We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Data Mining Techniques in Pharmacovigilance: Analysis of the Publicly Accessible FDA Adverse Event Reporting System (AERS)

Elisabetta Poluzzi, Emanuel Raschi, Carlo Piccinni and Fabrizio De Ponti

Additional information is available at the end of the chapter

1. Introduction

1.1. Data mining in a clinical pharmacology perspective

Drug use in medicine is based on a balance between expected benefits (already investigated before marketing authorization) and possible risks (i.e., adverse effects), which become fully apparent only as time goes by after marketing authorization. Clinical pharmacology deals with the risk/benefit assessment of medicines as therapeutic tools. This can be done at two levels: the individual level, which deals with appropriate drug prescription to a given patient in everyday clinical care and the population level, which takes advantage of epidemiological tools and strategies to obtain answers from previous experience. The two levels are intertwined and cover complementary functions.

Data mining has gained an important role during all stages of drug development, from drug discovery to post-marketing surveillance. Whereas drug discovery is probably the first step in drug development that resorts to data mining to exploit large chemical and biological databases to identify molecules of medical interest, in this chapter data mining will be considered within the context of long-term drug safety surveillance after marketing authorisation.

A pharmacological background is essential before considering data mining as a tool to answer questions related to the risk/benefit assessment of drugs. As a first step, it must be verified whether or not the available sources of data (e.g. spontaneous reporting systems, claim databases, electronic medical records, see below) are the most appropriate to address the research question. In other words, a prior hypothesis is required and one should

INTECH
open science | open minds

consider which tool is the best option for the specific aim. In addition, the actual impact in clinical practice of any research question depends on the communication and dissemination strategies and relevant indicators to evaluate this impact should be developed as well.

The use of data mining techniques in clinical pharmacology can be broadly grouped into two main areas, each with specific aims:

1. **identification of new effects of drugs** (mostly adverse reactions, but sometimes also new therapeutic effects, and effects in special populations);
2. **appropriateness in drug use** (e.g., frequency of use in patients with contraindications, concomitant prescriptions of drugs known for the risk of clinically relevant interactions).

Both aims can be addressed using each of the three conventional sources of data listed below, although the inherent purpose for which they are created should be kept in mind when interpreting results: any secondary analysis of data collected for other purposes carries intrinsic biases.

- **Spontaneous reporting systems** (SRS) are mostly addressed to identify adverse reactions. Virtually anywhere in the world, notification of adverse drug events is mandatory for health professionals, but also other subjects can report events to the relevant regulatory Authorities. Main Drug Agencies routinely use algorithms of data-mining to process data periodically and to find possible unknown drug-effect associations. These algorithms identify drug-reaction pairs occurring with a significant disproportion in comparison with all other pairs. Clinical pharmacology knowledge is then requested to interpret those signals and to decide if further examination is needed (either within the same source of data or by other type of data) or specific bias affects the validity of the signal. More detailed description on this specific strategy will be provided in the next paragraphs. This source of data is currently the most frequently approached with data mining in pharmacovigilance [1], because of its usefulness and the ready availability of information. In some limited cases, also aim 2 above can be addressed by this source: the detailed analysis of patient related risk factors (demographic characteristics, concomitant disorders or medications) can show *foci* of lack of appropriateness in the use of specific drugs included in adverse event reports. No inference on the incidence of the adverse event among patients exposed to a specific drug or with risk factors can be performed.
- **Electronic medical records** (namely, patient registries) are mainly collected with the aim to assist physicians in daily appropriate prescription. For each subject, these registries usually include information on socio-demographic characteristics, diagnoses, risk factors, treatments and outcomes. Primary care is the most frequent setting for the development of this kind of registry and many authoritative examples exist: GPRD (General Practice Research Database), HEALTH SEARCH, THIN (The Health Improvement Network), IPCI (Interdisciplinary Processing of Clinical Information). Hospital examples are also important as they clearly cover complementary therapeutic areas. The high quantity of information and the high quality of data included in these

tools makes them valuable sources for data mining aimed to address clinical pharmacology questions, both in terms of new effects of drugs (especially on primary endpoints, to confirm premarketing evidence) and of assessment of appropriate drug use (closer to the main purpose of the registries).

• **Claim databases** originate for administrative purposes: for instance, in many European countries, health care costs are provided by the National Health Service to Local Health Authorities on the basis of the costs of the medical interventions provided to citizens. Claim databases include all data useful to this purpose (e.g., diagnoses of hospital admissions, reimbursed prescriptions of drugs, diagnostic procedures in ambulatory care) and, as a secondary aim, they represent an important source for epidemiological questions. These data can be equally useful for both aims (1 and 2 above), provided that their intrinsic limitations are duly acknowledged, in particular it should be recognized that information on outcomes are not strictly related to drug use (namely, adverse drug reactions) and patho-physiological plausibility supporting drug-reaction associations should be more stringently verified.

Among assessments of appropriate drug use, there is growing interest in the study of drug-drug interactions, which are usually dealt with by analyzing claim databases searching for specific drug-drug pairs known to interact with clinically relevant consequences. The final aim of this strategy is to compare the frequency of such drug-drug pairs in different settings, or in different periods of time for the same setting, to quantify the risk at population level and the impact of specific educational interventions. Also SRSs can be mined for drug-drug interactions, but in this case the aim is usually quite different: specifically assessing the actual contribution of drug-drug interaction in the occurrence of defined ADRs [2,3]. Virtually, this last step (possible association between specific drug-drug interaction and adverse events) should be performed first in the drug evaluation process; once the risk of a given drug-drug interaction is documented and the percentage of patients developing the ADR is quantified (by using causal association methodologies), the risk at the population level can be evaluated for clinical and regulatory purposes.

After identification of the most appropriate source of data to address the research question, the appropriate methodology to approach data analysis should be identified. From data cleaning (a mere data managing step, see below) to statistical methodologies (e.g., multiple regression analysis), all steps of data management are considered parts of data mining techniques. Usually, each source of data is analyzed by the own natural data mining approach (e.g., disproportion calculation for SRSs, multiple regression analysis for electronic medical records), but emergent strategies to better exploit the more accessible sources are now appearing in the biomedical literature (e.g., self-controlled time series [4]). In fact, data mining could virtually provide as many associations as possible between drug and effect, but, without a consensus among experts on the methodological steps and a confirmation of patho-physiological pathways, the association can easily conduct to interpretation errors. For instance, the risk of pancreatitis by antidiabetics is currently a matter of debate and recently Elashoff et al. [5] claimed a 6-fold increase of pancreatitis in subjects exposed to exenatide. This raised criticisms by several authors [6,7]. Elashoff et al., indeed, tried to

approach SRSs by a specific strategy for analytical studies (case-control analysis), disregarding the absence of pure control subjects (see below).

Text-mining [8] and information mining [9] are also frequently used in searching possible associations between drugs and adverse events, although this is virtually ignored by regulatory agencies. Especially in this case, the source of data is a key step for a reliable result of the analysis: both free-text search into electronic patient records and text analysis of any document freely published online represent possible sources of data for these strategies. Because of the huge variety of information processed by this approach, a more strict plausibility of the associations found between drug and effect should be claimed, because of, for instance, the high risk of inverse causality.

2. Definition and objectives of data mining in pharmacovigilance

Before addressing methodological issues, we provide an overview on the potential of data mining in the field of pharmacovigilance.

2.1. The need for pharmacovigilance

Pharmacovigilance (PhV) has been defined by the World Health Organization (WHO) as "the science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other possible drug-related problems" [10]. In the past, it was regarded as being synonymous with post-marketing surveillance for adverse drug reactions. However, the concept of surveillance in pharmacovigilance and pharmacoepidemiology has evolved from the concept of surveillance in epidemiology. Surveillance and monitoring are different - surveillance involves populations, while monitoring involves individuals. Surveillance reflects real-world pharmacovigilance processes and has been recently defined "a form of non-interventional public health research, consisting of a set of processes for the continued systematic collection, compilation, interrogation, analysis, and interpretation of data on benefits and harms (including relevant spontaneous reports, electronic medical records, and experimental data)" [11].

The specific aims of pharmacovigilance are:

- the identification of previously unrecognized Adverse Drug Reactions (ADRs, novel by virtue of nature, frequency and/or severity);
- the identification of subgroups of patients at particular risk of adverse reactions;
- the continued surveillance of a product throughout the duration of its use, to ensure that the balance of its benefits and harms are and remain acceptable;
- the description of the comparative adverse reactions profile of products within the same therapeutic class;
- the detection of inappropriate prescription and administration;
- the further elucidation of a product's pharmacological and toxicological properties and the mechanism(s) by which it produces adverse effects;

- the detection of clinically important drug–drug, drug–herb/herbal medicine, drug–food, and drug–device interactions;
- the communication of appropriate information to health-care professionals;
- the confirmation or refutation of false-positive signals that arise, whether in the professional or lay media, or from spontaneous reports (see below for definition of a "signal") [12].

Therefore, PhV is becoming a holistic discipline embracing overall risk/benefit assessment, which necessarily takes place both at the individual patient level and at the population level (epidemiological perspective). It is now recognized as a key proactive approach for appropriate drug prescription and rational use of medicines, as it provides the tools to prevent, detect, monitor and counteract adverse drug reactions, by perfectly complementing the branch of pharmaco-epidemiology. A complete taxonomy of terminologies and definitions used in PhV (including medication errors) is beyond the aim of this chapter, as it has been extensively addressed by a number of Authors [10,13]. Indeed, the European Medicines Agency (EMA) has recently adopted the new **pharmacovigilance legislation** (**Regulation (EU) No 1235/2010** and **Directive 2010/84/EU**), approved by the European Parliament and European Council in December 2010 [14,15]. This legislation is the biggest change to the regulation of human medicines in the European Union (EU) since 1995 and is now in force as of July 2012. Great emphasis is explicitly mentioned on the role of PhV to *"promote and protect public health by reducing burden of ADRs and optimizing the use of medicines"*.

2.2. Pharmacovigilance tools: An overview

All the objective of PhV can be achieved through different types of studies, which are intended to be either "hypothesis-generating" or "hypothesis-testing" or to share these aims. The former are represented by spontaneous reporting and prescription event monitoring with the aim of identifying unexpected ADRs, whereas the latter are represented by case-control or cohort studies, which aim to prove (by risk quantification) whether any previous suspicious that has been raised is justified. Although rules and regulations governing the spontaneous reporting of ADRs vary among Countries, there are some basic commonalities: 1) with the exception of pharmaceutical companies that are legally "forced" to report ADRs to health authorities, it is an independent voluntary notification of the reporter (e.g., healthcare professional, patient); 2) it is sufficient that a suspicion arises (no certainty is required); 3) there must be an identifiable drug, patient, and event; 4) the total number of exposed to the drug and the total number that experiences/did not experience an event are unknown. In other words, the exact numerator and denominator, a pre-requisite for quantifying risk, are unavailable.

Although spontaneous reports are placed at the bottom of the hierarchy of the evidence, they represent a timely source in the early evaluation of safety issues associated with specific drugs, in particular for new compounds. Clinical trials are considered as the best source of evidence, although they suffer of several methodological issues (e.g., limited

sample size, reduced follow-up, evaluation of surrogate markers) that undermine their fully external validity. By contrast, the analysis of spontaneous reports reflects the real-world scenario, in which patients experienced the primary outcome in a complex pharmacological context. This underlines the importance of a high-quality analysis according to standard procedure and strengthens the need for transparency among different data miners (i.e. independent Research Organization, Universities, Regulators, Manufacturers) to provide the best outcome for patients. It is only the convergence of proofs which allows final conclusions and decisions in pharmacovigilance. Thus, the notion of 'levels of evidence', widely used for evaluating drug efficacy, cannot be actually applied in the field of ADRs; all methods are of interest in the evaluation of ADRs [16].

Spontaneous reports are collected by drug companies and at regional, national and international level through different databases. The Eudravigilance database (European Union Drug Regulating Authorities Pharmacovigilance) is held by the European Medicines Agency and collects/exchanges electronically ADRs coming from national regulatory authorities, marketing authorization holders and sponsors of interventional clinical trials and non-interventional studies in Europe. The Uppsala Monitoring Centre in Sweden is responsible for the worldwide gathering of all serious ADRs received by regulatory authorities and companies. The FDA Adverse Event Reporting System (FDA_AERS) collects ADRs from the US as well as rare and severe events from Europe. Unfortunately, only a minority of spontaneous reporting systems offers a public access to external researchers/consumers. These include the FDA_AERS and the Canada Vigilance Adverse Reaction Online Database [17]. At national level, several regional centers support the national Agency by publicly posting a list of potential signals. For instance, the Netherlands Pharmacovigilance Centre Lareb provides a quarterly newsletter service including new emerging signals [18].

Although SRSs have been criticized due to inherent pitfalls, a recent survey found that the Population-based reporting ratio (PBRR, defined as the total number of ADR reports collected in a safety database per year per million inhabitants) revealed an increased reporting activity at national level (half of the European Countries exceed the standard threshold of 300 for PBRR, indicating a sufficient quality for signal detection), with strong correlation with the relative increase at the international level [19].

While the analysis of multiple databases has the advantage to cover a very large population and heterogeneous patterns of reporting for ADRs, it is important to note that data cannot be simply aggregated/pooled to perform DMAs for the following reasons:

- differences exists among databases, especially in terms of accessibility, drug and adverse event coding.
- overlap are likely to exist among national and international archives, especially for rare and serious ADRs. The precise extent of overlap has not been assessed yet and may vary depending on the ADR under investigation.

Therefore, combination of different databases has been proposed to achieve the highest statistical power in signal detection [20].

2.3. The term "signal" in pharmacovigilance: An evolving concept

The major aim of PhV is signal detection (i.e., the identification of potential drug-event association that may be novel by virtue of their nature, severity and/or frequency). In the literature, the term signal has been subject to debate and discussion. Some authors attempted to provide a comprehensive overview of definitions that have been proposed and implemented over the years. The WHO has defined 'signal' as ''Reported information on a possible causal relationship between an adverse event and a drug, the relationship being unknown or incompletely documented.'' The definition is not self-contained, being followed by a qualifier that stipulates that ''usually more than one case is required'' [10]. This qualifier is important, because it implies that the information content of the signal must sufficiently reduce uncertainty to justify some action, and therefore partially mitigates the former limitation. This should be part of the definition. Meyboom et al. [21] stated that ''*A signal in pharmacovigilance is more than just a statistical association. It consists of a hypothesis together with data and arguments, arguments in favor and against the hypothesis. These relate to numbers of cases, statistics, clinical medicine, pharmacology (kinetics, actions, previous knowledge) and epidemiology, and may also refer to findings with an experimental character.''* It is important to note that a single report or a few reports can sometimes constitute a signal of suspected causality. This is the case of well documented, high-quality reports with positive re-challenge. These so-called anecdotes provide a definitive ('between-the-eyes') drug-related reactions, which do not necessarily need further formal verification [22,23]. In PhV, anecdotes have been also technically defined as "Designated Medical Event" (DME), which is a rare but serious reactions with high drug-attributable risk (i.e., a significant proportion of the occurrences of these events are drug induced) [24]. Stevens-Johnson Syndrome/Toxic Epidermal Necrolysis (SJS/TEN) and Torsade de Pointes (TdP) are typical example of DMEs, although a formal and universally accepted list of DMEs does not exist. The most balanced way to analyze DMEs is to use disproportionality in conjunction with a case-by-case evaluation, in order to capture all possible information from spontaneous reports [25]. However, in most of the cases, information that arises from one or multiple sources (including observations and experiments) should be taken into account for causality assessment [26]. Indeed, after a suspicion has been raised (signal generation), it needs to be corroborated by the accumulation of additional data (signal strengthening). The final step involves the confirmation and quantification of the relation between drug and ADR by epidemiological methods (signal quantification). Although in every step of the process both individual case reports and analytical techniques are involved, the relative importance of these approaches differs per step [26]. Once a signal is verified, a number of regulatory actions may be considered, depending on the level of priority assigned to the signal. This prioritization process is mainly based on the assessment of public health impact, the severity of the adverse event and the strength of disproportionality. Regulatory measures may range from close monitoring of the signal over time (without practical interventions, the so-called watchful waiting) to drug withdrawal from the market due to a negative risk/benefit profile.

2.4. The need for Data Mining Algorithms (DMAs)

Currently, the "art" of signal detection can be performed on a case-by-case basis (traditional approach, especially for DMEs) or through automated procedures to support the clinical evaluation of spontaneous reports, the so-called "data mining approach". In general terms, data mining can be considered an activity related to "knowledge discovery in databases", i.e., the process of extracting information form a large database [27]. In this context, data mining is refereed to as the computer-assisted procedures, starting from processing of dataset by data "cleaning" and culminating into the application of statistical techniques, often known as data mining algorithms (DMAs). DMAs are currently and routinely used by pharmacovigilance experts for quantitative signal detection [28]. The purposes of quantitative signal detection are many-fold and may vary depending on the local habit of PhV experts. For instance, DMAs can be used as an aid to the traditional case-by-case assessment; as a screening tool to periodically generate a list of signals requiring in depth investigation (i.e., to prioritize signals); on *ad hoc* basis to detect complex data dependencies, which are difficult to be manually detected (e.g., drug-drug interactions or drug-related syndromes) [29].

2.5. The use of DMAs

Although DMAs are relatively new as compared to clinical trials and epidemiological studies, the first published attempt to assess the extent of reporting in drug safety through disproportionality, to the best of our knowledge, was by Bruno Stricker [30]. In 1997, the study by Moore et al., [31] introduced the concept of *"case/non case method"* when performing pharmacovigilance analyses. They postulated that this term was a more effective representation than case-control: controls are not actual controls since they are all exposed to at least one drug, and have at least one event (there are no untreated 'healthy' controls). They are simply not cases of the event of interest: they are cases of something else [32].

The accuracy of data mining techniques has been already tested retrospectively to determine if already known safety issues would have been detected 'earlier' [33]. However, it is generally difficult to determine when a known safety concern was first detected. Moreover, the surrogate endpoint that have been used (e.g., the date of implementation of new labeling) is unlikely to truly represent the time of first detection of a new safety signal, thus affecting the results in favor of DMAs. Overall, DMAs often provided a high level of accuracy in terms of timely prediction of risk and, therefore their use have been encouraged as an early source of information on drug safety, particularly new drugs, thus guiding the proper planning of subsequent observational studies [34].

Although the rationale and the methodology of the various approaches differ, all DMAs query databases for disproportionality and express the extent to which the reported ADR is associated with the suspected drug compared with all other drugs (or a subgroup of drugs) in the database. The reporting of ADRs related to other drugs in the database is used as a *proxy* for the background occurrence of ADRs. In other words, they assess whether

statistically significant differences in the reporting exist among drugs (the so-called "unexpectedness") and provide an answer to the question: "does the number of observed cases exceed the number of expected cases?". It is important to underline that, although these approaches are known as "quantitative" signal detection methodologies, no risk quantification can be assessed. Moreover, the presence of a statistically significant result does not necessarily imply an actual causal relationship between the ADR and the drug, nor does the absence of a statistically significant result necessarily disprove the possible relationship. As a matter of fact, the term "signal of disproportionate reporting" has been suggested by Hauben & Aronson [26] to emphasize the uncertainty in causality assessment.

DMAs can be classified in frequentist and Bayesian approach. Among the former, the Reporting Odds Ratio (ROR) is applied by the Netherlands Pharmacovigilance Centre Lareb, whereas the Proportional Reporting Ratio (PRR) was first used by Evans et al. [35]. Bayesian methods such as Multi-item Gamma Poisson Shrinker (MGPS) [36] and Bayesian Confidence Propagation Neural network (BCPN) [37] are based on Bayes' law to estimate the probability (posterior probability) that the suspected event occurs given the use of suspect drug.

Frequentist or classical methods are particularly appealing and therefore widely used due to the fact that they are relatively easy to understand, interpret and compute as they are based on the same principles of calculation using the 2x2 table (see **Table 1** with **Figure 1**).

	Drug of interest	**All other drugs in the database**	**Total**
Adverse drug reaction of interest	A	B	A+B
All other adverse drug reaction	C	D	C+D
Total	A+C	B+D	A+B+C+D

A= number of reports containing both the suspect drug and the suspect adverse drug reaction
B= number of reports containing the suspect adverse drug reaction with other medications (except the drug of interest)
C= number of reports containing the suspect drug with other adverse drug reactions (except the event of interest)
D= number of reports containing other medications and other adverse drug reactions

Table 1. Formal 2x2 contingency table

The PRR involves the calculation of the rate of reporting of one specific event among all events for a given drug, the comparator being this reporting rate for all drugs present in the database (including the drug of interest). Usually, a disproportion is considered on the basis of three pieces of information: PRR≥2, χ2≥4 and at least 3 cases [35]. The ROR is the ratio of the odds of reporting of one specific event versus all other events for a given drug compared to this reporting odds for all other drugs present in the database. A signal is considered when the lower limit of the 95% confidence interval (CI) of the ROR is greater than one. Basically, the higher the value, the stronger the disproportion appears to be. In both these methods, for a given drug, the particular risk of an event being reported versus other events is compared to a reference risk: that observed for all drugs in the database for PRR and for all other drugs for ROR. This reference risk is thought to reflect the baseline risk of the reporting of an event in a subject taking a drug, provided that there is no specific association

between the drug and the event of interest. The question arises as to whether this reference risk always provides an accurate estimate of the baseline risk of an event for a patient receiving any drug, since it is obtained by considering a reference group of drugs that can include drugs known to be at risk for a particular event. This aspect will be specifically addressed in the following section of the chapter. Rothman et al., [38] proposed to treat spontaneous reporting system as a data source for a case-control study, thus excluding from the control series those events that may be related to drug exposure; therefore, the ROR may theoretically offer an advantage over PRR by estimating the relative risk. However, this apparent superiority has been called into question, because both disproportionality measures do not allow risk quantification, but only offer a rough indication of the strength of the signal [39].
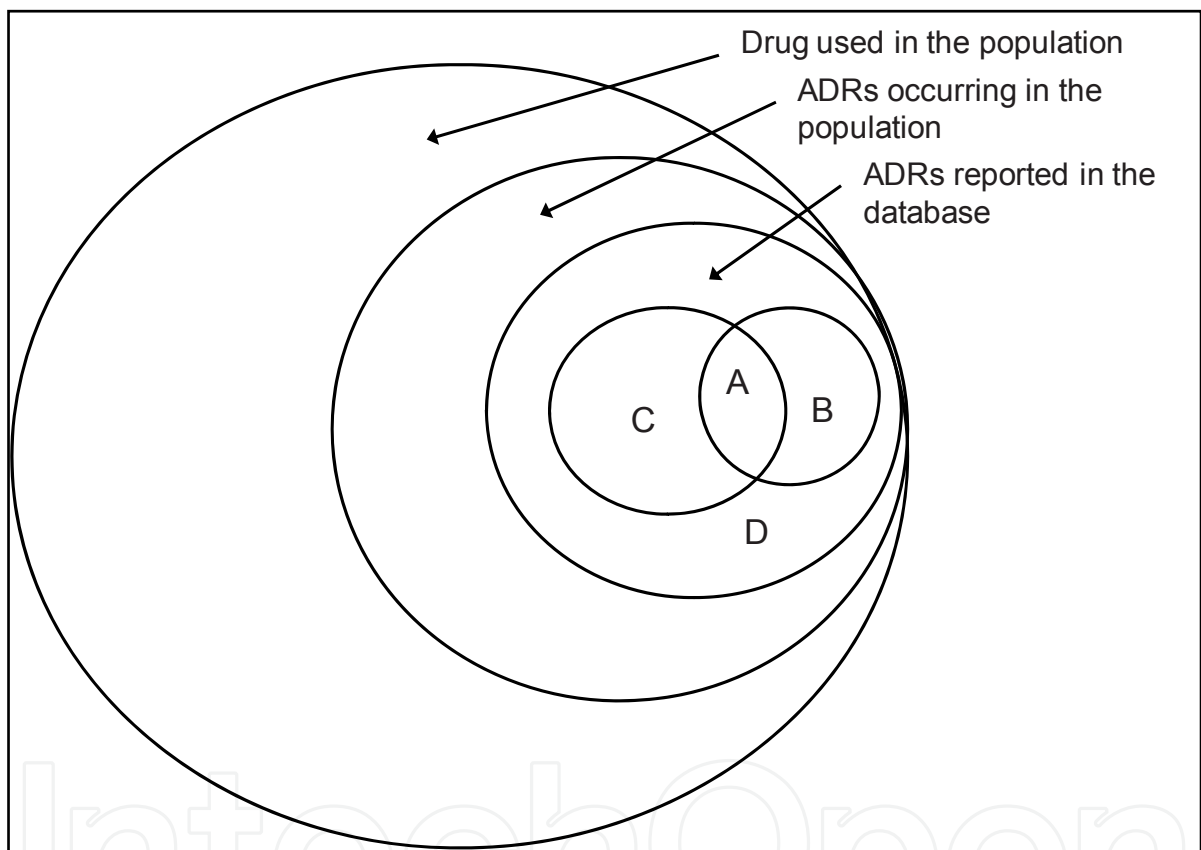

Figure 1. Relationships among patients exposed to drugs, experiencing ADRs and actual reporting (used to calculate disproportionality). Areas of circles do not necessarily reflect the precise proportion and relationships existing among circles.

As compared to the view of "frequency probability", Bayesian methods interpret the concept of probability as the degree to which a person believes a proposition. Bayesian inference starts with a pre-existing subjective personal assessment of the unknown parameter and the probability distribution (called prior distribution). They are based on the Bayes' law, assuming that there are two events of interest (D and E), which are not independent. From the basic theory of probability, it is known that the conditional probability of E given that D has occurred is represented as P(E/D)=P(E,D)/P(D), where P(D)=probability of a suspected drug being reported in a case report;

P(E)= probability of a suspected event being reported in a case report;

P(E,D)= probability that suspected drug and event being simultaneously reported in a case report;

P(E/D)= probability that suspected event being reported given the suspected drug being reported;

Assuming that the probability that D and E simultaneously occur is the same as the probability that D and E occur and rearranging the formula, we have P(E/D)=P(E,D)/P(D)=P(E)P(D/E)/P(D), which is Bayes' law.

The signal metric or signal score in BCPNN is the information component (IC) = $\log_2$ P(E,D)/P(E)P(D). If drug and event are statistically independent, the ratio of the joint probability of drug and event [P(E,D)] to the product of the individual probabilities [P(E)P(D)] will equal 1 and the IC will equal zero. The use of the logarithm of the ratio is derived from information theory. The elementary properties of logarithms make them suitable for quantifying information. The IC can be conceptualized as the additional information obtained on the probability of the event (or the additional uncertainty eliminated) by specifying a drug. Separate prior probabilities are constructed to represent the possible unconditional probabilities of drug [P(D)], event [P(E)] and the joint probability of drug and event [P(E,D)]. Uninformed priors (meaning all probabilities are weighed equally in the absence of data) are used for the unconditional probabilities.

The parameters of the joint probability distribution are then selected to make the limit of the IC approach zero for very low cell counts. As data accumulate, the influence of these priors will diminish more for cells with high counts. This means that, for drugs or events for which the counts are low, the signal score is shrunk more toward the prior probabilities based on statistical independence. A signal usually requires that the lower 95% CI of the IC exceed zero [40]. A comprehensive review of the principles subtending calculation of Bayesian methods is beyond the aim of this chapter and the reader is referred to Hauben & Zhou [41] for sophisticated, yet intuitive discussion of this issue.

Although the discussion on performance, accuracy and reliability of different DMAs is fascinating, actually, there is no recognized *gold standard* methodology. Therefore, several studies have been conducted to examine and compare the performance of different DMAs, in terms of sensitivity, specificity, accuracy and early identification of safety issues. A number of investigations explored whether differences exist between frequentist and bayesian approaches and found that PRR is more sensitive than MGPS, although the estimation from the MGPS is believed to be more robust when the number of reports is small [42-45]. While, to the best of our knowledge, no studies has been conducted to compare the MGPS and BCPN, van Puijenbroek et al., [46] first attempted to compare four DMAs ROR, PRR, Yule's Q and χ2 with IC. Since the IC is not the *gold standard*, these comparisons may have affected the findings, leading to an overestimation of the sensitivity. Notably, a high level of concordance was found, especially when the number of reports exceeds four. Kubota et al., [47] using a Japanese SRS, analyzed 38,731 drug-event combinations (DECs) and found a highly variable percentage of detected DEC among DMAs, ranging from 6.9% (GPS) to 54.6% (ROR). However, a misclassification bias may

have affected the results. It is clear, however, that the volume of signals generated by itself is an inadequate criterion for comparison and that the clinical nature of events and differential timing of signals needs to be considered [48]. In this context, a recent pilot study by Chen et al., [49] performed on the FDA_AERS, showed a better performance of ROR in terms of timing of early signal detection as compared to other DMAs, when tested on ten confirmed DEC. The issue of timely detection is of utmost importance in PhV, because early detection of safety-related problems may trigger signal substantiation and quantification through other pharmacoepidemiological research.

Apart from sensitivity, specificity, positive and negative predictive value, the possibility for correcting for covariates should be taken into account. Due to the nature of data source (mainly focused on drug-related information rather than patient's characteristics), the reporting of specific concomitant drugs may be used as a *proxy* for underlying disease, which may act as confounder. The use of these agents may be therefore considered in the multivariate regression analysis to calculate an adjusted disproportionality [50]. The ROR is a transparent and easily applicable technique, which allows adjustment through logistic regression analysis. An additional advantage of using ROR is the fact that non-selective underreporting of a drug or ADR has no influence on the value of the ROR compared with the population of patients experiencing an ADR [51]. For these reasons, our Research Unit of Clinical and Experimental Pharmacology at the University of Bologna at present uses the ROR with 95%CI to calculate disproportionality (data provided in the following section of the chapter are presented through this DMA). An overview on the most frequently used DMAs is provided in **Table 2** to summarize operative information for the reader.

The arbitrary nature of threshold criteria for signal detection could cause the identification of potential false positive or false negative associations. A recent review of published threshold criteria for defining signals of disproportionate reporting highlighted a considerable variation in defining a significant disproportionality among practitioners of pharmacovigilance data miners. For instance, ROR-1.96 Standard Error>1 may be used instead of 95%CI. The impact of this change is actually unexplored [52]. Indeed, changing the thresholds or selecting DMAs based on sensitivity considerations alone can have major implications: a more stringent criterion increases the sensitivity of the test by lowering the number of false positives, with the risk of missing credible signals. It is necessary to find an optimum balance, not just with regard to the use of statistics (frequentist *vs* Bayesian) but also among thresholds used for signal detection. Without a clinical evaluation of the signals, it is unclear how signal volume relates to signal value – the ability to identify real and clinically important problems earlier than they would be identified using current pharmacovigilance methods. In the wake of this challenging task, data mining methods based on false discovery rates have recently been proposed with promising results [53]. In addition, novel data mining techniques are emerging, such as those based on biclustering paradigm, which is designed to identify drug groups that share a common set of adverse events in SRS [54]. Another emerging issue pertains the identification of drug-drug interaction in a pharmacovigilance database. Very recently, a new three-way disproportionality measure (Omega, based on the total number of reports on two drugs and one ADR together) was developed within the WHO_Vigibase [55,56].

DMA	Computation	Published threshold criteria	Advantage	Limitations	Regulatory Agencies
Bayesian methods					
Multi-item Gamma Poisson Shrinker (MGPS) [36]	$$\frac{a\,(a+b+c+d)}{(a+c)\,(a+b)}$$	EBGM$_{05}$ > 2 N>0	Always applicable More specific as compared to frequentist method*	Relatively non-transparent for people non familiar with Bayesian statistics Lower sensitivity	FDA (AERS)
Bayesian Confidence Propagation Neural network (BCPN) [37]	$$\log_2 \frac{a\,(a+b+c+d)}{(a+c)\,(a+b)}$$	IC-2 SD>0	Always applicable More specific as compared to frequentist method* Can be used for pattern recognition in higher dimension	Relatively non-transparent for people non familiar with Bayesian statistics Lower sensitivity	UMC (WHO-Vigibase)
Frequentist methods					
Proportional Reporting Ratio (PRR) [35]	$$\frac{a/(a+b)}{c/(c+d)}$$ 95%CI=$e^{\ln(PRR)\pm1.96}$ $$\sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}$$	PRR≥2, χ^2≥4, N≥3	Easily applicable Easily interpretable More sensitive as compared to Bayesian method*	Cannot be calculated for all drug-event combinations Lower specificity	EMA (Eudravigilance) Italian Regulatory Agency (AIFA)
Reporting Odds Ratio (ROR) [46]	$$\frac{a/c}{b/d}$$ 95%CI=$e^{\ln(ROR)\pm1.96}$ $$\sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}$$	95%CI> 1, N≥2	Easily applicable Easily interpretable More sensitive as compared to Bayesian method* Different adjustment for covariates in logistic regression analysis	Odds ratio not calculated if denominator is zero (specific ADRs) Lower specificity	Lareb (Netherlands)

* when commonly cited thresholds are used.

χ^2= chi-squared; N= number of cases; IC= Information Component; CI=Confidence Interval; SD=Standard Deviation; EBGM=Empirical Bayesian Geometric Mean; $_{05}$=fifth percentile of the posterior distribution, i.e., there is a 95% probability that the "true" relative reporting ratio exceeds the EBGM$_{05}$

Table 2. Summary of major DMAs used for signal detection.

2.6. DMAs: Current debate

In the literature, there is debate on the advantage of using the number of reports instead of sales/prescriptions as denominator for signaling approach. We believe that both methods are useful, but should be used according to the research objective. In particular, the use of drug utilization data is of utmost importance to calculate reporting rates for drugs with already known association with the event, thus estimating the lowest incidence (assuming the under-reporting is limited due to the notoriety of the ADR or at least equally distributed within the database).

Recently, the publication of PhV analyses through disproportionality measures has been subject to debate and criticism [57]. Some benefits and strengths of using DMAs are undisputed, since they are quick and inexpensive analyses routinely performed by regulators and researchers for drug safety evaluation [16]. Apart from the hypothesis-generating purpose of signal detection, other important application of this method are (a) validation of a pharmacological hypothesis about the mechanism of occurrence of ADRs [58]; (b) characterization of the safety profile of drugs [59]. Nevertheless, it is important to underline that the identification of potential safety issues does not necessarily imply the need for wider communication to healthcare professional through publication, as it may cause unnecessary alarm without a real safety alert. The number of papers on *case/non case* evaluation are increasing exponentially and calls for minimum requirements before publication in the medical literature. This would ensure dissemination of high quality studies, which offer innovative methodology or provide novel insight into drug safety. The transparency policy recently adopted by the FDA is important to share with pharmacovigilance experts current safety issues requiring close monitoring before publicly disseminating results to consumers. Likewise, disproportionality analyses submitted for publication to relevant journals should address (and optimistically try to circumvent) bias related to selective reporting, in order to provide meaningful comparison among drugs and allow provisional risk stratification.

2.7. DMAs: Caveats

When planning a pharmacovigilance analysis and discussing its results, there are a number of limitations requiring careful consideration in view of the potential clinical implications. These caveats are related both to data source, namely individual reports with relevant spontaneous reporting system, and the adopted DMA [48,60]. Underreporting is one the most important limitation on SRS as it prevents to precisely calculate the real incidence of the event in the population [61]. As a matter of fact, only a small proportion of the ADRs occurring in daily practice is reported.

Substantial deficits in data quality and data distortion occur at two levels in SRS databases: at the level of the individual case records (i.e., quality and completeness of the reported information) and at the level of the overall sample (e.g., the framework of ADRs, which may vary depending on local rules for reporting). The FDA, for example, requires the reporting of events that occur in other countries only if they are serious and unlabeled based on the

US label. Hence a drug for which a serious adverse event has been included in the label earlier will appear to have fewer such events in the FDA AERS database than one for which the event was added later. The geographic distribution of market penetration of a new drug may also influence the apparent safety profile based on the SRS. For example, a drug whose use is predominantly in the U.S. may appear to have a very different profile than a drug whose use is predominantly in Europe, with the apparent difference potentially being due to differences in reporting expectations and behavior rather than real differences in the effects of the drugs. At present, there are also several limitations regarding the use of public-release version of the FDA_AERS. Apart from the 6-month lag time in data release through the FDA website, the most significant caveats concern the presence of duplicate reports and missing data as well as the lack of standardization in recording drug names of active substances. All these technical issues must be considered when exploring the FDA_AERS and will be specifically discussed in the following section of the chapter.

Another key danger in over-reliance on data mining may be described as "seduction bias" in which an extensive mathematical veneer of certain algorithms may seduce the safety reviewer/data miner into believing that the enormous deficiencies and distortions in the data have been neutralized [48]. Another aspect is the so called "self-deception bias", which can occur when a data miner with a strong incentive to believe in a particular outcome may consciously or subconsciously try to avoid results that contradict pre-existing expectations. In other words, the data miner may apply nonspecific case definitions of uncertain clinical relevance and/or sequential mining with different subsets of the database and/or candidate data mining parameters, until the "desired" output is achieved [48].

Concerning the overall quality of reports, there is still room for improvement due to the vast amount of missing data, especially on clinically relevant information regarding the time-to-onset of the reaction, the dechallenge/rechallenge information (an important component of the causality assessment). There should also be a commitment to improving the quality of the data, which is ultimately the rate limiting step. A recent study address some challenges and limitations of current pharmacovigilance processes in terms of data completeness and resorted to use the case of flupirtine to exemplify the need for refined ADR reporting [62].

The pattern of reporting is also widely influenced by external factors, which may affect the reliability of detected signal. Among the most important confounders, the product age (i.e., the time on the market of the compound) and stimulated reporting should be acknowledged. It is widely accepted that when a drug first receives marketing authorization, there is generally a substantial increase in the spontaneous reporting of ADRs (especially during the first two years on the market), which then plateaus and eventually declines. This epidemiological phenomenon is called "Weber effect" and was repeatedly shown for non-steroideal antinflammatory drugs [63-65]. This aspect may be related to the increased attention of clinicians towards a novel drug and may intuitively imply that the number of new signals detected reaches a peak over time with a subsequent decline. However, a new therapeutic indication or dose regimen may result in a new reporting pattern, thus one should be aware of the lifecycle status of the drug under investigation as well as significant change in its use. In addition, media attention and publicity resulting

from advertising or regulatory actions (e.g., dear doctor letter or warnings against drug-related safety issues) may result in increased reporting and can generate a higher-than-expected reporting ratio, a phenomenon known as "notoriety bias" [66,67]. Notoriety bias can also affect drugs other than those directly involved in alerts, thus causing a sort of "ripple effect" [7,68]. Routine incorporation of time-trend axis should be therefore recommended when planning pharmacovigilance analysis to gain insight into the temporal appearance of the signal, especially when regulatory interventions may have affected the life cycle of drug. Even if all the studied drugs are similarly affected by notoriety at a given time, false positive signals could be generated if this notoriety effect on reporting is differentially diluted among prior reports for older drugs compared with more recently marketed drugs. This differential effect related to the time on the market was recently demonstrated by Pariente et al., [69] for five "old" antidepressants *versus* escitalopram. Finally, physicians' prescribing are affected by a number of factors, including the severity of the disease, which create the potential for confounded drug-effect associations, the so-called "channeling bias" [70].

Concerning DMAs, it should be acknowledged that disproportionality methods do not estimate reporting rates and cannot provide incidence (see **Figure 1**). No drug usage is involved in the calculation. Also while reporting rates can increase for all DEC, measures of disproportionality are interdependent: an increase in a measure of disproportionality for one combination causes a decrease for other related combinations. A similar scenario occurs when two drugs have very different reporting patterns. If one has many more reports in general than the other, but they have similar rates for a particular term, the drug with the lower overall reporting will have a disproportionality detected more easily [29].

2.8. DMAs: Current perspective in signal refinement

The issue of selecting *a priori* a reference group (among drugs or events) as control to calculate disproportionality is a matter of debate in the literature, because it can significantly impact the results and their clinical implication [5,71,72]. Indeed, this aspect represents a commendable effort towards the application and implementation of novel data mining tools for clinical application. Under certain circumstances, this approach may be of benefit to provide the clinical perspective of PhV, with the attempt to move beyond *mere* signal detection. For instance, a recent study on SJS/TEN underlined several drugs with no association, which may be considered as alternative treatment options [73]. Basically, there are four scenarios deserving mention: (a) the calculation of an intraclass ROR, (b) the selection of specific control drugs/events, (c) the removal of already known drug-event associations and (d) the removal of non serious ADRs such as nausea, vomiting, abdominal pain, which are highly likely to be reported and may mask serious ADRs.

a. The calculation of an intraclass ROR is based on the only analysis of reports recording drugs of interest instead of the entire database (e.g., calculate the ROR of pancreatitis for the antidiabetic drug exenatide by using ADRs reported with all antidiabetic agents) [7]. This would allow to assess the disproportionate reporting of a given compounds in comparison with other agents within the same therapeutic class, and

may therefore be viewed as a secondary sensitivity analysis to provide provisional risk stratification. Indeed, one of the most important aspect and a compelling need for a clinician is to identify safer molecules within therapeutic classes. Moreover, this strategy helps to mitigate the so-called "confounding by indication bias" [74], by limiting the analysis to a population of patients that presumably share at least a set of common risk factors and diseases. Nonetheless, the so-called "channeling bias" (i.e., the possibility that drugs may be differently prescribed in relation to the severity of disease) still remains.

b.   The selection of specific control group of events or drugs is currently perceived as an attractive strategy for signal detection, and consists in the identification and exclusion of a cluster of events with no proved relation with the drug of interest. From one side, this approach may be of benefit to detect the hidden signals, which may escape from identification when the standard approach including all events is employed [71]. The most important limitations are related to the fact that the identification of control events/drugs is based on the expert opinion and therefore cannot be automatically implemented for all ADRs. Most remarkably, the risk is to select controls on the basis of the lack of evidence on possible association, which does not necessarily mean evidence of no association.

c.   A different approach for selecting group is based on the removal of already known drug-event association, which are usually over-reported in a pharmacovigilance database and may therefore mask possible novel associations. This bias is defined in the literature as the "competition bias" [72]. Indeed, if the event is significantly associated with other drugs, this will modify the denominator by increasing the background reporting rate, thereby possibly decreasing the sensitivity of the signal detection process. However, this *ad hoc* strategy precludes its automated application signal detection. Notably, the relevance of competition bias may depend on the nature and severity of the ADR, as it may have significant impact for clinically serious and relatively common adverse events.

d.   Another approach is based on the so-called "masking" or "cloaking" effect, which has been largely recognized, but only recently has its impact been explored in the FDA_AERS [75]. One simple application could be to exclude from the analysis all non serious ADRs, which are often submitted by relevant manufacturers. Notably, the FDA_AERS database offers the possibility to identify these reports, which are categorized as: 15-day reports, serious periodic reports, or non serious periodic reports for new molecular entity within the first 3 years following FDA approval.

While one is particularly stimulated to publish and disseminate results on positive associations (i.e., statistically significant disproportionality), it is very important to report also negative findings, i.e., the lack of drug-event association, which may be of benefit especially for prescribers [73,76].

In the light of the inherent limitations affecting pharmacovigilance analyses, the most important step in data mining approach is related to the *a priori* management of database before the application of DMAs (i.e., the definition and processing of the initial raw

dataset). Therefore, the following section will describe key issues before applying statistical analysis to spontaneous reports and explore how they can impact results. We will provide insight into the FDA_AERS and address major methodological issues encountered when approaching data mining. This choice is mainly based on the fact that the FDA_AERS is a worldwide publicly available pharmacovigilance archive and, therefore, we believe that transparency may improve accuracy, allow comparison of results among researchers and foster implementation. By virtue of its large population coverage and free availability, the FDA_AERS is emerging as a cornerstone for signal detection in PhV and is progressively being exploited by US and European researchers [25,73,77,78]. In addition, a number of commercial tools, namely query engines are now available. For instance, the FDAble [79] and OpenVigil [80], which is a free search query available through the University of Kiel. Another public source on drug-induced adverse reactions is represented by the DrugCite [81], which allows a number of tools, such as mobile applications to describe the primary safety profile of drugs or calculator to assess the probability of an adverse event being associated with a drug. It also allows searching the FDA_AERS database by providing easily interpretable graphics for the adverse events reported over time, stratified by relevant category, ages and gender, thus offering a quick check for clinicians on safety information. This would be of benefit for the entire drug safety assessment.

3. Methodological issues: the need for a pre-specified dataset

According to the Medwatch Program [82] founded in 1993, the US Food and Drug Administration (FDA) collects spontaneous reports on adverse reactions of drug, biologic, medical device, dietary supplement or cosmetic. In addition to reports coming from products' manufacturer, as required by regulation, FDA receives adverse event reports directly from healthcare professionals (such as physicians, pharmacists, nurses and others) and consumers (such as patients, family members, lawyers and others). Healthcare professionals and consumers may also report these events to the products' manufacturers, which are required to send the report to FDA as specified by regulations.

Raw data from the MedWatch system, together with ADR reports from manufacturers, are part of a public computerized information database called AERS (Adverse Event Reporting System) [83].

The FDA_AERS database contains over 4 million reports of adverse events of worldwide human drugs and biological products (concerning about 3 million of patients) and reflects data from 1997 to the present.

In this section we describe the anatomy of the FDA_AERS database and analytically illustrate dictionaries and coding systems used by this database. Furthermore, we describe our method to perform a systematic mapping of information originally not-codified. Subsequently, we discuss two important limitations that affect this dataset, duplicates and missing data, and we suggest a strategy for their detection and handling.

3.1. Anatomy of the FDA_AERS database

FDA_AERS is a relational databases structured in compliance with the international safety reporting guidance (ICH E2B) [84] issued by the International Conference on Harmonization. FDA_AERS includes different file-organized tables linkable through specific items and grouped in quarterly periods. Data from first quarter of 2004 to present are freely available into FDA website and can be easily downloaded [85], whereas previous data can be purchased from NTIS (National Technical Information Service) website [86]. Each file-package covers reports within one quarter of year, with the exception of the first available year (1997) where file-package covered the November-December 1997 period.

All data are provided in two distinct formats: SGML and ASCII. Although the first format conforms to the guidelines of the ICH E2b/M2 concerning transmission of individual Case safety reports [84], the ASCII data files are more useable. Indeed, these files can be imported into all popular applications for relational database such as ORACLE®, Microsoft Office Access, MySQL® and IBM DB2®.

Each quarterly file package includes, besides a text file containing a comprehensive description, the following 7 data files:

- DEMO file (demographic characteristics), including information on "event date", patient "age" and "gender", "reporter country" and "reporter's type of occupation".
- DRUG file (information for reported medications), including role codes assigned to each drug: "primary suspect drug" (PS), "secondary suspect drug" (SS), "interacting" (I) or "concomitant" (C).
- REACTION file, including all adverse drug reactions coded by MedDRA terminology [87] (for an overview of this tool see below);
- OUTCOME file (type of outcome, such as death, life-threatening, hospitalization);
- RPSR file, with information on the source of the reports (i.e. company, literature);
- THERAPY file, containing drug therapy start dates and end dates for the reported drugs (0 or more per drug per event).
- INDICATIONS file, containing all MedDRA terms coded for the indications of use (diagnoses) for the reported drugs;

Each file will generate a specific table that is linkable to each other using "ISR number" (Individual Safety Report) as primary key field. The ISR is a seven digit number that uniquely identifies an AERS report and allows linking all data files (see **Figure 2**).

Another important field is the "CASE Number" identifying an AERS case, which can include one or more reports (ISRs) due to possible follow-up of the same drug-reaction pair. If correctly linked, Case number allows identifying all ISRs of the same reports. It is though important to introduce this concept since this field is crucial in de-duplication process. Indeed "duplicate" ISRs (multiple reports of the same event) will normally have the same CASE number (but different ISR numbers).
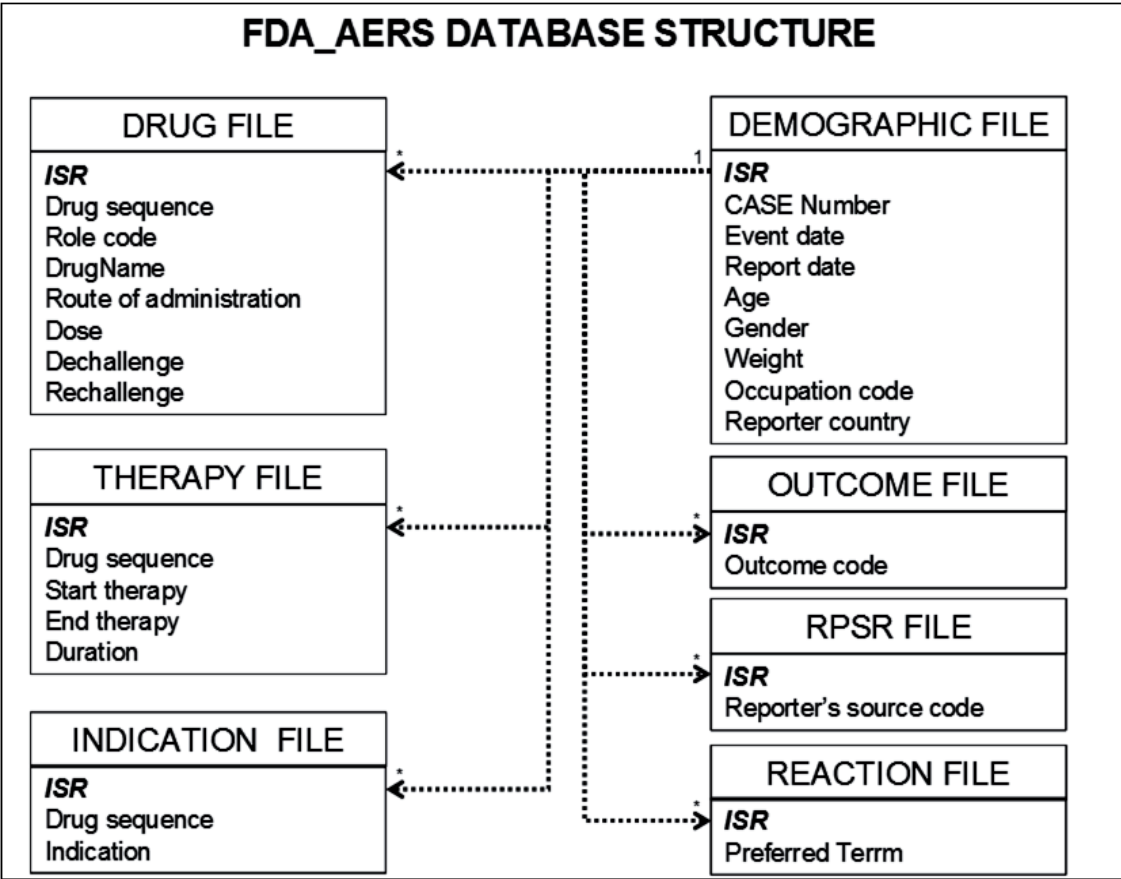
Figure 2. FDA_AERS database structure pictured according the Unified Modeling Language (UML) standardization. In each object, major fields of interest are reported in the FDA_AERS data files.

After downloading and assembling database, to proceed with pharmacovigilance data-mining, data management is essential for specific fields reporting (i) the adverse drug reaction ("PT-Preferred Term" in the REACTION file) and (ii) the suspected drug ("DrugName" in DRUG file). Unlike the adverse event, which is coded using the Medical Dictionary for Regulatory Activities (MedDRA) terminology [87], drugs are entered as textual terms and therefore an internal mapping procedure is required.

3.2. Adverse drug reaction identification: The multi-axial MedDRA hierarchy

In FDA_AERS, adverse events are coded according the Medical Dictionary for Regulatory Activities (MedDRA) terminology, which allows to classify adverse event information associated with the use of pharmaceutical and other medical products (e.g., medical devices and vaccines). This terminology is clinically validated, developed and maintained by international medical experts.

The dictionary was designed in the early 1990s for the use in the pharmaceutical industry and regulatory environment, to support all stages of the regulatory process concerning human medicines. It is owned by the International Federation of Pharmaceutical Manufacturers and Associations (IFPMA) and it was approved by the International Conference on Harmonisation (ICH) [88,89]. MedDRA is maintained by MSSO (MedDRA

Maintenance and Support Service Organization) [87] and JMO (Japanese Maintenance Organization) [90], which are independent organizations with missions to provide international support and continued development of terminology and to foster use of this tool. Access rights to MedDRA ASCII files are free for non-profit organizations (i.e. regulatory authorities, academic institutions, patients care providers and non-profit medical libraries), while a purchase subscription is required to business companies (i.e. pharmaceutical industries, contract research organizations, software developers).

The MedDRA terminology contains more than 60,000 terms for medical conditions, syndromes, diagnosis, clinical signs, laboratory and clinical investigations and social circumstances. These are organized within 5 hierarchical levels that help to bring together similar medical conditions. The structural elements of the MedDRA terminology are:

- SOC (System Organ Class) - Highest level of the terminology, and distinguished by anatomical or physiological system, etiology, or purpose
- HLGT (High Level Group Term) – Subordinate to SOC, superordinate descriptor for one or more HLTs
- HLT (High Level Term) – Subordinate to HLGT, superordinate descriptor for one or more PTs
- PT (Preferred Term) – Represents a single medical concept
- LLT (Lowest Level Term) – Lowest level of the terminology, related to a single PT as a synonym, lexical variant, or quasi-synonym

Although LLTs represent the bottom of the hierarchy, they consist of synonymous, lexical variants and representations of similar conditions; therefore PT level is generally considered the favourite for use in pharmacovigilance analysis, because it coresponds to a unique medical concept.

Since a PT (with its subordinate LLTs) may be represented in more than one of the superordinate levels, the MedDRA hierarchy is multi-axial. When a PT is included in two o more axes, a single SOC is designed as "primary" and others as "secondary" locations. This feature provides more flexibility in choosing and retrieving conditions to investigate. An example of the multi-axial hierarchical structure of MedDRA is shown in **Figure 3**.

To assist terminology searches and data retrieval, especially when MedDRA is used in post-marketing surveillance, utilities, called Standardised Medical Query (SMQs), were built and continually updated in a cooperative effort between the ICH and the Council for the International Organization of Medical Sciences (CIOMS). SMQs are groups of PTs and HLTs (with all their subordinate PTs) related to a defined medical condition or area of interest. SMQs may include very specific as well as less specific terms. The term lists are tested to verify their sensitivity and to avoid "noise" presence. SMQs may have certain specific design features: narrow and broad scope, algorithmic approach and hierarchical structure. "Narrow and broad searches" allow to identify cases that are highly likely to represent the condition of interest (a "narrow" scope) or to identify all possible cases, including some that may prove to be of little or no interest on closer inspection (a "broad" scope). In addition, for

some SMQs a combination of search terms from various sub-categories of the broad search terms is available (algorithmic search approach). This strategy is able to further refine the identification of cases of interest compared to the broad search category, therefore it yields greater sensitivity compared to the narrow search and greater specificity compared to the broad search. Moreover, in order to create more inclusive queries, some SMQs are a series of queries related to each other in a hierarchical relationship [91].
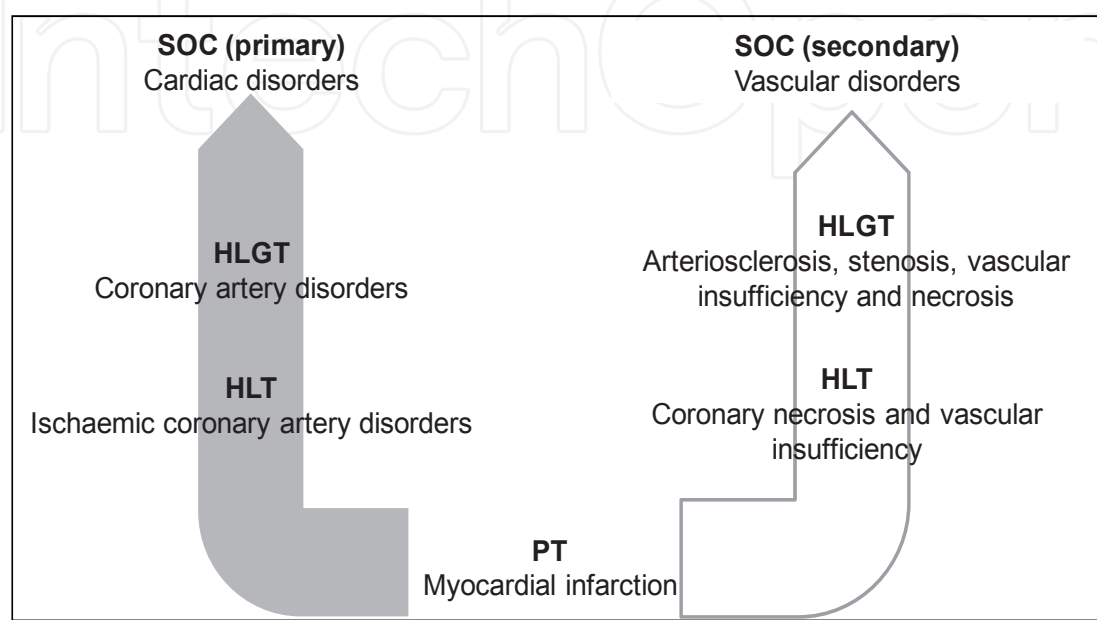

Figure 3. Example of MedDRA multi-axial structure from preferred term (PT) to system organ class (SOC), through high level term (HLT) and high level group term (HLGT).

Concerning the analysis of FDA_AERS, a specific knowledge of MedDRA terminology is essential in order to link "PT" field of REACTION table of FDA_AERS with this terminology. This represents the basis for different data-mining approaches to detect pharmacovigilance signals. Indeed, whereas an approach based on PT level is useful to investigate a particular DEC, it is unfeasible for large-scale screening of all possible DEC [92]. Furthermore, as demonstrated by Pearson et al. [93] the performance of a disproportionality detection improved from PT level to HLT level and to SMQ. However, a broader-level terms might generate false signals because it contains disparate PTs, not all clinically relevant.

3.3. Drug mapping: From free text to the ATC code

In FDA_AERS, the drug(s) used by the patient are reported in the "DRUGNAME" field as free text: either brand name or generic name can be reported, but also a combination of both, and misspellings can be present because of the lack of drug codification. This issue represents an important limitation of FDA_AERS data mining, as recognized by various authors [54,94]. To address this aspect, an *a priori* mapping process to analyse the possible association between drug and adverse reaction is necessary. Therefore, we created an *ad hoc* drug-name archive, including all generic and trade names of drugs marketed in US and in

most European Countries by using public lists freely available on authoritative websites and
public accessible drug dictionaries. In particular we used the following lists:

- **Drugs@FDA Data Files** [95]: the freely downloadable version of the Drugs@FDA online
  version. It contains information about FDA-approved brand name and generic
  prescription and over-the-counter human drugs and biological therapeutic products. It
  comprises drug products approved since 1998 to present. This list supplies drug name,
  as brand or generic, with the relevant active ingredient and other information (e.g.
  dosage form, route of administration, approval date). While the online application is
  updated daily, downloadable data files are updated once per week. Data files consist in
  9 text tables, which can be imported into a database or spreadsheet.
- **WHO Drug Dictionary** [96]: a computer registry containing information on all drugs
  mentioned on adverse reaction reports submitted by 70 countries participating in the
  WHO International Drug Monitoring Programme from 1968 onwards. This registry
  provides proprietary drug names together with all active ingredients and the chemical
  substances. Drugs are classified according to the Anatomical-Therapeutic-Chemical
  classification (ATC) [97] (for more details on this classification see below). Drugs
  recorded are those which have occurred in adverse reaction reports, and therefore the
  database covers most drugs used in countries involved in the WHO Programme. In
  collaboration with IMS Health, the WHO Drug Dictionary has been enhanced to
  include commonly used herbal and traditional medicines [96]. The registry is property
  of the Uppsala Monitoring Centre (UMC) and can be used by WHO collaborative
  centres or under license subscription. The Dictionary is available in different formats as
  data-files, with pre-defined relationships between the tables: a sample of registry and its
  technical description is freely available [98]. Our research group used this tool through
  the collaboration with the Pharmacology Unit of University of Verona as Reference
  Centre for Education and Communication within the WHO Programme for
  International Drug Monitoring.
- **EUROMEDSTAT** (European Medicines Statistics) Project [99]: a collaborative project
  promoted by the Italian National Research Council and funded by the European
  Commission. It started in 2002 and involved academics and government agencies from
  most European Union Member States, representatives of the WHO and the Council of
  Europe. Among aims of the project there was to build a European database of licensed
  medicines in order to perform comparison of drug utilization and drug expenditure
  among European countries. This database covers only a two-year period (2002-2003)
  and includes trade names of licensed medicines with relevant active ingredients and
  ATC codes, in addition to other information (e.g. price, reimbursement rules and
  licensed country).

The three lists are separately used to map "DrugName" reported in FDA_AERS into the
corresponding active substance. This mapping approach allowed allocating a substance
name to about 90% of all records in the entire database. Then, substance names have been
indexed according to the Anatomical Therapeutic Chemical (ATC) classification [97]. This
task was performed to facilitate data elaboration and allow analysis of database at different

level (from specific active substance to the broad therapeutic group). Indeed, in the ATC classification, the active substances are divided into different groups according to the organ or system on which they act and their therapeutic, pharmacological and chemical properties. This international classification is developed and maintained by the WHO Collaborating Centre for Drug Statistics Methodology (WHOCC), and was recommended as a standardized nomenclature since 1981. The WHOCC releases a new issue of the complete ATC index annually.

In the ATC classification, drugs are divided into different groups according to the organ or system on which they act and/or their therapeutic and chemical characteristics. The classification assigns to each substance one (or more) 7-digit alphanumeric code (ATC code), where it can recognize 5 different levels (**Table 3**):

- 1st level indicates the anatomical main group and consists of one letter (there are 14 main groups);
- 2nd level indicates the therapeutic main group and consists of two digits;
- 3rd level indicates the therapeutic/pharmacological subgroup and consists of one letter;
- 4th level indicates the chemical/therapeutic/pharmacological subgroup and consists of one letter;
- 5th level indicates the chemical substance and consists of two digits.

Level	ATC Code	Description
1st	A	ALIMENTARY TRACT AND METABOLISM
2nd	A10	DRUGS USED IN DIABETES
3rd	A10B	BLOOD GLUCOSE LOWERING DRUGS, EXCL. INSULINS
4th	A10BA	Biguanides
5th	A10BA02	Metformin

Table 3. Example of ATC codification for the antidiabetic metformin

Therefore, each bottom-level ATC code stands for a specific substance in a single indication (or use). This means that one drug, with more therapeutic uses, can have more than one ATC code. On the other hand, several different trade names share the same code if they have the same active substance and indications.

Although the ATC classification has a particular codification for combination products, in our methodology we decided to split the single active substances of the combination. This allows to analyse risk profile of a given substance and not of a particular product.

The drug mapping strategy described above is not the only possible to manage DrugName field of FDA_AERS. Indeed, other researchers used different methods and tools to index drugs and analyse the same database. For example, a research group of the Columbia University [54] first recognised drug name by using MedLEE, a text processor that translates clinical information from textual to vocabularies of the Unified Medical Language System (UMLS); afterwards they mapped drug names to generic name using RxNorm, a tool linking substance names (normalized names) of drugs with many drug names commonly used. A

Japanese research team [94] attempted to unify all drug names into generic names by detecting spelling errors through GNU Aspell, an open source spell checker.

3.4. How to handle missing data

As expected in large medical databases, also FDA_AERS is troubled by missing data, especially in old case reports. The fraction of missing data varies strongly with the variable considered and with the file analyzed. For example, in cumulative quarter DEMO files of 2004 – 2010 period, gender is 8% missing, age is 34% missing and event_date is 34% missing. These percentages are higher in corresponding THERAPY files, where end_therapy_date is not-filled in 64% of records. Reasons for the different fractions of missing data are still unclear and could be related to the reporter type, as argued by Pearson [100]. Pearson reported a difference of 10% in completeness of age field between reports coming from manufacturers (21.3% of missing age data) and those directly from patients or health professionals (31.1%).

Missing data is a well-recognized problem in large databases and widely discussed in the literature [101,102]. Various approaches have been used to deal with it with different consequences. Generally, these strategies fall into three classes: omission of incomplete records (deletion of records), imputation of missing values (single imputation) and computational modifications of missing fields (multiple imputation). The first approach (deletion of records) reduces the sample size and the subsequent statistical efficiency; it should be used only for small fractions of missing data. For large fractions of missing data one alternative is represented by single imputation of missing data on the basis of those available, but also this approach could modify the dataset. Another approach is the multiple imputation strategy, that allows to complete automatically missing information creating several different plausible imputed datasets and appropriately combining their results. It is obvious that neither this latter is free from drawbacks.

In our studies on FDA_AERS, we preferred the single imputation approach to manage missing data. We applied it only to demographic data (DEMO file) in order to improve quality and completeness of information for each case and, subsequently, to identify and remove duplicates (for details see below). To this purpose, we chose the following 4 key-fields, essential to characterize a single case: event_date, age, gender, reporter_country. Since each case, identified by a "CASE number", consists of one or more reports (ISRs), ideally representing the follow-up reports, we firstly linked each case with owner ISRs. Then, when a high level of similarity was present between two records (at least 3 on 4 of considered key-fields), we filled missing data. On the contrary, records with more missed key-fields were deleted. We acknowledge that this missing data handling could generate some biases, but, in the intent of our analyses, we consider these possible biases less hazardous than the deletion of a high number of records.

Noteworthy that FDA_AERS files different from DEMO could require other missing data handlings. Therefore, we advise to choose the appropriate methodology on the basis of both type of field and the magnitude of missing data.

3.5. Detection and removal of duplicates

A major problem in spontaneous reporting data is the presence of duplicates (i.e. the same report submitted by different reporters) and multiple reports (i.e. a follow-up of the same case with additional and updated information). Unfortunately, the exact extent of duplication present in FDA_AERS, and in other similar databases, is unknown. As a matter of fact, although this form of data corruption and its distortion in disproportionality analysis are well-recognized by researchers, only a few of studies examined the issue [103]. There are two major reasons for duplicate generation: different sources (e.g. health professionals, patients and manufacturers providing separate case reports related to the same event) and failure in linking of follow-up case reports to first event records. Some circumstances could generate extreme duplication [104]. For example, the reporting of the same event related to various concomitant drugs by manufacturers of each suspected drug (manufacturers have a statutory obligation to report the adverse event to FDA and to forward it to other involved manufacturers). Another possible source of duplication can be retrieved in the transfer of reports from national centers to the FDA.

Since in FDA_AERS the drawback of duplication is still unsolved and it can distort the final results, it becomes essential to mitigate this phenomenon before initiating the actual data mining process. To this aim, we performed a deduplication procedure firstly linking CASE number with relevant ISRs, secondly through the 4-key fields used in missing data handling (event_date, age, gender, reporter_country) and thirdly revising drug and event reported.

CASE	EVENT_DT	AGE	GNDR_COD	REPORTER_COUNTRY	Substance_Name	PT
6112494	20060624	59	M	UNITED STATES	AZITHROMYCIN	SYNCOPE
6112494	20060624	59	M	UNITED STATES	DELAVIRDINE	SYNCOPE
6112494	20060624	59	M	UNITED STATES	LAMIVUDINE	SYNCOPE
6112494	20060624	59	M	UNITED STATES	RITONAVIR	SYNCOPE
6112494	20060624	59	M	UNITED STATES	SAQUINAVIR	SYNCOPE
6112494	20060624	59	M	UNITED STATES	sulfamethoxazole and trimethoprim	SYNCOPE
6112494	20060624	59	M	UNITED STATES	ZIDOVUDINE	SYNCOPE
6112494	20060624	59	M	UNITED STATES	ABACAVIR	SYNCOPE
6120628	20060624	59	M	UNITED STATES	SAQUINAVIR	SYNCOPE
6120628	20060624	59	M	UNITED STATES	ABACAVIR	SYNCOPE
6120628	20060624	59	M	UNITED STATES	AZITHROMYCIN	SYNCOPE
6120628	20060624	59	M	UNITED STATES	DELAVIRDINE	SYNCOPE
6120628	20060624	59	M	UNITED STATES	RITONAVIR	SYNCOPE
6120628	20060624	59	M	UNITED STATES	SULFAMETHOXAZOLE	SYNCOPE
6120628	20060624	59	M	UNITED STATES	TRIMETHOPRIM	SYNCOPE
6120628	20060624	59	M	UNITED STATES	ZIDOVUDINE	SYNCOPE
6120628	20060624	59	M	UNITED STATES	LAMIVUDINE	SYNCOPE
6139100	20060624	59	M	UNITED STATES	ZIDOVUDINE	SYNCOPE
6139100	20060624	59	M	UNITED STATES	ABACAVIR	SYNCOPE
6139100	20060624	59	M	UNITED STATES	AZITHROMYCIN	SYNCOPE
6139100	20060624	59	M	UNITED STATES	DELAVIRDINE	SYNCOPE
6139100	20060624	59	M	UNITED STATES	LAMIVUDINE	SYNCOPE
6139100	20060624	59	M	UNITED STATES	RITONAVIR	SYNCOPE
6139100	20060624	59	M	UNITED STATES	SAQUINAVIR	SYNCOPE
6139100	20060624	59	M	UNITED STATES	SULFAMETHOXAZOLE	SYNCOPE
6139100	20060624	59	M	UNITED STATES	TRIMETHOPRIM	SYNCOPE

Record: ◄◄ ◄ [1] ► ►► ►* di 26 (Filtrati)

Figure 4. Example of duplicates in FDA_AERS for the 3 same reports of syncope

As advised by the FDA, we link the Case number with (ISRs) because of a "case" consists of one or more reports (ISRs) and if correctly linked, a follow-up report will have the same CASE number as the initial report (but a different ISR number). Then we grouped all records of a given case number and we associated the demographic information (event_date, age, gender, reporter country) the suspected drug/s and the reported event/s. Starting from this dataset, we performed a semi-automated de-duplication procedure in order to delete reports with different case number but reporting the same event. This procedure is based on the assumption that two or more apparently identical events are unlikely to be different events if the following conditions are fulfilled: they occurred in the same data with the same drugs, they were reported for same patients coming from the same country (**Figure 4** shows an example of duplication in FDA_AERS).

The following scheme (**Figure 5**) summarizes all the illustrated steps of our datamining strategy in analyzing the FDA_AERS.
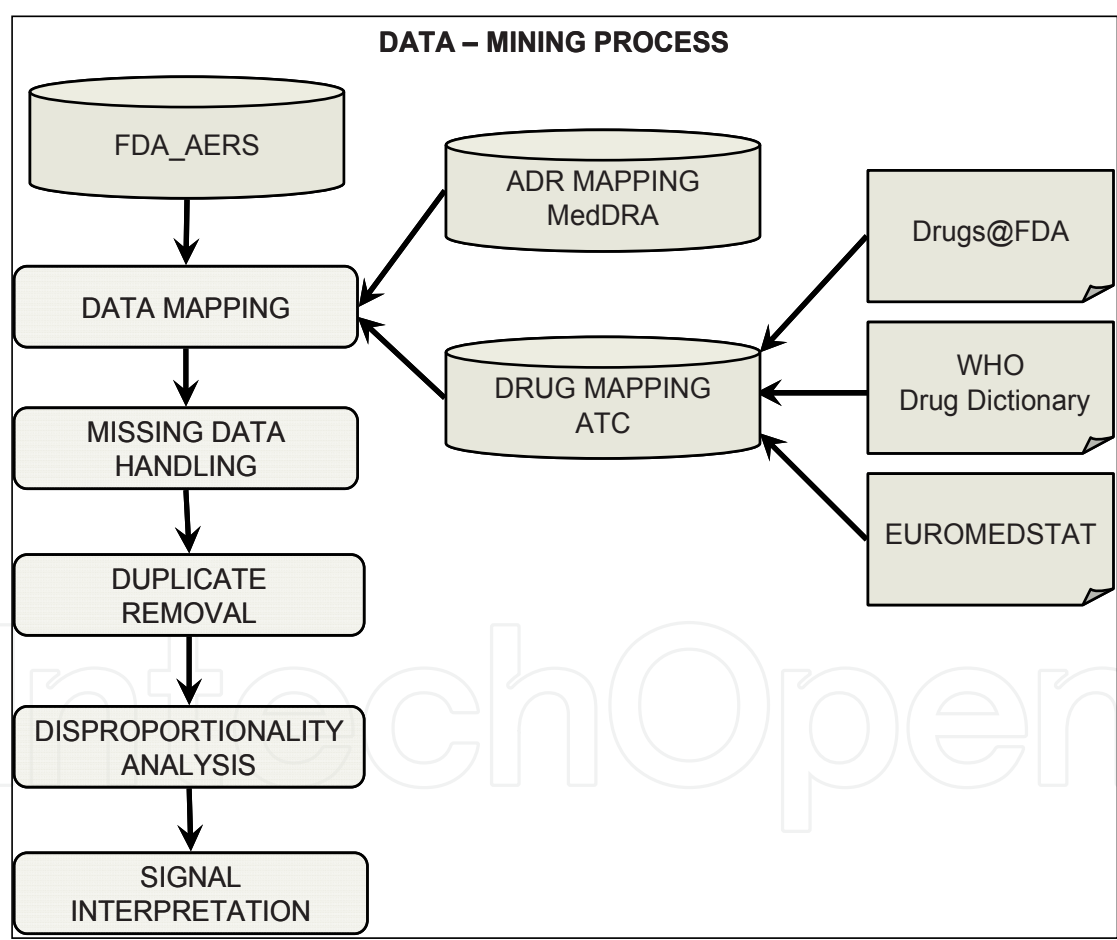

Figure 5. Steps of FDA_AERS data-mining process

4. Results of FDA_AERS data mining: some examples

Once the dataset has been processed for duplicates and missing data, there are some issues requiring further considerations to perform the complete analysis.

First, the analysis is usually conducted in terms of drug-event pairs, where every event is considered as many times as the reported drugs. Each of these drug-event pairs will be classified as cases or non cases, depending on the event of interest.

Second, the disproportionality can be performed at each level of MedDRA terminology (i.e., SOC, HLGT, HLT, PT) and the level of specificity of terms could be best defined by case-specific clinical judgment, rather than by rigid use of one level of terminology [105]. The importance of the choice of the level is well represented by a study aimed to investigate the safety profile of thiazolidinediones by using FDA_AERS [34]: while the disproportionality analysis at SOC level "injuries, poisoning, …" did not find any statistically significance for "injury", the analysis of immediately lower level HLGT "Bone and joint injuries" showed a disproportionality signal for both thiazolidinediones.

Third, disproportionality can be referred to all drugs included in the entire database or to a specific therapeutic class (see also reference group issues, above), isolated through the ATC code. The latter approach is needed to avoid the confounding by indication that occurs when a given class of drugs is preferentially used in subjects who have *a priori* higher or lower risk of presenting the event under investigation [74].

Fourth, disproportionality can be calculated by using two different approaches: a sequential cumulative strategy or a quarter-by-quarter approach. The cumulative analysis is the most adopted technique, both by regulators, academic researchers and manufacturers: it consists in the identification of a precise time point and a retrospective evaluation of reports, which are cumulatively considered. Indeed, several published papers on disproportionality cover a time period encompassing a number of years. Notably, this time period can be also split in different "time windows" (e.g., years, months), which can be analyzed for instance on a quarter-by-quarter basis (i.e., by calculating the ROR independently for a given window). While this approach may be useful to evaluate the behavior of the disproportionality over time and check whether the signal appears or disappears, it should be acknowledged that this strategy is highly influenced by fluctuations of the reporting, caused by seasonal changes in reporting and mostly by the typical attitude of recording the submitted reports and entering the database in January. This may cause an erroneous peak in reporting, which does not reflect the actual occurrence of the ADR. As an explanatory example, **Figure 6** compares results from the cumulative *versus* quarter-by-quarter analysis, performed on the association between exenatide and pancreatitis. Notably, the quarter-by-quarter analysis found a striking correspondence between the relevant FDA warning and the appearance of disproportionality (which, by contrast was found in the subsequent quarter in the cumulative analysis), but is characterized by large fluctuations of ROR (with very broad confidence intervals). The cumulative approach is therefore usually recommended as the most balanced and accurate approach for the analysis. Finally, an important aspect that should be kept in mind regards the selection of the time period of the analysis on the basis of the marketing authorization of the primary suspect drug. In **Figure 6**, only the period starting from the second quarter was described because exenatide received marketing approval in April 2005.
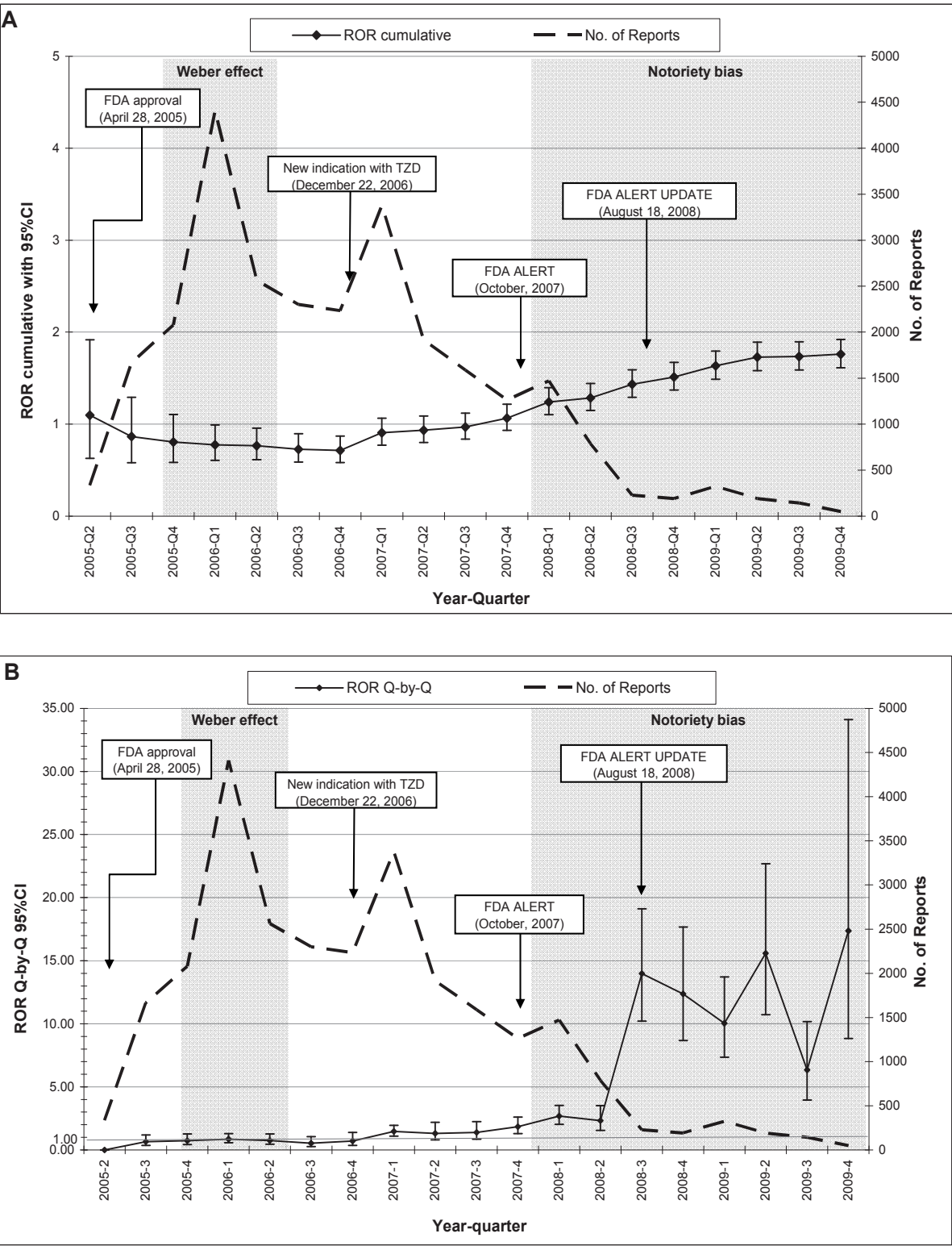
Figure 6. Comparison between cumulative (panel A, reprinted from [7], with permission of the copyright holder: Springer) and a quarter-by-quarter approach (panel B). Time trends of ROR with corresponding 95%CI of pancreatitis associated with exenatide use (full line) and relevant total number of spontaneous reports (dotted line). Boxes with relevant arrows indicate major regulatory measures in exenatide marketing history. Grey areas indicate the underlying epidemiological phenomena that may explain the observed peaks (see text for details).

5. Conclusions

On the basis of our experience, we have provided insight into the FDA_AERS to exemplify how to address major methodological issues discussed in the first part of the chapter. Because the FDA_AERS is a worldwide publicly available pharmacovigilance archive, we believe that fostering discussion among researchers will increase transparency and facilitate definition of the most reliable approaches. By virtue of its large population coverage and free availability, the FDA_AERS has the potential to pave the way to a new way of looking to signal detection in PhV. The existence of private societies developing and marketing software to analyze FDA_AERS data calls for the need to agree on data mining approaches.

PhV is a clinically oriented discipline, which may guide appropriate drug use through a balanced assessment of drug safety. Although much has been done in recent years, efforts are needed to expand the border of pharmacovigilance. For instance, although patients are important stakeholders in pharmacovigilance, little formal evaluation has been undertaken of existing patient reporting schemes. Notwithstanding some differences in the way various countries handle patient reports of ADRs, patient reporting has the potential to add value to pharmacovigilance by: reporting types of drugs and reactions different from those reported by clinicians; generating new potential signals; describing suspected ADRs in enough detail to provide useful information on likely causality and impact on patients' lives [106].

Registries (and their linkage to other data sources) have become increasingly appealing in postmarketing surveillance of medications, but their role is still very variable among countries. Notwithstanding significant limitations due to the lack of a control group and the need for complete case ascertainment to maintain data integrity, they are of utmost value for providing drug safety data through monitoring the incidence of rare adverse events particularly for highly specialized medications with significant financial cost [107]. Indeed, one of the most important aim of disease-based or drug-based registries is to promote appropriate use and guarantee access to innovative drugs.

One of the most important aspects to improve accuracy of signal detection in pharmacovigilance is to combine multiple methods, each in a distinct perspective (e.g., healthcare databases, spontaneous reporting systems), to increase the precision and sensitivity of the approach and complement limitations of each method [108]. In this context, the EU-ADR [109] and ARITMO projects [110] represent examples of multidisciplinary consortia aiming to create a federation of databases, namely healthcare databases and spontaneous reporting systems, to address drug-related safety issues.

In conclusion, our key messages are: (1) before applying statistical tools (i.e., DMAs) to pharmacovigilance database for signal detection, all aspects related to data quality should be considered (e.g., drug mapping, missing data and duplicates); (2) at present, the choice of a given DMA mostly relies on local habits, expertise and attitude and there is room for improvement in this area; (3) DMA performance may be highly situation dependent; (4) over-reliance on these methods may have deleterious consequences, especially with the so-called "designated medical events", for which a case-by-case analysis is mandatory and

complements disproportionality; and (5) the most appropriate selection of pharmacovigilance tools needs to be tailored to each situation, being mindful of the numerous biases and confounders that may influence performance and incremental utility of DMAs.

We support the implementation of DMAs, although one should not automatically assume that greater complexity is synonymous of greater precision and accuracy. Overall, data derived from DMAs should be considered with caution and guided by appropriate clinical evaluation. This clinical perspective should always be considered to support really appropriate drug use, balancing drug effectiveness, safety and, above all, actual patients' needs.

Author details

Elisabetta Poluzzi, Emanuel Raschi, Carlo Piccinni and Fabrizio De Ponti[*]
Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy

Acknowledgments

6. References

[1]   Edwards IR (2012) An agenda for UK clinical pharmacology. Pharmacovigilance. Br. J Clin Pharmacol 73: 979-982.

[2]   Leone R, Magro L, Moretti U, Cutroneo P, Moschini M, Motola D, Tuccori M, Conforti A (2010) Identifying adverse drug reactions associated with drug-drug interactions: data mining of a spontaneous reporting database in Italy. Drug Saf 33: 667-675.

[3]   Tatonetti NP, Denny JC, Murphy SN, Fernald GH, Krishnan G, Castro V, Yue P, Tsao PS, Kohane I, Roden DM, Altman RB (2011) Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels. Clin Pharmacol Ther 90: 133-142.

[4]   Maclure M, Fireman B, Nelson JC, Hua W, Shoaibi A, Paredes A, Madigan D (2012) When should case-only designs be used for safety monitoring of medical products? Pharmacoepidemiol. Drug Saf 21 Suppl 1: 50-61.

[5]   Elashoff M, Matveyenko AV, Gier B, Elashoff R, Butler PC (2011) Pancreatitis, pancreatic, and thyroid cancer with glucagon-like peptide-1-based therapies. Gastroenterology 141: 150-156.

[6]   Moore N (2011) Journal's withdrawal of article. Paper OK, title wrong. BMJ 342: d2732-

---

[*] Corresponding Author

[7] Raschi E, Piccinni C, Poluzzi E, Marchesini G, De Ponti F. (2011) The association of pancreatitis with antidiabetic drug use: gaining insight through the FDA pharmacovigilance database. Acta Diabetol [in press].

[8] Warrer P, Hansen EH, Juhl-Jensen L, Aagaard L (2012) Using text-mining techniques in electronic patient records to identify ADRs from medicine use. Br. J Clin Pharmacol 73: 674-684.

[9] Shetty KD, Dalal SR (2011) Using information mining of the medical literature to improve drug safety. J Am Med Inform. Assoc. 18: 668-674.

[10] Edwards IR, Aronson JK (2000) Adverse drug reactions: definitions, diagnosis, and management. Lancet 356: 1255-1259.

[11] Aronson JK, Hauben M, Bate A (2012) Defining 'surveillance' in drug safety. Drug Saf 35: 347-357.

[12] Aronson JK (2012) Adverse Drug Reactions: History, Terminology, Classification, Causality, Frequency, Preventability. In: John Talbot JKA, editors. Stephens' Detection and Evaluation of Adverse Drug Reactions: Principles and Practice. Chichester, UK: John Wiley & Sons, Ltd. pp. 1-119.

[13] Aronson JK (2009) Medication errors: definitions and classification. Br. J Clin Pharmacol 67: 599-604.

[14] (2010) Regulation (EU) No 1235/2010. Available: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:348:0001:0016:EN:PDF. Accessed 2012 Apr 27.

[15] (2010) Directive 2010/84/EU. Available: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:348:0074:0099:EN:PDF. Accessed 2012 Apr 27.

[16] Montastruc JL, Sommet A, Bagheri H, Lapeyre-Mestre M (2011) Benefits and strengths of the disproportionality analysis for identification of adverse drug reactions in a pharmacovigilance database. Br. J Clin Pharmacol 72: 905-908.

[17] (2012) Canada Vigilance Adverse Reaction Online Database. Available: http://www.hc-sc.gc.ca/dhp-mps/medeff/databasdon/index-eng.php. Accessed 2012 Apr 30.

[18] (2012) Netherlands Pharmacovigilance Centre Lareb. Available: http://www.lareb.nl/default.aspx. Accessed 2012 Apr 30.

[19] Srba J, Descikova V, Vlcek J (2012) Adverse drug reactions: Analysis of spontaneous reporting system in Europe in 2007-2009. Eur. J Clin Pharmacol [in press].

[20] Hammond IW, Gibbs TG, Seifert HA, Rich DS (2007) Database size and power to detect safety signals in pharmacovigilance. Expert Opin Drug Saf 6: 713-721.

[21] Meyboom RH, Hekster YA, Egberts AC, Gribnau FW, Edwards IR (1997) Causal or casual? The role of causality assessment in pharmacovigilance. Drug Saf 17: 374-389.

[22] Aronson JK, Hauben M (2006) Anecdotes that provide definitive evidence. BMJ 333: 1267-1269.

[23] Hauben M, Aronson JK (2007) Gold standards in pharmacovigilance: the use of definitive anecdotal reports of adverse drug reactions as pure gold and high-grade ore. Drug Saf 30: 645-655.

[24] Hauben M, Reich L (2004) Drug-induced pancreatitis: lessons in data mining. Br. J Clin Pharmacol 58: 560-562.

[25] Poluzzi E, Raschi E, Motola D, Moretti U, De Ponti F. (2010) Antimicrobials and the risk of torsades de pointes: the contribution from data mining of the US FDA Adverse Event Reporting System. Drug Saf 33: 303-314.

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

[27] Bate A, Lindquist M, Edwards IR (2008) The application of knowledge discovery in databases to post-marketing drug safety: example of the WHO database. Fundam. Clin Pharmacol 22: 127-140.

[28] Balakin KV, Hauben M, Bate A (2009) Data Mining in Pharmacovigilance. In: Balakin KV, editors. Pharmaceutical Data Mining: Approaches and Applications for Drug Discovery. Hoboken, NJ, USA: John Wiley & Sons, Ltd. pp. 341-378.

[29] Bate A, Evans SJ (2009) Quantitative signal detection using spontaneous ADR reporting. Pharmacoepidemiol. Drug Saf 18: 427-436.

[30] Stricker BH, Tijssen JG (1992) Serum sickness-like reactions to cefaclor. J Clin Epidemiol. 45: 1177-1184.

[31] Moore N, Kreft-Jais C, Haramburu F, Noblet C, Andrejak M, Ollagnier M, Begaud B (1997) Reports of hypoglycaemia associated with the use of ACE inhibitors and other drugs: a case/non-case study in the French pharmacovigilance system database. Br. J Clin Pharmacol 44: 513-518.

[32] Moore N, Thiessard F, Begaud B (2005) The history of disproportionality measures (reporting odds ratio, proportional reporting rates) in spontaneous reporting of adverse drug reactions. Pharmacoepidemiol. Drug Saf 14: 285-286.

[33] Hauben M, Reich L (2005) Potential utility of data-mining algorithms for early detection of potentially fatal/disabling adverse drug reactions: a retrospective evaluation. J Clin Pharmacol 45: 378-384.

[34] Motola D, Piccinni C, Biagi C, Raschi E, Marra A, Marchesini G, Poluzzi E (2012) Cardiovascular, ocular and bone adverse reactions associated with thiazolidinediones: a disproportionality analysis of the US FDA adverse event reporting system database. Drug Saf 35: 315-323.

[35] Evans SJ, Waller PC, Davis S (2001) Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports. Pharmacoepidemiol. Drug Saf 10: 483-486.

[36] DuMouchel W. (1999) Bayesian data mining in large frequency tables, with an application to the FDA spontaneous reporting. Am Stat 53: 177-201.

[37] Bate A, Lindquist M, Edwards IR, Olsson S, Orre R, Lansner A, De Freitas RM (1998) A Bayesian neural network method for adverse drug reaction signal generation. Eur J Clin Pharmacol 54: 315-321.

[38] Rothman KJ, Lanes S, Sacks ST (2004) The reporting odds ratio and its advantages over the proportional reporting ratio. Pharmacoepidemiol. Drug Saf 13: 519-523.

[39] Waller P, van PE, Egberts A, Evans S (2004) The reporting odds ratio versus the proportional reporting ratio: 'deuce'. Pharmacoepidemiol. Drug Saf 13: 525-526.

[40] Hauben M (2003) A brief primer on automated signal detection. Ann. Pharmacother. 37: 1117-1123.

[41] Hauben M, Zhou X (2003) Quantitative methods in pharmacovigilance: focus on signal detection. Drug Saf 26: 159-186.

[42] Hauben M, Reich L (2004) Safety related drug-labelling changes: findings from two data mining algorithms. Drug Saf 27: 735-744.

[43] Almenoff JS, LaCroix KK, Yuen NA, Fram D, DuMouchel W (2006) Comparative performance of two quantitative safety signalling methods: implications for use in a pharmacovigilance department. Drug Saf 29: 875-887.

[44] Li C, Xia J, Deng J, Jiang J (2008) A comparison of measures of disproportionality for signal detection on adverse drug reaction spontaneous reporting database of Guangdong province in China. Pharmacoepidemiol. Drug Saf 17: 593-600.

[45] Chen Y, Guo JJ, Healy DP, Lin X, Patel NC (2008) Risk of hepatotoxicity associated with the use of telithromycin: a signal detection using data mining algorithms. Ann. Pharmacother. 42: 1791-1796.

[46] Van Puijenbroek EP, Bate A, Leufkens HG, Lindquist M, Orre R, Egberts AC (2002) A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. Pharmacoepidemiol. Drug Saf 11: 3-10.

[47] Kubota K, Koide D, Hirai T (2004) Comparison of data mining methodologies using Japanese spontaneous reports. Pharmacoepidemiol. Drug Saf 13: 387-394.

[48] Hauben M, Patadia V, Gerrits C, Walsh L, Reich L (2005) Data mining in pharmacovigilance: the need for a balanced perspective. Drug Saf 28: 835-842.

[49] Chen Y, Guo JJ, Steinbuch M, Lin X, Buncher CR, Patel NC (2008) 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 22: 359-365.

[50] Poluzzi E, Raschi E, Moretti U, De Ponti F. (2009) Drug-induced torsades de pointes: data mining of the public version of the FDA Adverse Event Reporting System (AERS). Pharmacoepidemiol. Drug Saf 18: 512-518.

[51] van der Heijden PG, Van Puijenbroek EP, van BS, van der Hofstede JW (2002) On the assessment of adverse drug reactions from spontaneous reporting systems: the influence of under-reporting on odds ratios. Stat. Med 21: 2027-2044.

[52] Deshpande G, Gogolak V, Smith SW (2010) Data Mining in Drug Safety. Review of Published Threshold Criteria for Defining Signals of Disproportionate Reporting. Pharm Med 24: 37-43.

[53] Ahmed I, Thiessard F, Miremont-Salame G, Begaud B, Tubert-Bitter P (2010) Pharmacovigilance data mining with methods based on false discovery rates: a comparative simulation study. Clin Pharmacol Ther 88: 492-498.

[54] Harpaz R, Perez H, Chase HS, Rabadan R, Hripcsak G, Friedman C (2011) Biclustering of adverse drug events in the FDA's spontaneous reporting system. Clin Pharmacol Ther 89: 243-250.

[55] Strandell J, Bate A, Hagg S, Edwards IR (2009) Rhabdomyolysis a result of azithromycin and statins: an unrecognized interaction. Br. J Clin Pharmacol 68: 427-434.

[56] Strandell J, Caster O, Bate A, Noren N, Edwards IR (2011) Reporting patterns indicative of adverse drug interactions: a systematic evaluation in VigiBase. Drug Saf 34: 253-266.

[57] de Boer A. (2011) When to publish measures of disproportionality derived from spontaneous reporting databases? Br. J Clin Pharmacol 72: 909-911.

[58] De Bruin ML, Pettersson M, Meyboom RH, Hoes AW, Leufkens HG (2005) Anti-HERG activity and the risk of drug-induced arrhythmias and sudden death. Eur Heart J 26: 590-597.

[59] Giezen TJ, Mantel-Teeuwisse AK, Meyboom RH, Straus SM, Leufkens HG, Egberts TC (2010) Mapping the safety profile of biologicals: a disproportionality analysis using the WHO adverse drug reaction database, VigiBase. Drug Saf 33: 865-878.

[60] Stephenson WP, Hauben M (2007) Data mining for signals in spontaneous reporting databases: proceed with caution. Pharmacoepidemiol. Drug Saf 16: 359-365.

[61] Hazell L, Shakir SA (2006) Under-reporting of adverse drug reactions : a systematic review. Drug Saf 29: 385-396.

[62] Anderson N, Borlak J (2011) Correlation versus causation? Pharmacovigilance of the analgesic flupirtine exemplifies the need for refined spontaneous ADR reporting. PLoS One 6: e25221-

[63] Wallenstein EJ, Fife D (2001) Temporal patterns of NSAID spontaneous adverse event reports: the Weber effect revisited. Drug Saf 24: 233-237.

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

[65] McAdams MA, Governale LA, Swartz L, Hammad TA, Dal Pan GJ (2008) 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 17: 882-889.

[66] De Bruin ML, Van Puijenbroek EP, Egberts AC, Hoes AW, Leufkens HG (2002) Non-sedating antihistamine drugs and cardiac arrhythmias -- biased risk estimates from spontaneous reporting systems? Br. J Clin Pharmacol 53: 370-374.

[67] Motola D, Vargiu A, Leone R, Conforti A, Moretti U, Vaccheri A, Velo G, Montanaro N (2008) Influence of regulatory measures on the rate of spontaneous adverse drug reaction reporting in Italy. Drug Saf 31: 609-616.

[68] Pariente A, Gregoire F, Fourrier-Reglat A, Haramburu F, Moore N (2007) Impact of safety alerts on measures of disproportionality in spontaneous reporting databases: the notoriety bias. Drug Saf 30: 891-898.

[69] Pariente A, Daveluy A, Laribiere-Benard A, Miremont-Salame G, Begaud B, Moore N (2009) Effect of date of drug marketing on disproportionality measures in pharmacovigilance: the example of suicide with SSRIs using data from the UK MHRA. Drug Saf 32: 441-447.

[70] Petri H, Urquhart J (1991) Channeling bias in the interpretation of drug effects. Stat. Med 10: 577-581.

[71] Ooba N, Kubota K (2010) Selected control events and reporting odds ratio in signal detection methodology. Pharmacoepidemiol. Drug Saf 19: 1159-1165.

[72] Pariente A, Didailler M, Avillach P, Miremont-Salame G, Fourrier-Reglat A, Haramburu F, Moore N (2010) A potential competition bias in the detection of safety signals from spontaneous reporting databases. Pharmacoepidemiol. Drug Saf 19: 1166-1171.

[73] Papay J, Yuen N, Powell G, Mockenhaupt M, Bogenrieder T (2012) Spontaneous adverse event reports of Stevens-Johnson syndrome/toxic epidermal necrolysis: detecting associations with medications. Pharmacoepidemiol. Drug Saf 21: 289-296.

[74] Gregoire F, Pariente A, Fourrier-Reglat A, Haramburu F, Begaud B, Moore N (2008) A signal of increased risk of hypoglycaemia with angiotensin receptor blockers caused by confounding. Br. J Clin Pharmacol 66: 142-145.

[75] Wang HW, Hochberg AM, Pearson RK, Hauben M (2010) An experimental investigation of masking in the US FDA adverse event reporting system database. Drug Saf 33: 1117-1133.

[76] Hauben M, Hochberg A (2008) The Importance of Reporting Negative Findings in Data Mining. The Example of Exenatide and Pancreatitis. Pharm Med 22: 215-219.

[77] Jones SC, Sorbello A, Boucher RM (2011) Fluoroquinolone-associated myasthenia gravis exacerbation: evaluation of postmarketing reports from the US FDA adverse event reporting system and a literature review. Drug Saf 34: 839-847.

[78] Piccinni C, Motola D, Marchesini G, Poluzzi E (2011) Assessing the association of pioglitazone use and bladder cancer through drug adverse event reporting. Diabetes Care 34: 1369-1371.

[79] (2012) FDAble. Available: http://www.fdable.com/. Accessed 2012 Apr 30.

[80] (2012) OpenVigil. Available: http://webcl3top.rz.uni-kiel.de/pharmacology/pvt/queryaers.php. Accessed 2012 Apr 30.

[81] (2012) DrugCite. Available: http://www.drugcite.com/. Accessed 2012 Apr 30.

[82] (2012) MedWatch: The FDA Safety Information and Adverse Event Reporting Program. Available: http://www.fda.gov/Safety/MedWatch/default.htm. Accessed 2012 Apr 30.

[83] (2012) Adverse Event Reporting System (AERS). Available: http://www.fda.gov/drugs/guidancecomplianceregulatoryinformation/surveillance/adversedrugeffects/default.htm. Accessed 2012 Apr 30.

[84] (2012) International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Available: http://www.ich.org/. Accessed 2012 Apr 30.

[85] (2012) The Adverse Event Reporting System (AERS): Quarterly Data Files. Available: http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Surveillance/AdverseDrugEffects/ucm082193.htm. Accessed 2012 Apr 30.

[86] (2012) National Technical Information Service. Available: http://www.ntis.gov/about/index.aspx. Accessed 2012 Apr 30.

[87] (2012) MedDRA - the Medical Dictionary for Regulatory Activities. Available: http://www.meddramsso.com/. Accessed 2012 Apr 30.

[88] Brown E (2007) Medical Dictionary for Regulatory Activities (MedDRA). In: Mann R, Andrews E, editors. Pharmacovigilance. Chichester, UK: John Wiley & Sons, Ltd. pp. 167-186.

[89] Mozzicato P (2009) MedDRA. An Overview of the Medical Dictionary for Regulatory Activities. Pharm Med 23: 65-75.

[90] (2012) Japanese Maintenance Organization (JMO). Available: http://www.pmrj.jp/jmo/php/indexe.php. Accessed 2012 Apr 30.

[91] (2012) Introductory Guide for Standardised MedDRA Queries (SMQs). Version 15.0. Available: http://www.meddramsso.com/files_acrobat/SMQ_intguide_15_0_English_update.pdf. Accessed 2012 Apr 30.

[92] Bate A, Brown EG, Goldman SA, Hauben M (2012) Terminological challenges in safety surveillance. Drug Saf 35: 79-84.

[93] Pearson RK, Hauben M, Goldsmith DI, Gould AL, Madigan D, O'Hara DJ, Reisinger SJ, Hochberg AM (2009) Influence of the MedDRA hierarchy on pharmacovigilance data mining results. Int. J Med Inform. 78: e97-e103.

[94] Kadoyama K, Miki I, Tamura T, Brown JB, Sakaeda T, Okuno Y (2012) 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: 33-39.

[95] (2012) Drugs@FDA Data Files. Available: http://www.fda.gov/Drugs/InformationOnDrugs/ucm079750.htm. Accessed 2012 Apr 30.

[96] (2012) The WHO Drug Dictionary Enhanced. Available: http://www.umc-products.com/DynPage.aspx?id=73588&mn1=1107&mn2=1139. Accessed 2012 Apr 30.

[97] (2011) ATC/DDD Index 2012. Available: http://www.whocc.no/atc_ddd_index/. Accessed 2012 Apr 30.

[98] (2012) WHO Drug Dictionary Samples & Documents. Available: http://www.umc-products.com/DynPage.aspx?id=73551&mn1=1107&mn2=1139&mn3=6040. Accessed 2012 Apr 30.

[99] (2002) EURO-MED-STAT (European Medicines Statistics). Available: http://ec.europa.eu/health/ph_projects/2001/monitoring/fp_monitoring_2001_exs_12_en.pdf. Accessed 2012 Apr 30.

[100] Pearson R (2006) The Problem of Disguised Missing Data. Available: http://www.sigkdd.org/explorations/issues/8-1-2006-06/12-Pearson.pdf. Accessed 2012 Apr 30.

[101] Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, Wood AM, Carpenter JR (2009) Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ 338: b2393-

[102] Marston L, Carpenter JR, Walters KR, Morris RW, Nazareth I, Petersen I (2010) Issues in multiple imputation of missing data for large general practice clinical databases. Pharmacoepidemiol. Drug Saf 19: 618-626.

[103] Norèn GN, Orre R, Bate A, Edwards IR (2007) Duplicate detection in adverse drug reaction surveillance. Data Min Knowl Disc 14: 305-328.

[104] Hauben M, Reich L, DeMicco J, Kim K (2007) 'Extreme duplication' in the US FDA Adverse Events Reporting System database. Drug Saf 30: 551-554.

[105] Bate A, Lindquist M, Orre R, Edwards IR, Meyboom RH (2002) Data-mining analyses of pharmacovigilance signals in relation to relevant comparison drugs. Eur J Clin Pharmacol 58: 483-490.

[106] Avery AJ, Anderson C, Bond CM, Fortnum H, Gifford A, Hannaford PC, Hazell L, Krska J, Lee AJ, McLernon DJ, Murphy E, Shakir S, Watson MC (2011) Evaluation of patient reporting of adverse drug reactions to the UK 'Yellow Card Scheme': literature review, descriptive and qualitative analyses, and questionnaire surveys. Health Technol. Assess. 15: 1-iv.

[107] Willis CD, McNeil JJ, Cameron PA, Phillips LE (2012) Monitoring drug safety with registries: useful components of postmarketing pharmacovigilance systems. J Clin Epidemiol. 65: 121-125.

[108] Trifirò G, Patadia V, Schuemie MJ, Coloma PM, Gini R, Herings R, Hippisley-Cox J, Mazzaglia G, Giaquinto C, Scotti L, Pedersen L, Avillach P, Sturkenboom MC, van der LJ, Eu-Adr Group (2011) EU-ADR healthcare database network vs. spontaneous reporting system database: preliminary comparison of signal detection. Stud. Health Technol. Inform. 166: 25-30.

[109] EU-ADR project (2012). Available: http://www.alert-project.org/. Accessed 2012 Apr 30.

[110] ARITMO project (2012). Available: http://www.aritmo-project.org/. Accessed 2012 Apr 30.