; data mining; signal; signal detection; proportional reporting ratio; reporting odds ratio; information component; empirical Bayes geometric mean; pharmacoepidemiology; pharmacovigilance.

Introduction

The US Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS, formerly AERS) is a database that contains information on adverse event and medication error reports submitted to the FDA [1-3]. The database is designed to support the FDA's post-marketing safety surveillance program for drug and therapeutic biologic products [1-3]. Its structure adheres to the international safety reporting guidance issued by the International Conference on Harmonisation, ICH E2B [1-3]. Adverse events and medication errors are coded using terms in the Medical Dictionary for Regulatory Activities (MedDRA) terminology [4]. The reports from manufacturers are either expedited, which must be submitted within 15 days, or periodic [1-3]. Reports can also be submitted by health care professionals and the public through

the "MedWatch" program [1-3]. The original system was started in 1969, but since the last major revision in 1997, reporting has markedly increased [5, 6]. To date, the FAERS is the largest repository of spontaneously reported adverse events in the world with more than 4 million reports [5, 6]. The FDA releases the data to the public, and public access offers the possibility to external researchers and/or pharmacovigilance experts to explore this data source, that is, allows conducting pharmacoepidemiological studies and/or pharmacovigilance analyses.

Pharmacoepidemiology is defined as the study of the use and effects of drugs in large numbers of people, whereas pharmacovigilance is the science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other drug-related problems [7-9]. The FAERS database is useful for analyzing associations with adverse events, but if used for pharmacovigilance, where early and timely detection is important [10-13], has two major deficiencies [14-16]: the lag time from collection to public release of the data, and the form in which the data are released, i.e., ASCII or SMGL files, which is not readily amenable to query or analysis [14-16]. Additionally, spontaneous reports can have missing data, and more than 2 reports can be submitted for one patient (duplicated reports) [17]. Through an attempt to address these shortcomings, a novel system, named the CzeekV system, has been developed by Dr. Okuno in collaboration with Kyoto Constella Technologies Co., Ltd., Japan, and we have employed to evaluate the safety profiles of several drugs [18-26]. Data mining algorithms are applied for the quantitative detection of signals [17, 27-30], where a signal means a statistical association between a drug and an adverse event or a drug-associated adverse event, including the proportional reporting ratio (PRR) [31], the reporting odds ratio (ROR) [32], the information component (IC) given by a Bayesian confidence propagation neural network [33], and the empirical Bayes geometric mean (EBGM) [34]. In this article, the 4 methods will be compared in terms of the signal contents, and the latest information on this area is summarized for future investigations.

FAERS database

A data set consists of 7 data tables; 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 drugs in the DRUG table are assigned as any of primary suspect, secondary suspect, concomitant, or interacting. The adverse events in the REAC table are coded using the preferred terms (PTs) in the MedDRA. It should be noted that there are a number of duplicated entries and the data occasionally contain misspelling and miswords. In our previous reports [18-26], input data were taken from the public release of the FAERS database from the first quarter of 2004 through the end of 2009, and the total number of reports used was 2,231,029. Prior to data mining, duplicated reports were deleted according to the FDA's recommendation of adopting the most recent CASE number, resulting in a reduction in the number of reports from 2,231,029 to 1,644,220. Several adverse events coded by PT terms and several drugs can be listed in a report. All drug names were unified into generic names by a text-mining approach, because FAERS permits the registering of arbitrary drug names, including trade names and abbreviations. Spelling errors were detected by a spell checker software, GNU Aspell, and carefully confirmed by working pharmacists. The total number of errors was 223,239. Foods, beverages, treatments (e.g. X-ray radiation), and unspecified names (e.g. beta-blockers) were omitted for this study, and the total number of omissions was 164,384. Consequently, a total of 22,017,956 co-occurrences were found in 1,644,220 reports, where a co-occurrence was a pair of a drug and an adverse event, and they were the basis for the signal detection. Of the 1,644,220 reports used for data mining, gender data were available for 1,520,994 (92.5%); 605,271 (36.8%) for males and 915,723 (55.7%) for females. Age data were available for 1,084,999 reports (66.0%), and the average (±SD) was 52.7±23.2 years.

Data mining algorithms

Data mining algorithms have been developed to identify drug-associated adverse events (signals) that are reported more frequently than expected by estimating expected reporting frequencies on the basis of information on all drugs and all events in the database [17, 27-30]. For example, PRR [31], ROR [32], IC [33], and EBGM [34] are widely used, and indeed, currently employed by the Medicines and Healthcare products Regulatory Agency (MHRA), UK, the Netherlands Pharmacovigilance Centre, the World Health Organization (WHO), and the FDA, respectively. All of these algorithms calculate signal scores, i.e., the values for PRR, ROR, IC, and EBGM, to assess whether a drug is significantly associated with an adverse event or not. These calculations or algorithms, so-called the disproportionality analyses or measures, however, differ from one another in that the PRR and ROR are frequentist (non-Bayesian), whereas the IC and EBGM are Bayesian.

A two-by-two contingency table is the framework for analysis (Table 1). The number of co-occurrences of interest is defined as n11. The number of co-occurrences with a drug of interest, but without an adverse of interest, is defined as n10, and the number n01 is assigned to those without a drug of interest, but with an adverse event of interest. The number of co-occurrences without either is as n00, and using them, the PRR and ROR are defined as:

$$PRR = [\ n11 \times (\ n01 + n00\ )\ ]\ /\ [\ n01 \times (\ n11 + n10\ )\ ]$$

$$ROR = (\ n11 \times n00\ )\ /\ (\ n10 \times n01\ )$$

The expected number of co-occurrences of interest, n11(expected), is defined as:

$$n11(expected) = [\ (\ n11 + n10\ ) \times (\ n11 + n01\ )\ ]\ /\ (\ n11 + n10 + n01 + n00\ )$$

The observed-to-expected ratio, n11/n11 (expected), is used for calculation of the IC and EBGM. The IC is a logarithmic metric of n11/n11 (expected) that is implemented in a Bayesian framework, and n11/n11 (expected) is known as the EBGM, when implemented within an empirical Bayesian framework.

In this section, only the scoring thresholds are given. The reader is referred to the articles for more extensive details on each statistical test [31-34]. Using the PRR, a signal is detected if the number of co-occurrences is 3 or more and the PRR is 2 or more with an associated χ^2 value of 4 or more [31]. For the ROR, a signal is detected, if the lower limit of the 95% two-sided confidence interval exceeds 1 [32]. Signal detection using the IC is done using the IC025 metric, a lower limit of the 95% two-sided confidence interval of the IC, and a signal is detected if the IC025 value exceeds 0 [33]. Finally, for the EBGM, the EB05 metric, a lower one-sided 95% confidence limit of the EBGM, is used and a signal is detected when the EB05 is greater than or equal to the threshold value 2.0 [34]. In our studies [18-26], these 4 methods were used to detect signals, and the adverse events were listed as drug-associated, when at least 1 of 4 indices met the criteria indicated above.

In 2003, the Pharmaceutical Research and Manufacturers of America-FDA Collaborative Working Group on Safety Evaluation Tools, consisting of statisticians, pharmacoepidemiologists, and pharmacovigilance professionals from the pharmaceutical industry and the FDA, reviewed the best practices for the use of these methods [28]. In summary, they stated that there is evidence that data mining may be useful, but the evidence is not sufficient to fully judge the value of data mining in pharmacovigilance, and that time and experience will reveal the value and utility

[28]. It is important to understand both the strengths and weaknesses of data mining algorithms to minimize their misapplication and misuse [27]. We recommend that readers refer to the many excellent review articles on data mining algorithms published previously [17, 27-30].

Table 1. A two-by-two contingency table for analysis.

	With an adverse event of interest	Without an adverse event of interest	Total
With a drug of interest	n11	n10	n11+n10
Without a drug of interest	n01	n00	n01+n00
Total	n11+n01	n10+n00	n11+n10+n01+n00

n11: the number of co-occurrences of interest. n11+n10: the total number of co-occurrences with a drug of interest. n11+n01: the total number of co-occurrences with an adverse event of interest. n11+n10+n01+n00: the total number of co-occurrences in the database.

Comparison of 4 data mining algorithms

In Table 2, an example of output data is provided to both familiarize readers with the analysis and show the differences between the 4 signal scores. These data are on the warfarin-, aspirin- and clopidogrel-associated haematemesis, and those on haemorrhage, haematoma, melaena and haematochezia have already published previously [26]. The total number of warfarin-associated adverse events was 736, and 848 for aspirin and 838 for clopidogrel [26]. According to the number of co-occurrences, haematemesis ranked 56th among 736 warfarin-associated adverse events, 13th among 848 aspirin-associated ones, and 61st among 838 clopidogrel-associated ones. The number of co-occurrences is an important index for monitoring the emergence of an adverse event, but is independent of the signal scores and/or performance of detection of signals. For example, haematemesis ranked 559th, 375th, and 581st, respectively, according to the ROR score. Moreover, it should be noted that whether an adverse event is detected as a signal or not depends on the algorithm used for signal detection. Indeed, as shown in Table 2, based on the ROR and IC, the analysis indicated the association of these 3 drugs with haematemesis; however, it was not associated with warfarin when the PRR was used, and the EBGM failed to detect the associations for warfarin and clopidogrel.

Table 2. Signal scores for warfarin-, aspirin- and clopidogrel-associated haematemesis.

Statins	N	PRR (kai2)	ROR (95% two-sided CI)	IC (95% two-sided CI)	EBGM (95% one-sided CI)
Warfarin	268	1.991 (131.982)	2.006 (1.778, 2.234) *	0.985 (0.811, 1.158) *	1.968 (1.778)
Aspirin	332	6.469 (1525.210) *	6.566 (5.889, 7.244) *	2.661 (2.504, 2.818) *	6.425 (5.864) *
Clopidogrel	235	2.254 (163.238) *	2.270 (1.995, 2.544) *	1.160 (0.975, 1.346) *	2.218 (1.991)

N: the number of co-occurrences; PRR: the proportional reporting ratio; ROR: the reporting odds ratio; IC: the information component; EBGM: the empirical Bayes geometric mean; CI: the confidence interval. The asterisk (*) means statistically significant association, i.e., the adverse events are detected as signals.The data are on the associations with haematemesis, and those on haemorrhage, haematoma, melaena and haematochezia have already published previously [26].

The relationships between the 4 algorithms and numbers of signals are summarized in Table 3. This table was constructed using the data in our previous reports [19, 20, 22-26]. Among the 4 methods, the ROR provided the highest number of signals, and the EBGM the lowest. The difference in the number of signals can be explained by a higher rate of false positives or a lower ability to detect the signals. Without standards, one cannot know which the case is. Additionally, an analysis was done here using the data on warfarin, aspirin and clopidogrel, and it was indicated that all EBGM-based signals were included in the PRR-based signals, and also in the IC- or ROR-based ones, and that the PRR- and IC-based signals were in the ROR-based ones. In other words, the ROR-based signals could be stratified into 5 groups; the signals detected by the ROR only, the signals detected by the ROR and PRR, the signals detected by the ROR and IC, the signals detected by the ROR, PRR, and IC, and the signals detected by the 4 methods. If this relationship will be confirmed for other drugs, it can be concluded that the EBGM is the most conservative algorithm.

Several studies have compared data mining algorithms [32, 35-40]; however, as Bate and Evans recently concluded [17], the different algorithms have slightly different properties and consequently one might be preferable in a particular application. If used for pharmacovigilance, the data mining algorithms should be assessed from the standpoint of early and timely signal detection [10-13]. There are limited published comparative data, but recently, Chen et al. compared the timing of early signal detection with the PRR, ROR, IC and EBGM using the FAERS database, and concluded that the ROR showed the better performance [10]. This does not conflict our finding mentioned above.

Several strategies have been proposed to increase the power to detect signals, and consequently to heighten pharmacovigilance, including integration with other databases [41] and removing already known drug-event associations [42]. Detecting signals not described in the prescribing information for a drug at the time of its approval, but supported by published evidence for an association (unlabeled supported signals), is important, and a high rate of detection of unlabeled supported signals is attained by data mining using higher level terms than the PT, i.e., the HLT or the Standard MedDRA Queries (SMQ), a group of PT terms [43]. New statistical methodologies will be continuously developed [44, 45], and therefore we will not be able to draw a conclusion concerning which is best for some time. Prior to discussion, it is important to elucidate the difference in the table of adverse events listed as drug-associated.

Table 3. Effects of 4 data mining algorithms on the number of signals (drug-associated adverse events).

	PRR	ROR	IC	EBGM
Cisplatin	479	884	430	175
Carboplatin	409	810	412	144
Oxaliplatin	410	732	406	150
Colistin	71	238	54	23
5-Fluorouracil	461	864	441	161
Capecitabine	411	802	395	146
Pravastatin	218	701	285	19
Simvastatin	192	744	399	30
Atorvastatin	284	883	514	55
Rosuvastatin	227	619	282	63
Tigecycline	91	248	75	44
Omeprazole	213	818	370	14
Esomeprazole	194	743	317	17
Warfarin	331	736	426	110
Aspirin	348	848	377	100
Clopidogrel	366	838	476	104

PRR: the proportional reporting ratio; ROR: the reporting odds ratio; IC: the information component; EBGM: the empirical Bayes geometric mean. This table was constructed using the data in our previous reports [19, 20, 22-26], and the permission of replication was obtained from the publishers.

Advantages of FAERS data mining

It is well-accepted that a randomized, prospective, large-scale and long-term clinical trial is the best way to assess the association between a drug and an adverse event; however, such trials are not practical

due to great expenses of time and cost, especially for rare but clinically important adverse events [46, 47]. Data mining of the FAERS database might provide previously unknown, but clinically important associations, and give us useful suggestions to guide clinical decision making. Additionally, the FAERS database might be a useful tool for pharmaceutical companies' post-marketing activities [48], and given an association between pre-marketing data and post-marketing signal [49], the database might provide constructive suggestions about the method of pre-marketing data collection.

Our studies suggested that FAERS data mining reproduced some well-established clinical associations, including cisplatin and nephrotoxicity [19], carboplatin and myelosuppression [19], oxaliplatin and peripheral sensory neuropathy [19], capecitabine and hand-foot syndrome [22], statins and muscular events [23], proton pump inhibitors and hypomagnesaemia [25], and antiplatelets and bleeding complications [26]. These results indicate the usefulness of the database and algorithms used, but do not certify an ability to provide previously unknown, but clinically important associations.

In 2003, an editorial comment concerning the efficacy and safety of a drug, published in a respected scientific journal, had a considerable impact, since it claimed that no reliable data were provided by a manufacturer [50]. Debate about this issue continued in the journal until Meyboom and Edwards concluded that there was a need to improve and accelerate pharmacovigilance [51]. The FAERS database relies on reports not only from manufacturers but also from health care professionals and the public, and one advantage is that the database includes end-user assessments.

The coding of adverse events and medication errors using the PT terms in the MedDRA has considerable advantages. For example, the terminology used to describe statin-associated muscular symptoms varies and therefore the incidence varies among reports [52, 53]. The National Lipids Association's Muscle Expert Panel and other statin experts have emphasized the importance of standardizing related terms to allow reliable comparisons among studies and to improve care for statin users [53]. This does not apply only to statins, and the employment of MedDRA ensures higher quality results.

In 2009, a pilot study performed by Hochberg et al. concerning drug-versus-drug comparisons found the rank-order of adverse event rates in the FAERS database to be consistent with the results of published studies, encouraging the use of the database for comparisons [54]. The number of reports with or without normalization by usage or sales during the corresponding period was used to compare drugs [55], but adverse events are underreported (discussed later), which might lead to incorrect conclusions [46, 47, 56, 57]. To date, we have no evidence that the signal scores can be used to determine the rank-order of drugs in terms of risk. Indeed, they were calculated for several drugs in the same class, but a discussion about the difference in susceptibility to adverse events was pending [58-61].

Limitations of FAERS data mining

In general, adverse events are underreported in spontaneous reporting systems [62-66]. The rate of reporting can vary with the particular adverse event [1], but averages just 6% [62]. Various factors can be determinants of underreporting, but the knowledge and attitude of health professionals seem to be most important [63]. Indeed, educational intervention was shown to improve the rate of reporting [64, 65]. A patient-targeted survey found that 87% of patients spoke to their physicians about a possible connection, but the physicians were more likely to exclude than affirm the possibility [66]. Pharmacists, nurse practitioners, and physician assistants play an important role [67, 68], and more publicity for the FAERS database and/or education should be considered to promote patient reporting [69]. Even though the reporting rate has dramatically improved, the FAERS database is still not appropriate for estimating incidence rates, due to the absence of a denominator [1].

The number of reports and signals, and the signal scores are influenced by various factors. The number of reports increases over the first 2 years after launching, and then starts going down [5, 6]. This is known as the Weber effect [70], although it is not always observed [71]. The number of signals and signal scores also possibly fluctuate during several years after launching, and the number of unlabeled supported signals depends on the time window after launching [72]. Generally speaking, reporting can be accelerated after a drug-associated adverse event is highlighted [73-75], and this is known as the notoriety effect [74]. Additionally, the notoriety for a drug can accelerate the reporting of other drugs in the same class, known as the ripple effect [74]. In contrast, signal scores can be suppressed by a large number of reports in which the same adverse event is connected with other drugs [76]. This is the masking or cloaking effect [76]. Close attention should be paid to the results of signal detection, especially for drugs launched only recently, and it is important to investigate a temporal axis, when planning a pharmacovigilance analysis [73].

It should be noted that there is no credible counterfactual means, e.g., a randomized control group, to list signals, and therefore disease-oriented adverse events can be listed also. For example, 238 colistin-associated adverse events included sepsis, pseudomonas infection, *Acinetobacter* infection, and influenza-like illness [20]. Generally, the results can be biased by unmeasured confounding factors; e.g., the association of a drug with an adverse event might be explained by those of other drugs which are often co-administered. Although a comparison of drugs in the same class can offset confounding factors, enabling drug-versus-drug comparisons, a statistically well-organized methodology should be established to minimize their effects.

The data mining algorithms are designed to identify bivariate associations, and the possibility that an adverse event occurred synergistically with more than 2 drugs is excluded. Harpaz et al. have recently analyzed the possibility of multi-item adverse event associations, i.e., associations relating multiple drugs to possibly multiple adverse events [77]. Multi-item associations are rarely reported but are important because they can indicate drug-drug interactions, and a total of 1,167 multi-item associations were identified using 162,744 reports [77]. Signal scores after stratification by the presence or absence of co-administration might provide information about the drug-drug interaction.

Closing remarks

The FAERS database has been used to analyze the safety profiles of various drugs. Additionally, the highly suspicious drugs to induce some serious adverse events are listed by analyzing the FAERS database, including torsades de pointes [78, 79] and Stevens-Johnson syndrome/toxic epidermal necrolysis [80]. Several organizations maintain their own well-organized databases of spontaneously reported adverse events, and use them to analyze associations with drugs. The WHO Programme for International Drug Monitoring, starting in 1968, is a cornerstone of world-wide pharmacovigilance, and the Uppsala Monitoring Centre, Uppsala, Sweden, maintains the WHO Global Individual Case Safety Report Database, VigiBase [81-84]. In France, pharmacovigilance activities are carried out by 31 regional centers based in clinical pharmacology departments of university hospitals [85, 86], and the French Pharmacovigilance Database has been used to analyze the safety profiles of various drugs [87-91]. Additionally, two European databases, that is, the General Practitioners Research Database (GPRD) in the UK, and the PHARMO Record Linkage System in the Netherlands, are also frequently used for analytical studies [92]. In the near future, a device or infrastructure enabling real-time analysis at a doctor's office or by a patient will be provided by information technology service companies.

Data mining does not provide sufficient evidence on causality, and merely suggests the necessity for well-organized clinical studies with respect to associations. The WHO defines a "signal" as "reported information on a possible causal relationship between an adverse event and a drug, the relationship being unknown or incompletely documented previously" [7, 8]; however, considerable ambiguity remains in the definition in reports [93-95]. Poluzzi et al. used the FAERS database to list drugs liable to induce torsades de pointes, and thereafter authorized reports were used as references to stratify the signals into expected and unexpected signals [78, 79]. Torsades de pointes is considered a designated medical event, i.e., a low-probability event with drug-attributed risk, and a case-by-case analysis is of primary importance [40]. As for unexpected signals, they emphasized the necessity for further investigation and close surveillance [78, 79].

A debate recently published in a respected journal indicates both the advantages and limitations of data mining of spontaneously reported adverse event databases [96, 97]. de Boer emphasized that the disproportionality measures as used in spontaneous reporting databases have important limitations and more advanced way might generate new relevant knowledge worth publishing [96]. In contrast, Montastruc et al. commented that none of the methods (e.g., case-control studies and cohort studies), if taking alone, should be considered as definitive for evaluating drug risk, and disproportionality studies appear to be important today, due to a growing demand for safer drugs [97]. A report in the FAERS database is a story, sometimes only a rumor, but numerous reports can reflect reality. With larger numbers of faithful reports, the FAERS database and other spontaneously reported databases should help to optimize pharmacotherapy.

Acknowledgments

Competing Interests

The authors have declared that no competing interest exists.

References

1. Rodriguez EM, Staffa JA, Graham DJ. The role of databases in drug postmarketing surveillance. Pharmacoepidemiol Drug Saf. 2001; 10: 407-410.
2. Wysowski DK, Swartz L. Adverse drug event surveillance and drug withdrawals in the United States, 1969-2002: the importance of reporting suspected reactions. Arch Intern Med. 2005; 165: 1363-1369.
3. [Internet] U.S. Food and Drug Administration (FDA). http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/default.htm
4. [Internet] MedDRA MSSO. http://www.meddramsso.com/index.asp
5. Moore TJ, Cohen MR, Furberg CD. Serious adverse drug events reported to the Food and Drug Administration, 1998-2005. Arch Intern Med. 2007; 167: 1752-1759.
6. Weiss-Smith S, Deshpande G, Chung S, et al. The FDA drug safety surveillance program: adverse event reporting trends. Arch Intern Med. 2011; 171: 591-593.
7. World Health Organization (WHO). The importance of pharmacovigilance: Safety monitoring of medicinal products. 2002.
8. Lindquist M. The need for definitions in pharmacovigilance. Drug Saf. 2007; 30: 825-830.
9. Lindquist M. Data quality management in pharmacovigilance. Drug Saf. 2004; 27: 857-870.
10. Chen Y, Guo JJ, Steinbuch M, et al. Comparison of sensitivity and timing of early signal detection of four frequently used signal detection methods: An empirical study based on the US FDA Adverse Event Reporting System database. Pharm Med. 2008; 22: 359-365.
11. Motola D, Piccinni C, Biagi C, et al. Cardiovascular, ocular and bone adverse reactions associated with thiazolidinediones: a disproportionality analysis of the US FDA adverse event reporting system database. Drug Saf. 2012; 35: 315-323.
12. Sommet A, Grolleau S, Bagheri H, et al. Was the thrombotic risk of rofecoxib predictable from the French Pharmacovigilance Database before 30 September 2004? Eur J Clin Pharmacol. 2008; 64: 829-834.
13. Lindquist M, Ståhl M, Bate A, et al. A retrospective evaluation of a data mining approach to aid finding new adverse drug reaction signals in the WHO international database. Drug Saf. 2000; 23: 533-542.
14. [No authors listed]. Making a difference. Nat Biotechnol. 2009; 27: 297.
15. Pratt LA, Danese PN. More eyeballs on AERS. Nat Biotechnol. 2009; 27: 601-602.
16. Böhm R, Höcker J, Cascorbi I, et al. OpenVigil - free eyeballs on AERS pharmacovigilance data. Nat Biotechnol. 2012; 30: 137-138.
17. Bate A, Evans SJ. Quantitative signal detection using spontaneous ADR reporting. Pharmacoepidemiol Drug Saf. 2009; 18: 427-436.
18. Sakaeda T, Kadoyama K, Yabuuchi H, et al. Platinum agent-induced hypersensitivity reactions: Data mining of the public version of the FDA Adverse Event Reporting System, AERS. Int J Med Sci. 2011; 8: 332-338.
19. Sakaeda T, Kadoyama K, Okuno Y. Adverse event profiles of platinum agents: Data mining of the public version of the FDA Adverse Event Reporting System, AERS, and reproducibility of clinical observations. Int J Med Sci. 2011; 8: 487-491.
20. Sakaeda T, Kadoyama K, Tamon A, et al. Data mining of the public version of the FDA Adverse Event Reporting System, AERS: Colistin-associated adverse events. Jpn J Chemother. 2011; 59: 610-613.
21. Kadoyama K, Kuwahara A, Yamamori M, et al. Hypersensitivity reactions to anticancer agents: Data mining of the public version of the FDA Adverse Event Reporting System, AERS. J Exp Clin Cancer Res. 2011; 30: 93.
22. Kadoyama K, Miki I, Tamura T, et al. 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. 2012; 9: 33-39.
23. Sakaeda T, Kadoyama K, Okuno Y. Statin-associated muscular and renal adverse events: Data mining of the public version of the FDA Adverse Event Reporting System. PLoS ONE. 2011; 6: e28124.
24. Kadoyama K, Sakaeda T, Tamon A, et al. Adverse event profile of tigecycline: Data mining of the public version of the US Food and Drug Administrtation Adverse Event Reporting System. Biol Pharm Bull. 2012; 35: 967-970.
25. Tamura T, Sakaeda T, Kadoyama K, et al. Omeprazole- and esomeprazole-associated hypomagnesaemia: Data mining of the public version of the FDA Adverse Event Reporting System. Int J Med Sci. 2012; 9: 322-326.
26. Tamura T, Sakaeda T, Kadoyama K, et al. Aspirin- and clopidogrel-associated bleeding complications: Data mining of the public version of the FDA Adverse Event Reporting System. Int J Med Sci. 2012; 9: 441-446.
27. Hauben M, Reich L. Drug-induced pancreatitis: lessons in data mining. Br J Clin Pharmacol. 2004; 58: 560-562.
28. Almenoff J, Tonning JM, Gould AL, et al. Perspectives on the use of data mining in pharmaco-vigilance. Drug Saf. 2005; 28: 981-1007.
29. Almenoff JS, Pattishall EN, Gibbs TG, et al. Novel statistical tools for monitoring the safety of marketed drugs. Clin Pharmacol Ther. 2007; 82: 157-166.
30. Hauben M, Bate A. Decision support methods for the detection of adverse events in post-marketing data. Drug Discov Today. 2009; 14: 343-357.
31. Evans SJ, Waller PC, Davis S. Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. Pharmacoepidemiol Drug Saf. 2001; 10: 483-486.
32. van Puijenbroek EP, Bate A, Leufkens HG, et al. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. Pharmacoepidemiol Drug Saf. 2002; 11: 3-10.
33. Bate A, Lindquist M, Edwards IR, et al. A Bayesian neural network method for adverse drug reaction signal generation. Eur J Clin Pharmacol. 1998; 54: 315-321.
34. Szarfman A, Machado SG, O'Neill RT. Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the US FDA's spontaneous reports database. Drug Saf. 2002; 25: 381-392.
35. Gould AL. Practical pharmacovigilance analysis strategies. Pharmacoepidemiol Drug Saf. 2003; 12: 559-574.
36. Rothman KJ, Lanes S, Sacks ST. The reporting odds ratio and its advantages over the proportional reporting ratio. Pharmacoepidemiol Drug Saf. 2004; 13: 519-523.
37. Waller P, van Puijenbroek E, Egberts A, et al. The reporting odds ratio versus the proportional reporting ratio: 'deuce'. Pharmacoepidemiol Drug Saf. 2004; 13: 525-526.
38. Matsushita Y, Kuroda Y, Niwa S, et al. Criteria revision and performance comparison of three methods of signal detection applied to the spontaneous reporting database of a pharmaceutical manufacturer. Drug Saf. 2007; 30: 715-726.
39. Hochberg AM, Hauben M, Pearson RK, et al. An evaluation of three signal-detection algorithms using a highly inclusive reference event database. Drug Saf. 2009; 32: 509-525.
40. Hochberg AM, Hauben M. Time-to-signal comparison for drug safety data-mining algorithms vs. traditional signaling criteria. Clin Pharmacol Ther. 2009; 85: 600-606.
41. Hammond IW, Gibbs TG, Seifert HA, et al. Database size and power to detect safety signals in pharmacovigilance. Expert Opin Drug Saf. 2007; 6: 713-721.
42. Pariente A, Didailler M, Avillach P, et al. A potential competition bias in the detection of safety signals from spontaneous reporting databases. Pharmacoepidemiol Drug Saf. 2010; 19: 1166-1171.
43. Pearson RK, Hauben M, Goldsmith DI, et al. Influence of the MedDRA hierarchy on pharmacovigilance data mining results. Int J Med Inform. 2009; 78: e97-e103.
44. Gibbons RD, Segawa E, Karabatsos G, et al. Mixed-effects Poisson regression analysis of adverse event reports: the relationship between antidepressants and suicide. Stat Med. 2008; 27: 1814-1833.
45. Harpaz R, Perez H, Chase HS, et al. Biclustering of adverse drug events in the FDA's spontaneous reporting system. Clin Pharmacol Ther. 2011; 89: 243-250.
46. Chertow GM, Mason PD, Vaage-Nilsen O, et al. On the relative safety of parenteral iron formulations. Nephrol Dial Transplant. 2004; 19: 1571-1575.
47. Chertow GM, Mason PD, Vaage-Nilsen O, et al. Update on adverse drug events associated with parenteral iron. Nephrol Dial Transplant. 2006; 21: 378-382.
48. Bailey S, Singh A, Azadian R, et al. Prospective data mining of six products in the US FDA Adverse Event Reporting System: disposition of events identified and impact on product safety profiles. Drug Saf. 2010; 33: 139-146.
49. Moylan CA, Suzuki A, Papay JI, et al. A pre-marketing ALT signal predicts post-marketing liver safety. Regul Toxicol Pharmacol. 2012; 63: 433-439.
50. Editorial. The statin wars: why AstraZeneca must retreat. Lancet. 2003; 362: 1341.
51. Meyboom RH, Edwards IR. Rosuvastatin and the statin wars--the way to peace. Lancet. 2004; 364: 1997-1999.
52. Harper CR, Jacobson TA. Evidence-based management of statin myopathy. Curr Atheroscler Rep. 2010; 12: 322-330.

53. Di Stasi SL, MacLeod TD, Winters JD, et al. Effects of statins on skeletal muscle: a perspective for physical therapists. Phys Ther. 2010; 90: 1530-1542.

54. Hochberg AM, Pearson RK, O'Hara DJ, et al. Drug-versus-drug adverse event rate comparisons: a pilot study based on data from the US FDA Adverse Event Reporting System. Drug Saf. 2009; 32: 137-146.

55. Bailie GR. Comparison of rates of reported adverse events associated with i.v. iron products in the United States. Am J Health Syst Pharm. 2012; 69: 310-320.

56. Wysowski DK, Swartz L, Borders-Hemphill BV, et al. Use of parenteral iron products and serious anaphylactic-type reactions. Am J Hematol. 2010; 85: 650-654.

57. Auerbach M, Kane RC. Caution in making inferences from FDA's Adverse Event Reporting System. Am J Health Syst Pharm. 2012; 69: 922-923.

58. Piccinni C, Motola D, Marchesini G, et al. Assessing the association of pioglitazone use and bladder cancer through drug adverse event reporting. Diabetes Care. 2011; 34: 1369-1371.

59. Schmedt N, Andersohn F, Garbe E. Signals of progressive multifocal leukoencephalopathy for immunosuppressants: a disproportionality analysis of spontaneous reports within the US Adverse Event Reporting System (AERS). Pharmacoepidemiol Drug Saf. 2012; 21: 1216-1220.

60. Colman E, Szarfman A, Wyeth J, et al. An evaluation of a data mining signal for amyotrophic lateral sclerosis and statins detected in FDA's spontaneous adverse event reporting system. Pharmacoepidemiol Drug Saf. 2008; 17: 1068-1076.

61. Knudsen JF, Sokol GH, Flowers CM. Adjunctive topiramate enhances the risk of hypothermia associated with valproic acid therapy. J Clin Pharm Ther. 2008; 33: 513-519.

62. Hazell L, Shakir SA. Under-reporting of adverse drug reactions : a systematic review. Drug Saf. 2006; 29: 385-396.

63. Lopez-Gonzalez E, Herdeiro MT, Figueiras A. Determinants of under-reporting of adverse drug reactions: a systematic review. Drug Saf. 2009; 32: 19-31.

64. Figueiras A, Herdeiro MT, Polónia J, et al. An educational intervention to improve physician reporting of adverse drug reactions: a cluster-randomized controlled trial. JAMA. 2006; 296: 1086-1093.

65. Herdeiro MT, Polónia J, Gestal-Otero JJ, et al. Improving the reporting of adverse drug reactions: a cluster-randomized trial among pharmacists in Portugal. Drug Saf. 2008; 31: 335-344.

66. Golomb BA, McGraw JJ, Evans MA, et al. Physician response to patient reports of adverse drug effects: implications for patient-targeted adverse effect surveillance. Drug Saf. 2007; 30: 669-675.

67. Ehrenpreis ED, Sifuentes H, Ehrenpreis JE, et al. Suboptimal reporting of adverse medical events to the FDA Adverse Events Reporting System by nurse practitioners and physician assistants. Expert Opin Drug Saf. 2012; 11: 177-183.

68. Gavaza P, Brown CM, Lawson KA, et al. Examination of pharmacists' intention to report serious adverse drug events (ADEs) to the FDA using the theory of planned behavior. Res Social Adm Pharm. 2011; 7: 369-382.

69. Du D, Goldsmith J, Aikin KJ, et al. Despite 2007 law requiring FDA hotline to be included in print drug ads, reporting of adverse events by consumers still low. Health Aff (Millwood). 2012; 31: 1022-1029.

70. Hartnell NR, Wilson JP. Replication of the Weber effect using postmarketing adverse event reports voluntarily submitted to the United States Food and Drug Administration. Pharmacotherapy. 2004; 24: 743-749.

71. McAdams MA, Governale LA, Swartz L, et al. Identifying patterns of adverse event reporting for four members of the angiotensin II receptor blockers class of drugs: revisiting the Weber effect. Pharmacoepidemiol Drug Saf. 2008; 17: 882-889.

72. Hochberg AM, Hauben M, Pearson RK, et al. Systematic investigation of time windows for adverse event data mining for recently approved drugs. J Clin Pharmacol. 2009; 49: 626-633.

73. Raschi E, Piccinni C, Poluzzi E, et al. The association of pancreatitis with antidiabetic drug use: gaining insight through the FDA pharmacovigilance database. Acta Diabetol. 2011; in press.

74. Pariente A, Gregoire F, Fourrier-Reglat A, et al. Impact of safety alerts on measures of disproportionality in spontaneous reporting databases: the notoriety bias. Drug Saf. 2007; 30: 891-898.

75. Hauben M, Reich L, Gerrits CM. Reports of hyperkalemia after publication of RALES-a pharmacovigilance study. Pharmacoepidemiol Drug Saf. 2006; 15: 775-783.

76. Wang HW, Hochberg AM, Pearson RK, et al. An experimental investigation of masking in the US FDA adverse event reporting system database. Drug Saf. 2010; 33: 1117-1133.

77. Harpaz R, Chase HS, Friedman C. Mining multi-item drug adverse effect associations in spontaneous reporting systems. BMC Bioinformatics. 2010; 11(Suppl 9): S7.

78. Poluzzi E, Raschi E, Moretti U, et al. Drug-induced torsades de pointes: data mining of the public version of the FDA Adverse Event Reporting System (AERS). Pharmacoepidemiol Drug Saf. 2009; 18: 512-518.

79. Poluzzi E, Raschi E, Motola D, et al. Antimicrobials and the risk of torsades de pointes: the contribution from data mining of the US FDA Adverse Event Reporting System. Drug Saf. 2010; 33: 303-314.

80. Papay J, Yuen N, Powell G, et al. Spontaneous adverse event reports of Stevens-Johnson syndrome/toxic epidermal necrolysis: detecting associations with medications. Pharmacoepidemiol Drug Saf. 2012; 21: 289-296.

81. Lindquist M. VigiBase, the WHO global ICSR database system: basic facts. Drug Inf J. 2008; 42: 409-419.

82. Giezen TJ, Mantel-Teeuwisse AK, Meyboom RH, et al. Mapping the safety profile of biologicals: a disproportionality analysis using the WHO adverse drug reaction database, VigiBase. Drug Saf. 2010; 33: 865-878.

83. Strandell J, Caster O, Bate A, et al. Reporting patterns indicative of adverse drug interactions: a systematic evaluation in VigiBase. Drug Saf. 2011; 34: 253-266.

84. de Jong HJ, Saldi SR, Klungel OH, et al. Statin-associated polymyalgia rheumatica. An analysis using WHO global individual case safety database: a case/non-case approach. PLoS One. 2012; 7: e41289.

85. Montastruc JL, Sommet A, Lacroix I, et al. Pharmacovigilance for evaluating adverse drug reactions: value, organization, and methods. Joint Bone Spine. 2006; 73: 629-632.

86. Moore N, Noblet C, Kreft-Jais C, et al. French pharmacovigilance database system: examples of utilization. Therapie. 1995; 50: 557-562.

87. Grandvuillemin A, Disson-Dautriche A, Miremont-Salamé G, et al. Cetuximab infusion reactions: French pharmacovigilance database analysis. J Oncol Pharm Pract. 2012; in press.

88. Bondon-Guitton E, Bagheri H, Montastruc JL. Drug-induced gingival overgrowth: a study in the French Pharmacovigilance Database. J Clin Periodontol. 2012; 9: 513-518.

89. Taugourdeau-Raymond S, Rouby F, Default A, et al. Bevacizumab-induced serious side-effects: a review of the French pharmacovigilance database. Eur J Clin Pharmacol. 2012; 68: 1103-1107.

90. Moulis G, Sommet A, Durrieu G, et al. Trends of reporting of 'serious'vs. 'non-serious' adverse drug reactions over time: a study in the French PharmacoVigilance Database. Br J Clin Pharmacol. 2012; 74: 201-204.

91. Vodovar D, LeBeller C, Mégarbane B, et al. Drug Fever: a descriptive cohort study from the French national pharmacovigilance database. Drug Saf. 2012; 35: 759-767.

92. Härmark L, van Grootheest AC. Pharmacovigilance: methods, recent developments and future perspectives. Eur J Clin Pharmacol. 2008; 64: 743-752.

93. Hauben M, Reich L. Communication of findings in pharmacovigilance: use of the term "signal" and the need for precision in its use. Eur J Clin Pharmacol. 2005; 61: 479-480.

94. Hauben M, Aronson JK. Defining 'signal' and its subtypes in pharmacovigilance based on a systematic review of previous definitions. Drug Saf. 2009; 32: 99-110.

95. Edwards IR, Lindquist M. First, catch your signal! Drug Saf. 2010; 33: 257-260.

96. de Boer A. When to publish measures of disproportionality derived from spontaneous reporting databases? Br J Clin Pharmacol. 2011; 72: 909-911.

97. Montastruc JL, Sommet A, Bagheri H, et al. Benefits and strengths of the disproportionality analysis for identification of adverse drug reactions in a pharmacovigilance database. Br J Clin Pharmacol. 2011; 72: 905-908.