Safety and Risk Benefit Analyses

Jeff Jianfei Guo

Contents

J. J. Guo (✉)
Division of Pharmacy Practice & Administrative Sciences, University of Cincinnati College of Pharmacy, Cincinnati, OH, USA
e-mail: jeff.guo@uc.edu

S. Piantadosi, C. Meinert (eds.), *Principles and Practice of Clinical Trials*,
https://doi.org/10.1007/978-3-319-52677-5_136-1

Abstract

In the past two decades more than 20 high-profile brand-name drugs including rofecoxib, troglitazone, cisapride, cerivastatin, natalizumab, gemtuzumab, and sibutramine were withdrawn from the market due to drug safety concerns related to severe adverse events. In 2005, the Food and Drug Administration (FDA) issued risk management guidance for the pharmaceutical industry. Subsequently the FDA Amendments Act in 2007 gave the FDA the authority to require pharmaceutical companies to develop and implement a Risk Evaluation and Mitigation Strategy (REMS) for specified prescription drugs and initiated the FDA Sentinel safety surveillance program in order to enhance the benefit-risk balance for pharmaceutical products. This chapter describes some basic concepts of drug safety, post-marketing surveillance, pharmacovigilance, and risk management. The safety signal detection algorithms and regulatory interventions will be also discussed. Finally some commonly used benefit-risk assessments (BRA) will be reviewed and discussed briefly in this chapter since the BRA methods are becoming critical tools for the life cycle of drug development and enhancing decision-making and regulatory interventions. The BRA is only one of unique analyses for clinical trials. Some concepts and methods of BRA may crossover with other analyses discussed in other chapters like "Intention to treat and alternative approaches", "Cost effectiveness analyses ", and "Development and validation of risk prediction models" Bayesian adaptive designs.

Keywords

Drug safety · Pharmacovigilance · Post-marketing surveillance · Risk management · Risk evaluation and mitigation strategy · Signal detection · Regulatory intervention · Benefit-risk assessments

Introduction

In response to 1937 drug safety tragedy of sulfanilimide elixir, the US Food Drug and Cosmetic Act was established in 1938 (FDA 2019). Due to the infamous thalidomide birth defect disaster in 1961, the Kefauver-Harris Amendment Act was passed in 1962 (FDA 2019). Both safety and efficacy have been two primary considerations for any new drug approval. Since the mid-1990s more than 20 high-profile brand-name drugs, including rofecoxib (Vioxx®), troglitazone (Rezulin®), cisapride (Propulsid®), cerivastatin (Baycol®), pemoline (Cylert®), natalizumab (Tysabri®), gemtuzumab (Mylotarg®), sibutramine (Reductil®), trovafloxacin (Trovan®), and propoxyphene (Darvocet/Darvon®), were withdrawn from the market (FDA 2005; Guo et al. 2010; Lis et al. 2012). In March 2005, the FDA issued risk management guidance for the pharmaceutical industry, which included three separate guidelines: Premarketing Risk Assessment, Development and Use of Risk Minimization Action Plans, and Good Pharmacovigilance Practices and

Pharmacoepidemiologic Assessment (FDA 2005). Subsequently Title IX of the FDA Amendments Act in 2007 gave the FDA the authority to require pharmaceutical companies to develop and implement a Risk Evaluation and Mitigation Strategy (REMS) for specified prescription drugs (Lis et al. 2012).

Regarding drug safety, some concepts and regulation requirements may crossover with this textbook Sect. 4 in "Regulation and Oversight" chapter on data and safety monitoring and reporting. Below let us discuss briefly about pharmacovigilance, drug safety surveillance, and post-marketing surveillance.

Pharmacovigilance

The FDA Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment guideline defined pharmacovigilance as "generally regarded as all post-marketing scientific and data gathering activities relating to the detection, assessment, understanding, and prevention of adverse events or any other product-related problems" (FDA 2005). "Pharmaco-" refers to pharmacology for pharmaceutical products. "Vigilance" is watchfulness in guarding against danger or providing for safety. The International Conference on Harmonization recommended similar pharmacovigilance planning E2E guideline (ICH 2004). The EU European Medicines Agency (EMA) provided guidelines for risk management plan (RMP) and good pharmacovigilance practices for pharma industry (Lis et al. 2012).

The scope of pharmacovigilance is focused on post-marketing activities. It applies pharmacoepidemiology theory and methodology to guide its related research or evaluation and plan for intervention. It also covers research activities related to drug utilization, spontaneous reporting system like FDA adverse event reporting systems, clinical pharmacology, pharmacogenetics, and regulatory sciences.

Post-Marketing Surveillance

Similar to pharmacovigilance, post-marketing surveillance is also a practice of monitoring the safety of pharmaceutical products and medical device after the product is approved by FDA. From clinical perspective, many effects of pharmaceutical products are usually not clear or unknown at the market entry. It is important to conduct post-marketing surveillance in general or specific population. Here are four major reasons for post-marketing surveillance.

1. During the clinical trials phases I, II, and III, there are usually relatively small numbers of study subjects involved in clinical trials. It will be impossible to capture or confirm many clinical effects or adverse events occurring in less than 1/500 due to limited sample size and statistical power.
2. Usually clinical trials have restricted study population. There are many inclusion and exclusion criteria for trials. Recruited study subjects or patients usually do not include elderly patients, pregnant women, or children.

3. Clinical trials are usually conducted during relatively short durations of exposure. Trials cannot detect long-term or delayed effects, for example, hepatotoxicity effects.
4. Clinical trials are usually involved in rigorous controlled drug use and lack of drug interaction information.

During the post-marketing phase, large numbers of patients may be exposed to medication therapy. Many elderly, children, and even pregnant women are likely involved in these medication exposures. It will be possible to measure the drug use and effects in large population. Post-marketing surveillance becomes a useful approach to monitor drug safety using real-world data like FDA Adverse Event Reporting Systems (FAERS) data, health insurance claims, electronic medical records, patient registries, and other electronic device or application collected health records. Post-marketing surveillance is essential to detect and confirm the rare adverse event or specific effects.

Drug Safety Surveillance: Passive and Proactive Models

Unlike post-marketing surveillance, the modern drug safety surveillance involves multidisciplinary methods to document, monitor, and evaluate adverse drug events or adverse drug reactions, and plan effective interventions using risk management tools. It applies to the life cycle of a drug development including investigational new drug, clinical trials phases I, II, III, new drug application, post-marketing (phase IV), and even removal from market. Drug safety surveillance incorporates epidemiologic methodology to monitor drug use and effects in large population using real-world data and modern risk management tools. Real-world data often include spontaneous reporting systems (e.g., FDA FAERS, WHO ADR), product and disease registries, electronic medical records, health survey data, computerized medical insurance claims, data collected from social media or special applications or devices, even pragmatic clinical trial data, and other public health records like national death index.

Figure 1 describes the traditional time line for drug development life cycle from preclinical animal and cell studies to clinical trials, to FDA new drug application, and to post-marketing Phase IV studies. Both pharmacovigilance and post-marketing surveillance are focused on Phase IV after products are approved with a small circle or scope. Drug safety surveillance, however, is involved in the life cycle of drug development with a large circle or scope.

Traditionally the pharmaceutical industry relied on the passive model of safety surveillance for several decades, which is to use FDA spontaneous reporting system (SRS) to identify any adverse event signals or cases of suspected drug-induced disease/condition and plan for interventions. All adverse cases are voluntary

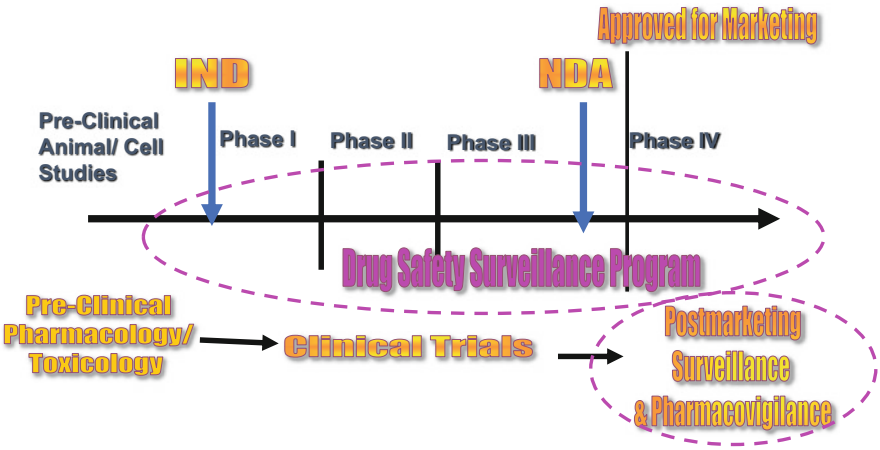

Fig. 1 Drug safety surveillance vs. post-marketing surveillance

based. No specific efforts are made to make sure more or all cases are reported. Passive surveillance is integrated to routine health-care delivery system.

In contrast, proactive safety surveillance involves "a systematic process for analyzing multiple observational health care data sources to better understand the effects of medical products" (Stang et al. 2010). Based on the proactive safety surveillance, there is a periodic solicitation of adverse event case reports, such as: (1) the FDA adverse event surveillance system; (2) acute severe hepatotoxicity surveillance program; (3) active post-marketing surveillance among the US hospitals; and (4) the US death certificate surveillance project. In 2008, the FDA launched the Sentinel Initiative which enhances the FDA's ability to proactively monitor the safety of drug, biologic and medical device products, which is a complement of existing FDA adverse event reporting system (Stang et al. 2010). With the Sentinel Initiative, the FDA can rapidly access medication use and effect at the patient level in large population using electronic health-care data. Based on a series of FDA-developed, standardized computer codes, all Sentinel partners (hospitals or health insurance companies) will generate a set of patient-level data accordingly. The large observational data allows one to identify potential cases actively through screening of hospital admission records, emergency department logs, medical wards, intensive care units, outpatient care, and medication use and effect. Appropriate screening terms and rigorous standardized procedures are necessary to minimize the number of missing cases.

The current practice involves a combination of active and passive systems using multiple resources and data, including: (1) real-world data from general population and special population; (2) disease history data of disease incidence and prevalence; (3) diagnosis and drug utilization data from outpatient and inpatient records. This kind of combination practice requires coordination and collaboration among various programs, clinicians, pharmacists, health-care providers, epidemiologists, information system specialists, research design efforts, and other health-care stakeholders.

Safety Signal Detections and Regulatory Interventions

All adverse events collected from clinical trials or post-marketing phase are required to document and evaluate accordingly. When relationship between drug and adverse event is established, a potential safety signal is created. The relevant adverse events are treated as signals that may involve in some suspicious adverse event or reaction. Commonly used safety signal detection algorithms and common regulatory interventions for drug safety surveillance are discussed below.

Signal Detection Algorithms

Compared to clinical trials and traditional epidemiologic studies like case-control and cohort study, the computerized drug safety signal detection algorithms or data mining algorithms are relatively new and characterized by providing a fast and efficient way of detecting possible adverse safety signals. Many large SRS databases have been developed in different countries and World Health Organization (WHO), which can be used for early safety detection using signal detection algorithms. Commonly used SRS available to the public access include the FAERS, the UK Yellow Card System, Canadian Adverse Drug Reaction Monitoring Program, WHO Uppsala ADR Center, etc.

Several signal detection algorithms have been described in literature, mainly including the reporting odds ratio (ROR) introduced by the Netherlands Pharmacovigilance Center Lareb, the proportional reporting ratio (PRR) introduced by the UK Yellow Card Scheme, the Multi-item Gamma Passion Shrinker (MGPS) adopted by the US FDA, and the information component (IC) introduced by WHO Uppsala ADR center (Chen et al. 2008b). Other algorithms including the Yule's Q and the Chi-square have been also introduced, but they are rarely used. Screening the adverse event reports based on safety signal detection algorithms could often save time and prevent some unnecessary efforts.

ROR is a measure of disproportionality. Compared to other algorithms, ROR is relatively easy to calculate. Like the traditional odds ratio in epidemiology, the ROR is an estimate of incidence rate ratio, calculating the odds of the exposure of suspected drug in those who had adverse events divided by the odds of the exposure of suspected drug in those without adverse events (Almenoff et al. 2005). The ROR with 95% confidence interval is computed using a 2 x 2 table: ROR = (A/B)/(C/D) = AD/BC. See Table 1 below.

The PRR is another early attempt of quantitative analysis of adverse drug reaction reports. It measures the strength of association between the suspected events and suspected drugs, using the similar calculation of the relative risk (RR). The higher value of PRR is the stronger strength of the safety signal appears to be. Literature suggests that criteria for the PRR consisting of PRR greater than 2, Chi-square greater than 4, and three or more reported cases are often used to identify the possible safety signals (Evans et al. 2001). In practice, the PRR with a value of more than 3 may highlight a need for evaluation of cases and further investigation. The

Table 1 The 2 x 2 table for the calculation of adverse event (AE) safety signals

	AE reports with the suspected drug	AE reports with all other drugs	Total
Reporting with suspected AEs	A	B	(A + B)
Reporting with other AEs	C	D	(C + D)
Total	(A + C)	(B + D)	N = A + B + C + D

Notes: A = the number of reports containing both suspected drug and suspected AE; B = the number of reports containing other drug use but with other AE; C = the number of reports containing the suspected AE but with other medications; D = the number of reports concerning other medications and other AEs

computation of PRR is same as the RR estimated in epidemiology and can be calculated using the 2x2 table.

$$\mathrm{PRR} = [\mathrm{A}/(\mathrm{A}+\mathrm{C})]/[\mathrm{B}/(\mathrm{B}+\mathrm{D})] = \mathrm{A}(\mathrm{B}+\mathrm{D})/\mathrm{B}(\mathrm{A}+\mathrm{C}).$$

Incorporating Bayesian approaches into data mining is a major initiative in the safety signal detection. Two slightly different Bayesian algorithms have been developed and applied in the post-marketing surveillance. One is IC of Bayesian Confidence Propagation Neural Network (BCPNN) (Bate et al. 2002). The other is the MGPS (DuMouchel 1999).

The WHO Uppsala Monitoring Center (UMC) has been applying the IC/BCPNN to detect drug safety signals using the WHO adverse drug reaction (ADR) database. The UMC collects ADR reporting data from more than 160 member countries, including the USA, the UK, German, France, Span, Italy, Japan, China, India, Canada, Mexico, Brazil, South Africa, etc. The computation for the IC/BCPNN basically includes two steps: (1) making an estimation of a prior probability (i.e., constructing a prior and a likelihood function); and (2) improving the estimation in light of incoming new information (i.e., updating the posterior mean) (Bate et al. 2002).

The MGPS is another approach that applies the Bayes' law into the adverse event signal detection. It was initially developed and applied to the FAERS database. This method assumes that the expected number of reports containing both the suspected drug and the suspected adverse event is fixed known, and could be computed after stratifying the reports based on the age, sex, or other factors that may influence the reports of drug and events (DuMouchel 1999).

The basic concepts involved in the AE signal detection or data mining algorithms are related to the 2 x 2 table presented in Table 1. They are related to analytical approaches of disproportionality analysis. Both ROR and PRR are relatively easy to understand. The computation is relatively inexpensive compared to IC/BCPNN and MGPS. Both IC and MGPS apply the Bayesian inference with slightly different models. Based on limited literature comparing the performance and sensitivity of these signal detections, Hauben and Reich study (2004) indicated PRR could signal

almost twice as many drug-event combinations as the MGPS did. Chen et al. (2008a, b) showed that ROR has better performance compared to PRR, IC, and MGPS. Both IC and MGPS require a set of mathematical knowledge and complicated calculations but often believed to provide more stable estimation compared to the ROR and PRR, especially with small frequency of drug-event combinations.

Regulatory Interventions for Drug Safety

Every pharmaceutical product has its unique experience of safety surveillance and government intervention story. Figure 2 illustrates a typical workflow of drug safety surveillance in post-marketing phase in the USA. After new drug application approval, FDA will actively document and evaluate the adverse drug events based on FAERS database. FAERS documented five key components including suspected drug and concomitant medication, suspected adverse event and outcomes, patient information, pharmaceutical information, and reporter's information which is confidential only use for FDA internal review. Once a safety signal is detected from an SRS, case review will be performed based on clinical information. For any significant case, pharmaceutical industry or government agency or academic institution would like to conduct pharmacoepidemiologic study using large real-world data, such as electronic health records, insurance claims database, or disease or drug registry data.

When the risk of drug adverse event is confirmed, FDA and pharmaceutical industry will plan for regulatory interventions, including label change or black box warning. Usually "Dear Doctor" and/or "Dear Pharmacist" "Dear Health Practitioner" letters will be sent out to relevant prescribers and health practitioners like pharmacists and nurse practitioners. Continue educations for relevant medical doctors and pharmacists/nurses will be conducted to inform the product risk and benefit.

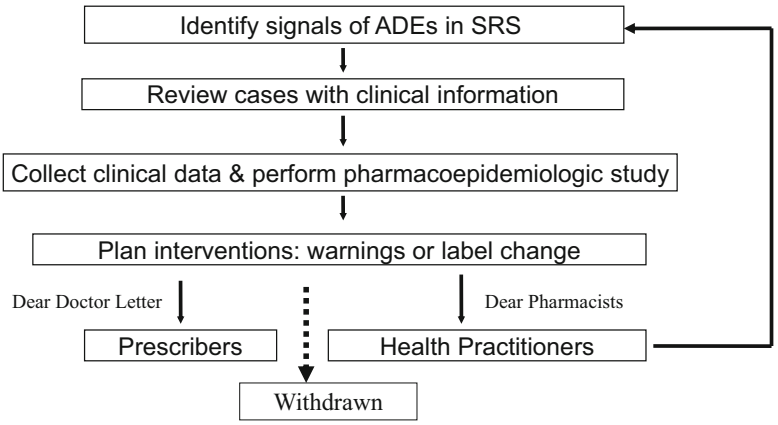

Fig. 2 Drug safety workflow for post-marketing surveillance

For some specific products, the restricted use may be applied for its distribution. Meanwhile government, pharmaceutical industry and academic institute will continue to assess the risk-event combinations and effectiveness of planned interventions. If the intervention is proved ineffective or risk remains high, the product will be withdrawn from the market either by FDA or by pharmaceutical industry voluntarily.

From the above discussion, the traditional common interventions include (1) labeling changes; (2) general warning or black-box warnings; (3) continue educations among medical doctors, pharmacists, and nursing communities; (4) "Dear Doctor Letter," or "Dear Health Practitioner (Pharmacist) Letter"; (5) restricted distribution; and (6) withdrawal (Guo et al. 2003, 2008).

In the recent years, other interventions have been practiced in the USA. These include: the third class or behind the counter of drug, modifying standards of new drug approval, new labeling models, patient package inserts (patient labeling), Medication Guides, special advertising campaign for product risk and benefit information, formal risk management plan with relevant informed consent and restrictions, and mandatory monitoring registries (e.g., isotretinoin, thalidomide).

The third class of drug refers to some certain pharmaceutical products to be used only under the supervision of a physician or dentist or pharmacist (Pray and Pray 2011). The other two classes of pharmaceutical products are related to prescription and over-the-counter medications. Some states have created a third class of medications including certain codeine-containing syrups, methamphetamine (meth), and pseudoephedrine due to overdose pain medication use.

For the need of safety assessment in post-marketing surveillance, some safety monitoring registries have been utilized for some specific products like isotrotenoin (Accutane®), thalidomide (Thalomid®), alosetron (Lotronex®), etc. The System for Thalidomide Education and Prescribing Safety (STEPS) is a good example of this kind of registry to prevent exposure to thalidomide during pregnancy (Zeldis et al. 1999). As a comprehensive program to control prescribing, dispensing, and use of the drug, the STEPS registry achieves the goals of controlling access to thalidomide; educating prescribers, pharmacists, and patients about risk and benefits of the product; and monitoring compliance of thalidomide use. Prescribing physicians need to ensure patient eligibility criteria and monitoring procedures. Pharmacies must also agree to comply with patient identification and monitoring criteria. Patients receive visual aids, written material, and verbal counseling about the benefit and risk of thalidomide therapy.

The iPLEDGE® program is a computer-based risk management registry designed to eliminate fetal exposure to isotretinoin during pregnancy through a special restricted distribution program. The brand-name was withdrawn from the US market in 2009. The iPLEDGE® is used for both prescribing and dispensing all isotretinoin with two goals: (1) no female patient starts isotretinoin therapy if pregnant; (2) no female patient on isotretinoin therapy becomes pregnant (see Website page: https://www.ipledgeprogram.com/iPledgeUI/home.u).

The Prescribing Program for Lotronex (alosetron) is another registry used for providers, pharmacists, and patients to reduce the risk of severe gastrointestinal

adverse events. Alosetron prescribing doctors have to sign up online and confirm that they can understand and diagnose irritable bowel syndrome with diarrhea and the possible side effects like constipation. (See Website page: https://www.lotronex.com/PrescribingProgramForLotronex.aspx.)

Benefit and Risk Assessments

From the public health perspective, the FDA evaluates the risks and benefits for the population perspective. For any approved medication, the prescriber serves as agent to manage risks and benefits for the individual patient. Meanwhile patients make their own decisions about treatment choices based on their personal valuation of benefits and risks. While the government agency cannot involve in any individual treatment decision, FDA's role is to ensure that accurate, substantiated, and balanced information about any approved drug and biologic product is available to the prescriber and the patient.

For ensuring drug safety, FDA issued risk management guidance for industry in 2005, including pre-marketing risk assessments, good pharmacovigilance and pharmacoepidemiological assessment, and risk minimization action plan (FDA 2005a, 2005b, 2005c). The risk management was defined as an iterative process designed to optimize the benefit-risk balance regulated products, which involves in assessing a product's risk-benefit balance, developing tools to minimize risk while preserving benefits, evaluating tools' effectiveness and reassessment risk-benefit balance, and making adjustments to risk management tools further to improve risk-benefit balance. (FDA 2005b). Incorporating with risk management guidance, the 2007 FDA Amendment Act gave the FDA the authority to require pharmaceutical companies to develop and implement an REMS for specified pharmaceutical products.

In March 2018, the US FDA published "benefit-risk assessment (BRA) in drug regulatory decision-making" which provides the general benefit-risk framework and calls for more creative approaches to conceptualizing, measuring, and applying BRA throughout the life cycle of a drug and biologic product (FDA 2018). The benefit-risk framework (BRF) is defined as "a structured, qualitative approach focused on identifying and clearly communicating key issues, evidence, and uncertainties in FDA's BRA and how those considerations inform regulatory decision." BRF should include four dimensions: analysis of condition, current treatment options, benefit, and risk management. For each dimension, two aspects of consideration need to be specified: (1) The detailed treatment of evidence and uncertainties about a drug's benefits and risks are considered in the context of the severity of the condition and the current medical needs for patients. (2) The conclusion and reasons about the drug benefit and risk in each dimension should be also documented.

Similarly, for the European Union, the Committee for Medicinal Products for Human Use is responsible to provide guidance for industry how to assess risks and benefits of authorized medicines on behalf of the EMA. The EMA also created the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance to

develop an algorithm to articulate risks and benefit profiles for pharmaceutical products.

Appropriate BRA can provide useful information for proactive intervention in health-care settings, which could save lives, reduce litigation, and lead to improved patient safety, better health outcomes, and lower overall health-care costs (Guo et al. 2010). Guo et al. (2010) systematically reviewed literature and summarized some common used quantitative BRA including:

- Quantitative framework for risk-benefit assessment (QFRBA)
- Benefit-less risk analysis (BLRA)
- Quality-adjusted time without symptoms and toxicity (Q-TWiST)
- Number needed to treat (NNT) versus number needed to harm (NNH)
- Relative value adjusted number needed to treat (RV-NNT)
- Minimum clinical efficacy (MCE)
- Incremental net health benefit (INHB)
- Risk-benefit plane (BRP) and risk-benefit acceptability threshold (RBAT)
- Probabilistic simulation methods (PSM) and Monte Carlo simulation (MCS)
- Multi-criteria decision analysis (MCDA)
- Risk-benefit contour (RBC)
- Stated preference method (SPM) or maximum acceptable risk (MAR)

In addition, other BRAs were found in the literature such as net efficacy adjusted for risk, net clinical benefit analysis, the principle of threes, and net-benefit-adjusted-for-utility analysis. Due to the uniqueness and complexity of each assessment, some commonly used BRA methods are discussed below.

Quantitative Framework for Benefit-Risk Assessment (QFBRA)

The QFBRA is widely used in drug safety surveillance by regulatory agencies and by the pharmaceutical industry, although it cannot combine risks and benefits into a single value. It includes all commonly used epidemiologic assessments such as incidence, relative risk, odds ratios, attributable risk, risk difference, relative risk reduction, and absolute risk reduction (Guo et al. 2010).

The Risk refers to a comprehensive set of all possible adverse drug events (ADEs) and a set of probabilities associated with these adverse outcomes. The basic expression of risk for pharmaceutical products is the **incidence** of an ADE, which can be defined as the number of new adverse drug events occurred in a defined population over a specific period of time divided by the population at risk over a specific period of time. There are two types of incidence measures: cumulative incidence and incidence density rate. The cumulative incidence is calculated as the number of new cases of ADE during a specified period divided by the total number of persons at risk for developing the ADE during the period. The incidence density rate is calculated as the number of new cases of ADEs during a specified period divided by the total number of person-time observed in population at risk during that

period. For example, 77 new cases of acute pancreatitis were identified among 1838 newly diagnosed patients receiving HIV treatment. These newly diagnosed patients had a total of 3943 treatment follow-up person-years. The incidence density rate was calculated as 77 / 3943 = 1.95 per 100 person-years (Guo 2005).

Relative risk (RR) or risk ratio can be calculated as probability of developing the ADEs with a risk factor divided by probability of developing the ADE without a risk factor. Using Table 1 2x2 table, RR = [A/(A + B)] / [C/(C + D)]. Confidence intervals (CI 95%) and p values are then calculated to answer the question as to whether there is a statistically significant risk. A relative risk without these parameters is of no value to the clinician. A relative risk of 1, of course, is the null or "no difference" relative risk. It is important to note that the relative risk does not indicate the magnitude of absolute risk. In other words, it reveals nothing about the incidence of the adverse events in either group. Similarly to RR, the odds ratios (OR) is often an estimate of RR and calculated as odds of exposed to risk factor divided by odds of no exposed to risk factor

$$\mathrm{OR} = (\mathrm{A}/\mathrm{B})/(\mathrm{C}/\mathrm{D}).$$

Attributable Risk (AR) otherwise known as **Risk Difference (RD)** is useful to measure absolute risk differences which can calculate the additional risk (incidence, probability) of ADEs in individual patient taking a new drug over or above that experienced by patients who are taking conventional drug therapy. Attributable risk is the additional incidence of an ADE related to exposure to either drug, taking into account the background incidence of ADEs, presumably from other causes that could include other drugs. AR or RD = probability of exposed to risk factor – probability of not exposed to risk factor. The population-based attributable risk is an extension of AR that uses total population data.

The ***benefit*** associated with a medication can measure how much risk reduced for adverse events associated with the disease. The **relative risk reduction (RRR)** is useful to calculate the ratio between the proportion of exposed individuals who experience a decline in adverse events divided by the proportion of unexposed individuals who experience such a benefit. Similarly, **attributable risk reduction or *absolute risk reduction* (ARR)** represents the decline in adverse events between exposed and unexposed groups. The main disadvantage of using the relative risk reduction in clinical decision-making is that it does not reflect the magnitude of the benefit over no therapy (e.g., the control group). Therefore, it will overestimate or underestimate the absolute impact of therapy when the outcome event incidence is very rare or very common, respectively. From clinical trials, benefit measurements might include clinically relevant efficacy parameters such as specific biomarkers, surrogates, and putative surrogate endpoints. For a safety surveillance study, benefit can be measured as medication adherence rate or treatment effectiveness.

Benefit-Less-Risk Analysis (BLRA)

The BLRA is a ratio combing both benefit and risk into one measurement. The benefit is measured by the response rate and risk is determined by the incidence of a serious side effect. Considering BRA only when there are serious side effects is not good enough because there are many situations when a treatment has a better efficacy outcome at the cost of more side effects even though none of the side effects are considered to be serious. In the latter case, it is often important to take into consideration the increased side effects when evaluating the benefit of the treatment. This approach of BR evaluation discounts the observed benefit of a treatment by the observed risk, and hence the name "benefit-less-risk." The discounting, applied to each individual in a trial, utilizes a method proposed by Chuang-Stein (Chuang-Stein 1994) to consolidate the safety data collected in a clinical trial. The collating of the safety information allows one to estimate quantitatively the risk experienced by each individual, and therefore enables the construction of a risk-adjusted benefit measure for the same individual.

Each patient's efficacy experience (benefit) of a therapy is represented by a binary response variable; "1" signifies that a therapy response is obtained, and "0" means that no response is achieved. The patient's side-effect experience (risk) from five different body functions is represented by a value ranging from 0.0 to 1.0, where the value of 1.0 represents the worst safety experience and 0.0 means no safety concern (Chuang-Stein 1994). If a pretreatment ordinal response is also available, the change in the response variable will serve as an estimate for the benefit; otherwise, the post-treatment response will be used for $\mathbf{E_i}$. To construct a risk-adjusted benefit measure $\mathbf{E^*_i}$ for individual $\boldsymbol{i}$, $\mathbf{E_i}$ has to be discounted by a suitable multiple of the $\mathbf{R_i}$, as: $E_i^* = E_i - F * R_i$, where **F** is a proportionality constant which determines how much penalty the side effects exert on the original benefit measure. The quantity $\mathbf{E^*_i}$ is the risk-adjusted benefit, or the estimated net benefit, for individual $\boldsymbol{i}$. The choice of **F** may depend on many factors (Chuang-Stein 1994).

Quality-Adjusted Time Without Symptoms and Toxicity (Q-TWiST)

Q-TWiST considers the importance of time without symptoms of disease and toxicity effects and adjusted the relevant quality (Gelber and Goldhirsch 1986; Gelber et al. 1996). This is known as quality-adjusted TWiST. The Q-TWiST is like a revised version of the quantitative comparisons of BLRA. The benefit is measured by drug attributed gain of quality-adjusted life years (QALY). Cumulative risks of toxicities and disease progression are calculated for drug attributed loss of QALY. Q-TWiST compares the relative therapeutic value of treatments based on the patient experience within the context of clinical outcomes related to cancer and its treatment (Gelber et al. 1996). This method assumes that cancer patients progress through a set of health states of varying utility value for the individual patient. This Q-TWiST has been used for risk-benefit of cancer treatment. By weighting the

durations of health states according to their quality of life, patient arrives at a single end point reflecting the duration of survival and the quality of life. Using survival analysis methods, the mean duration of each disease state and mean quality-adjusted survival time are estimated.

Number Needed to Treat (NNT) and Number Needed to Harm (NNH)

The number needed to treat NNT is defined as the number of patients who need to be treated to prevent or eliminate the occurrence of one additional event of disease of interest. NNT can be calculated by taking the reciprocal of the absolute risk reduction. $\mathrm{NNT} = \frac{1}{(P_1 - P_2)} = \frac{1}{Absolute\_Risk\_\mathrm{Re}duction}$, where P_1 and P_2 are proportions of the disease of interest in the control group and treatment group, respectively (Holden et al. 2003). In epidemiologic terms, P_1 and P_2 are risks associated with treatments and their difference is the absolute risk reduction, which can also be considered as the difference in outcome event incidence between the control and treatment groups. For example, patients with severe hyperlipidemia may receive a statin medication for a year and 8.2% mortality rate is observed compared to 11.5% mortality rate is observed for patients without receiving any statin. The absolute risk reduction is equal to (11.5% – 8.2% = 3.3%). The NNT = 1/0.033 = 30, which means 30 patients with hyperlipidemia needed to treat in order to avoid one death.

While the NNT is to measure the therapy benefit effect with a specific number of treated patients, the number needed to harm (NNH) can be calculated for adverse effects related to a specific therapy treatment. NNH can be calculated as follows: $\mathrm{NNH} = 1 / (Q_2 - Q_1)$, where Q_1 and Q_2 are the risks of AE of interest in the untreated and treated groups, respectively, and it is assumed that $Q_1 < Q_2$ because treated group has higher risk than untreated group.

Both NNT and NNH have been used widely for BRA across different therapeutic areas. A comparison between NNT and NNH can be used as a very basic comparison of benefit versus risk for a population of patients who may benefit from the treatment. Indeed, a *risk-benefit ratio* equal to NNH/NNT can be calculated between treatment and control groups. If the ratio is greater than 1 (i.e., NNH/NNT > 1 or NNT < NNH), then fewer patients need to be treated in order to achieve benefit than will be treated to have one additional occurrence of an ADE.

Minimum Clinical Efficacy (MCE)

For a new treatment, MCE can be defined as the minimal clinical efficacy needed for it to be worth considering as an alternative treatment after taking into account, including the efficacy of the standard treatment, the adverse event profiles associated with standard treatment, new treatment, and the risk of disease of interest associated with no treatment (Djulbegovic et al. 1998). MCE seeks to improve clinical care by a quantitative comparison of the potential benefit against the potential risk of a

particular treatment. It seeks to find the minimal therapeutic benefit at which a treatment is still worth administering. MCE takes into account not only the benefits and harms of the new and standard treatments but also the natural characteristics of the disease in the general population, represented by untreated group.

The relative efficacy of the new treatment as compared to the conventional treatment should be at least the same as the relative efficacy of the conventional treatment plus the difference in risk of the AE divided by the risk of the disease of interest in the untreated group. The MCE method seeks to find the minimal therapeutic benefit at which a treatment is worth administering, which can be used as a yardstick for acceptance of a new treatment alternative. The details required to balance the ADE profiles as well as efficacy impact can be extensive.

Incremental Net Health Benefit (INHB)

The incremental net health benefit (INHB) of new Drug 2 versus current therapy Drug 1 can be expressed as: $\mathbf{INHB = (E_2 - E_1) - (R_2 - R_1),}$ where effectiveness (E) is measured in quality-adjusted life years (QALYs) and risk (R) can also be measured in QALYs (Garrison et al. 2007). When $\mathbf{(E_2–E_1) > (R_2–R_1),}$ a favorable risk-benefit balance is achieved. That is, the expected QALYs gains as a result of efficacy exceed the expected losses to safety problems. The QALY represents an adjustment to length of life for the quality of life experienced. Quality of life is measured with a preference scale or index, where 0 represents the value for death and 1 represents normal health. This measure can be adapted to BRA by separating the outcomes into expected health improvements with positive QALYs (benefits) and adverse health impacts with negative QALYs (risks). Some of the literature mentioned that the benefits and risks for INHB could be measured using value-adjusted life years as opposed to QALYs (Garrison et al. 2007). The INHB approach is a theoretically sound modeling method with strong potential for usefulness in clinical and regulatory decision making.

Relative-Value-Adjusted Number Needed to Treat (RV-NNT) and RV-NNH

In order to account for patient preferences, both the NNT and NNH measures have been revised to incorporate patients' relative utility values. Relative utility values are obtained using either the standard-gamble method or the time-trade-off approach. The relative value (RV) is calculated from a numeric representation of patients' preferences for specific outcomes. Hence RV-NNT and RV-NNH can be calculated between treatment and control comparison groups. A favorable BRA outcome is obtained when RV-NNH /NNT $>$ 1 (Holden et al. 2003).

Risk-Benefit Plane (RBP) and Risk-Benefit Acceptability Threshold (RBAT)

A hypothetical model of the risk–benefit plane (RBP) is a two-dimensional plot with benefit and risk on the two axes, including four quadrants NE, SE, NW, and SW (Lynd and O'Brien 2004). The risk measurement can be incidence of ADEs or frequency of ADE. If the risk is on X-axis, the risk for the new therapy increases from left to right. The benefit measurement can be incidence of benefit or product of efficacy and responder rate. The benefit on Y-axis increases from bottom to top. Hence, risk-benefit ratios in NW depicts that experimental therapy dominates due to this treatment option with low risk and high benefit. In the SE quadrant, the active treatment option has higher risks and lower benefits (with a high benefit-risk ratio), and the control therapy is said to dominate. The remaining two quadrants involve high risk and more benefit in SW, and less risk and less benefit in NE. An appropriate risk-benefit acceptability threshold (RBAT) will be determined in BRP plot, which is indicated by a slope of line that crosses over the SW and NE quadrants (Lynd and O'Brien 2004).

Probabilistic Simulation Methods (PSM) and Monte Carlo Simulation (MCS)

Similar to the above RBP model, the average difference in the probability of achieving a benefit with the new therapy relative to conventional therapy can be plotted on the X-axis (ΔB) and the average difference in the probability of risk for the new therapy can be plotted on the Y-axis (ΔR). Both axes therefore range from −1 to 1, with 0 at the origin (Lynd and O'Brien 2004). Then, four quadrants are labeled with points of the compass NE, SE, NW, and SW. For a hypothetical model of **PSM,** using differences in the probability of achieving benefit and risk, the incremental risk-benefit ratio (IRBR) related to the new therapy can be defined as the incremental probability of an ADE (ΔR) with a new therapy relative to conventional treatment divided by the incremental probability of a beneficial effect (ΔB). The Y-axis represents average difference in the probability of risk for the new therapy versus conventional treatment. In a clinical study, an **MCS** was applied to compare the efficacy and safety of administering anticoagulants to trauma patients who are already at an elevated risk of bleeding (Lynd and O'Brien 2004).

Multi-Criteria Decision Analysis (MCDA)

MCDA is a decision tool aimed at supporting decision makers who are faced with making numerous risk and benefit evaluations. The risk can be measured by incidence of ADEs, discontinuation rate due to ADEs, and other risk factors such as potential drug interactions, off-label use leading to safety hazards, and safety issues observed in preclinical safety studies. The benefit involves clinical relevant

end-points from clinical trials and other benefit criteria. Using a decision value tree, a risk-benefit ratio for a specific drug therapy can be evaluated systematically. Both benefit and risk criteria can be split into multiple criteria in case of different primary endpoints, relevant subgroups, and relevant interactions. Although the MCDA model can be customized by adding or changing benefit and risk criteria, data extraction from clinical trials is critical for the internal validity assessment of the MCDA technique (Mussen et al. 2007).

Risk-Benefit Contour (RBC)

The RBC is to provide a two-dimensional graph showing both the probability of benefit from treatment, based, for example, on the survival rate, and the probability of drug toxicity or ADEs (the risk). The degree of drug benefit is captured along the X-axis, and the degree of drug risk is measured along the Y-axis. By finding out from each patient the amount of risk he or she is willing to accept to obtain a certain benefit, a set of individual risk-benefit contours can be determined (Shakespeare et al. 2001).

Stated Preference Method (SPM) or Maximum Acceptable Risk (MAR)

SPM or MAR is based on hedonic-utility principles, and therapeutic treatment options (commodities) over which consumers make choices. Consumer choices can be considered as a random utility function specified as $U_j = V_j + \varepsilon_j$ with $V_j = X_j\beta$ (Hauber et al. 2009), where V_j is the determinate part of the utility function for treatment j; X_j is a vector of attribute levels for treatment j; β is a vector of attribute parameters; and ε_j is a random error. Benefit-risk trade-off preferences can be estimated based on consumer experience or probability of AEs. Patient's preferences can be collected from survey questionnaires and interview techniques such as contingent valuation techniques. The current best practice standard requires participants to make trade-offs between choices using discrete choice experiments. Best-worst scaling methods are also being developed and may become more wisely used in the future. Using SPM or MAR, the risk-benefit trade-off can be calculated as the increase in risk of AEs that reduces the patients' satisfaction scores between two treatment options.

Conclusions

In summary, multiple BRAs are available for drug safety surveillance during new drug development. Various BRA methods have been utilized for clinical decision-making in different therapeutic areas, including oral contraceptives, antipsychotics, anti-hyperlipidaemia medications, cancer chemotherapy, iron-chelation, and anti-hypertensives. Above BRA methods discussed should cover mainstream BRA

methods, but may not cover all available methods due to the limited effort for literature review. All BRA methods can be utilized for safety surveillance not only in pharmaceutical and biologics but also in medical devices and other products. There are increasing safety considerations about medical devices such as implantable "Essure" device for sterilization (Johal et al. 2018), implantable metal device containing cobalt, and Davinci robotic system. Although all BRA methods have their limitations related to data requirements, statistical properties, and availability of patient preference (utility) measurement, the BRA methods are becoming more visible and useful supplemental tools for informed decision-making and regulatory interventions.

Key Facts/Points

- **Drug safety surveillance** is a multidisciplinary method to document, monitor, and evaluate adverse drug events, and plan effective interventions. It applies to the life cycle of a drug development.
- **Pharmacovigilance** is generally regarded as all post-marketing scientific and data gathering activities relating to the detection, assessment, understanding, and prevention of adverse events. It is often referred to **post-marketing surveillance**.
- **Risk management** is an iterative process designed to optimize the benefit-risk balance regulated products including assessment, development, evaluation, reassessment, and adjustment for risk-benefit balance.
- There are four key **safety signal detection algorithms** like reporting odds ratio (ROR), proportional reporting ratio (PRR), Multi-item Gamma Passion Shrinker (MGPS), and information component (IC).
- Many traditional and innovative regulatory are commonly used in USA, such as labeling changes, black-box warnings, continue educations, "Dear Doctor Letter," "Dear Health Practitioner Letter," restricted distribution, medication guides, withdrawal, etc.
- Common quantitative **benefit-risk assessments (BRAs)** include quantitative framework for risk-benefit assessment (QFRBA), benefit-less risk analysis (BLRA), quality-adjusted time without symptoms and toxicity (Q-TWiST), number needed to treat (NNT) versus number needed to harm (NNH), relative value adjusted number needed to treat (RV-NNT), minimum clinical efficacy (MCE), incremental net health benefit (INHB), risk-benefit plane (BRP) and risk-benefit acceptability threshold (RBAT), probabilistic simulation methods (PSM) and Monte Carlo simulation (MCS), multi-criteria decision analysis (MCDA), risk-benefit contour (RBC), and stated preference method (SPM) or maximum acceptable risk (MAR).
- This chapter review is based on the author's previous research work and expertise. A recent formal literature search for this review was not performed. The author is indebted to current and past graduate students and colleagues who worked together with the author for the related research subjects in drug safety and benefit-risk analysis.

Cross-References

► Bayesian Adaptive Designs

References

Almenoff JS, Tonning JM, Gould AL et al (2005) Perspectives on the use of data Mining in Pharmacovigilance. Drug Saf 28:981–1007

Bate A, Lindquist M, Edwards IR, Orre R (2002) A data mining approach for signal detection and analysis. Drug Saf 25:393–397

DuMouchel W (1999) Bayesian data mining in large frequency tables, with an application to FDA spontaneous reporting system. Am Stats 53:177–190

Chen Y, Guo JJ, Healy D, Lin X, Patel NC (2008a) Risk of hepatotoxicity associated with the use of Telithromycin: signal detection based upon the FDA's spontaneous reporting system. Annals Pharmacotherapy 42(12):1791–1796

Chen Y, Guo JJ, Steinbuck M, Lin XD, Buncher CR, Patel C (2008b) Comparisons of data mining algorithms for adverse drug reactions: an empirical study based on the adverse event reporting system of the Food and Drug Administration. J Pharm Med 22(6):359–365

Chuang-Stein C (1994) A New Proposal for Benefit-Less-Risk Analysis in Clinical Trials. *Controlled Clinical Trials* 15:30–43

Djulbegovic B, Hozo I, Fields K, Sullivan D (1998) High-dose chemotherapy in the adjuvant treatment of breast cancer: benefit/risk analysis. Cancer Control 5:394–405

Evans SJV, 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

Food and Drug Administration (FDA) (2005) FDA Guidance for industry: Premarketing risk assessment. Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment. Development and use of risk minimization action plans (RiskMAP). Online available at: https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm071696.pdf

Food and Drug Administration (FDA) (2018) Benefit-risk assessment in drug regulatory decision-making. Online available at: https://www.fda.gov/files/about%20fda/published/Benefit-Risk-Assessment-in-Drug-Regulatory-Decision-Making.pdf

FDA. (2019) Federal Food, Drug, and Cosmetic Act (FD&C Act). Kefauver-Harris Amendments revolutionized drug development. U.S. Department of Health and Human Services. Online available at: https://www.fda.gov/ForConsumers/ConsumerUpdates/ucm322856.htm.https://www.fda.gov/regulatoryinformation/lawsenforcedbyfda/federalfooddrugandcosmeticactfdcact/default.htm. Accessed on January 22, 2019

Garrison LP, Towse A, Bresnahan BW (2007) Assessing a structured, quantitative health outcomes approach to drug risk-benefit analysis. Health Aff 26(3):684–695

Gelber RD, Goldhirsh A (1986) A new endpoint for the assessment of adjuvant therapy in postmenopausal women with operable breast cancer. J Clin Oncol 4:1772–1779

Gelber RD, Goldhirsch A, Cole BF, Wieand HS, Schroeder G, Krook JE (1996) A quality-adjusted time without symptoms or toxicity (Q-TWiST) analysis of adjuvant radiation therapy and chemotherapy for resectable rectal cancer. J Natl Cancer Inst 88(15):1039–1045

Guo JJ, Curkendall S, Jones J, Fife D, Goehring E, She DW (2003) The impact of cisapride label change on codispending of contraindicated medications. J Pharmacoepidemiology Drug Safety 12:295–301

Guo JJ, Jang R, Louder A, Cluxton RJ (2005) Acute pancreatitis associated with different drug therapies among HIV-infected patients. Pharmacotherapy 25(8):1044–1054

Guo JJ, Goehring E, Jones JK (2008) The story of cisapride and its withdraw from market: a case study [book chapter 29]. In: Hartzema AG, Tilson HH, Chan AK (eds) Pharmacoepidemiology and therapeutic risk management. Harvey Whitney Books, Cincinnati, pp 727–738

Guo JJ, Pandey S, Doyle J, Bian B, Raisch D (2010) A review of current risk-benefit assessments for drug safety: report of ISPOR risk-benefit management working group. Value Health 13(5):657–666

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

Hauber AB, Mohamed AF, Johnson FR, Falvey H (2009) Treatment preferences and medication adherence of people with type 2 diabetes using oral glucose-lowering agents. Diabet Med 26:416–424

Holden WL, Juhaeri J, Dai W (2003) Benefit-risk analysis: a proposal using quantitative methods. Pharmacoepidemiol Drug Saf 12:611–616

International conference on harmonization (ICH) technical requirements for registration of pharmaceuticals for human use. (2004) ICH harmonized tripartite guideline pharmacovigilance planning E2E. November 2004. Online available at: https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E2E/Step4/E2E_Guideline.pdf. Accessed on Jan 22, 2019

Johal T, Kuruba N, Sule M, Mukhopadhyay S, Raje G (2018) Laparoscopic salpingectomy and removal of Essure hysteroscopic sterilization device: a case series. Eur J Contracept Reprod Health Care 23(3):227–230

Lis Y, Roberts MH, Kamble S, Guo JJ, Raisch DW (2012) Comparison of FDA and EMA risk management implementation for recent pharmaceutical approvals: report of the International Society for Pharmacoeconomics & outcomes research risk management working group. Value Health 15(8):1108–1118

Lynd LD, O'Brien BJ (2004) Advances in risk-benefit evaluation using probabilistic simulation methods: an application to the prophylaxis of deep vein thrombosis. J Clin Epidemiol 57:795–803

Mussen F, Salek S, Walker S (2007) A quantitative approach to benefit-risk assessment of medicines – part 1: the development of a new model using multi-criteria decision analysis. Pharmacoepidemiol Drug Saf 16:S2–S15

Pray WS, Pray GE (2011) Behind-the-counter products: a third class of drugs. US Pharmacist 36(9):11–15

Shakespeare TP, Gebski VJ, Veness MJ, Simes J (2001) Improving interpretation of clinical studies by use of confidence levels, clinical significance curves, and risk-benefit contours. Lancet 357:1349–1353

Stang PE, Ryan PB, Meng R, Racoosin JA, Overhage JM, Hartzema AG et al (2010) Advancing the science for active surveillance: rationale and design for the observational medical outcomes partnership. Ann Int Med 153:600–606

Zeldis JB, William BA, Thomas SD, Elsayed ME (1999) S.T.E.P.S. a comprehensive program for controlling and monitoring access to thalidomide. Clin Therap 21(2):319–330