
Jeff J. Guo, PhD, 1 Swapnil Pandey, MS, 2 John Doyle, PhD, 3, 4 Boyang Bian, MS, 1 Yvonne Lis, PhD, 4 Dennis W. Raisch, PhD 5

1University of Cincinnati Health Academic Center, College of Pharmacy, Cincinnati, OH, USA; 2Kendle International Inc., Cincinnati, OH, USA; 3Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, NY, USA; 4Centre for Socioeconomic Research, Welsh School of Pharmacy, Cardiff, UK; 5University of New Mexico, College of Pharmacy, Albuquerque, NM, USA

ABSTRACT

Objective: Although regulatory authorities evaluate the risks and benefits of any new drug therapy during the new drug-approval process, quantitative risk–benefit assessment (RBA) is not typically performed, nor is it presented in a consistent and integrated framework when it is used. Our purpose is to identify and describe published quantitative RBA methods for pharmaceuticals.

Methods: Using MEDLINE and other Internet-based search engines, a systematic literature review was performed to identify quantitative methodologies for RBA. These distinct RBA approaches were summarized to highlight the implications of their differences for the pharmaceutical industry and regulatory agencies.

Results: Theoretical models, parameters, and key features were reviewed and compared for the 12 quantitative RBA methods identified in the literature, including the Quantitative Framework for Risk and Benefit Assessment, benefit–less–risk analysis, the quality-adjusted time without symptoms and toxicity, number needed to treat (NNT), and number needed to harm and their relative-value-adjusted versions, minimum clinical efficacy, incremental net health benefit, the risk–benefit plane (RBP), the probabilistic simulation method, multicriteria decision analysis (MCDA), the risk–benefit contour (RBC), and the stated preference method (SPM). Whereas some approaches (e.g., NNT) rely on subjective weighting schemes or nonstatistical assessments, other methods (e.g., RBP, MCDA, RBC, and SPM) assess joint distributions of benefit and risk.

Conclusions: Several quantitative RBA methods are available that could be used to help lessen concern over subjective drug assessments and to help guide authorities toward more objective and transparent decision-making. When evaluating a new drug therapy, we recommend the use of multiple RBA approaches across different therapeutic indications and treatment populations in order to bound the risk–benefit profile.

Keywords: drug safety, incremental risk–benefit ratio, multicriteria decision analysis, number needed to treat, risk–benefit assessment, risk–benefit plane, stated preference method.

Introduction

In the past decade, over a dozen high-profile brand-name drugs, including rofecoxib, troglitazone, cisapride, and cerivastatin, were withdrawn from the market [1]. The US Food and Drug Administration (FDA) established a Drug Safety and Risk Management Division, which is charged with evaluating the safety, efficacy, and abuse potential of drugs, as well as risk management and risk communication. 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 Pharmacoeconomic Assessment [2–5]. In 2007, Title IX of the FDA Amendments Act gave the FDA the authority to require companies to develop and implement a risk evaluation and mitigation strategy for specified prescription drugs. In this era of renewed focus on drug safety, the FDA has called for more creative approaches to conceptualizing, measuring, and applying risk–benefit assessment (RBA) techniques to develop and improve a systematic RBA approach throughout the life cycle of a pharmaceutical product [6–9]. Appropriate RBA 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 [8–12].

In Europe, part of the mandate of the Committee for Medicinal Products for Human Use (CHMP) is to assess risks and benefits of authorized medicines on behalf of the European Medicines Agency (EMEA). In 2007, the CHMP revised its guidance and included quantitative RBA in the regulatory agenda with the publication of a report examining the potential value of existing benefit–risk models and methods [13, 14]. Although no specific method was recommended, several RBA features were noted as being of value, including 1) all important benefits and medically serious risks are identified; and 2) the risks and benefits are weighted according to their relative importance and the strength of the evidence available. It was also decided that a comprehensive review of available quantitative methods for RBA relevant to the CHMP was required to explore further development of tailored methodologies. The EMEA recently created the European Network of Centres for Pharmacoeconomics and Pharmacovigilance, which is in the process of developing an algorithm to articulate safety and benefit profiles for pharmaceutical products.

Regulatory agencies use various methods to discover rare toxic events. These include review of data from randomized controlled clinical trials, observational epidemiological studies (case-control, cohort, and cross-sectional analyses), drug-use surveys, automated databases linking drugs and disease, spontaneous reporting (passive surveillance, such as the FDA’s...
MedWatch program), and established patient registries [15,16]. When a regulatory authority such as the FDA or the EMEA obtains information about potentially significant drug toxicity, it thoroughly reviews the original new drug application data that were used for the initial approval. Concurrently, a formal analysis is initiated by reviewing all the spontaneous reports available from the FDA Adverse Events Reporting System, the United Kingdom’s Yellow Card Scheme, and the World Health Organization’s Uppsala Monitoring Centre. Additionally, the sponsor of a pharmaceutical product is required to review its safety database and report directly to the regulators. In some circumstances, further postmarketing studies are designed to answer specific questions about toxicity. In the process of reviewing risks and benefits, a regulatory authority typically seeks input from advisory committees that review safety and efficacy data and make recommendations [8,9,17]. There is a structured process for convening the appropriate advisory committee. The authority prepares a set of questions involving product safety, efficacy, study design, results interpretation, and risk–benefit profiles for committee discussion. After review of the advisory committee’s recommendations, the authority determines a course of action, which may include changes to labeling, direct correspondence to health professionals, or removal of the drug from the market.

This traditional process does not produce an explicit, consistent, transparent, and aggregate quantification of the risks and benefits and lacks clarity pertaining to the role of specific factors in the recommendations. Although some quantitative measures, such as quality-adjusted life-years (QALYs) and number needed to treat (NNT) [11,17–21], discussed later in this article, are used by regulators, there is lack of standardized and validated quantitative methodology. Challenges to developing such a methodology include heterogeneity and multiplicity of risks and benefits, uncertainty regarding attribution of risks and benefits to a particular treatment, and the temporality and paucity of drug-exposure and outcome data [18–23]. Furthermore, the traditional process does not allow for systematic reassessment of risks and benefits over time. Although an innovative drug may initially possess advantageous risk–benefit ratios versus older drugs, these ratios often change over time, as occurred, for example, with COX-II inhibitors.

Although none of the major regulatory agencies has a clear benchmark for what constitutes an acceptable level of risk, nevertheless, the pharmaceutical industry is functioning in an era of increased risk assessment, which requires proactive drug-safety analysis and additional commitment to safety. Several dozen risk-management programs or registries have been established by major pharmaceutical companies in the United States, including iPLEDGE® to prevent exposure to isotretinoin during pregnancy, the clozapine patient registry to prevent agranulocytosis, an alosetron prescribing program for reducing the risk of severe gastrointestinal adverse events, and the STEPS® program to provide risk and/or benefit assessment of drugs. Excluded were conference programs, educational catalogues, announcements in trade journals, veterinary-medicine publications, animal studies, ADR case reports, abstracts without specific information on risk or benefit assessment, and general discussions related to risk or benefit assessment for medical devices or biotechnology.

The perspective on RBA was not limited to that of any one stakeholder. Although both qualitative and quantitative RBA studies were retrieved from the literature, we reviewed only those using quantitative methods.

**Methods**

An extensive literature search was first performed to identify qualitative and quantitative approaches to RBA in drug evaluations. The MEDLINE and Cochrane Library databases were used as well as other Internet-based search engines to find scientific publications, books, and academic conference proceedings. Country-specific drug regulatory Internet Web sites were also searched. Key words included risk, benefit, risk–benefit assessment, postmarketing surveillance, drug safety surveillance, adverse drug reactions (ADRs), and pharmacovigilance. Our inclusion criteria were 1) the publications were in English; and 2) they focused on pharmaceutical or drug safety surveillance and provided risk and/or benefit assessment of drugs. Excluded were conference programs, educational catalogues, announcements in trade journals, veterinary-medicine publications, animal studies, ADR case reports, abstracts without specific information on risk or benefit assessment, and general discussions related to risk or benefit assessment for medical devices or biotechnology.

The perspective on RBA was not limited to that of any one stakeholder. Although both qualitative and quantitative RBA studies were retrieved from the literature, we reviewed only those using quantitative methods.

**Results**

Based on the key words, over 12,000 papers were identified. Using our inclusion and exclusion criteria, over 300 abstracts and Web site summaries related to RBA for drug safety were identified and reviewed. Further selection of only quantitative RBA resulted in 59 articles, which were used as the basis for this article. Figure 1 summarizes the process of study selection. Because of limited published information regarding net clinical benefit analysis, the principle of threes, and net-benefit-adjusted-for-utility analysis [10,21,37], we excluded these methods. Remaining were 12 quantitative RBA methodologies, which are individually described below and summarized in Table 1. They are described roughly in chronological order.

**Quantitative Framework for Risk–Benefit Assessment (QFRBA)**

Standard notions of risk and benefit are well documented in the literature as part of a basic quantitative framework (QFRBA). The QFRBA attempts to quantify these notions directly, and more recent RBA techniques build on these same ideas.

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 quantitative expression of risk is the incidence of adverse events, i.e., the number of new ADEs in a defined population over a specific period of time divided by the population at risk for the adverse event over this same time period [27,28,38,39]. Relative risk (RR) takes into account differences in exposure to the drug and is defined as the ratio between the proportions of exposed individuals who experience an ADE divided by the proportion of unexposed individuals who experience an ADE [38–41]. For example, whereas osteoporotic individuals are more likely to experience a fracture, there is some likelihood that people without osteoporosis will fracture as well, and this fact is accounted for in the RR denominator. Attributable risk (AR), otherwise known as risk difference, uses the same raw information as in the computation of RR but measures the absolute difference between the risk of ADEs between those exposed to the drug and those who were not and then divides this difference by the absolute difference between the size of the two...
The population-based attributable risk (PAR) is an extension of AR that uses total population data. PAR is particularly important for public-health decision-makers.

The benefit associated with a medication might be its ability to reduce either an ADE or another adverse event associated with the disease itself. In this case, relative risk reduction (RRR) is useful and is defined as 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 [27,28,38,40,43]. For instance, again in the case of osteoporosis, RRR rates are computed for different anti-osteoporosis medications based on the number of fractures averted by the treatments. Similarly, attributable risk reduction, otherwise known as absolute risk reduction (ARR), represents the decline in adverse events between exposed and unexposed groups [27,28,38,43]. However, “benefit” does not always mean absence of adverse events. From clinical trials, benefit measurements might include clinically relevant efficacy parameters such as specific biomarkers, surrogates, and putative surrogate end points [44–46]. From observational studies, benefit measurements could include the patient medication and treatment adherence rate or population-based treatment effectiveness, among many other possibilities.

The QFGBRA methodology is widely used in drug safety surveillance by regulatory agencies and by the pharmaceutical industry. It has a long history and a solid foundation. The QFGBRA does not, however, provide a methodology for combining risks and benefits into a single value that could potentially be used to compare risk-benefit profiles between alternative therapies.

Benefit-Less-Risk Analysis (BLRA)

For BLRA, risk and benefit relationships are presented as risks subtracted from benefits using weights assigned to one of five benefit–risk outcome categories: 1) efficacy without side effects; 2) efficacy with side effects; 3) no efficacy and no side effects; 4) no efficacy with side effects; and 5) side effects leading to withdrawal from an ongoing treatment or intolerable side effects [47,48]. The risk-adjusted benefit measure for each individual is obtained from the difference between an aggregate benefit score and an aggregate risk score.

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. Notably, these categories are primarily qualitative, but weights can be chosen by individuals to reflect the relative importance of the five categories. A mathematical model is derived from these weights and categories to calculate the observed benefit-less-risk score for an individual. If applied in a clinical trial comparing different treatments, statistical significance tests can be performed to compare treatments. The interpretation of the results remains focused on the individual.

The weighting method used is subject to differences in interpretation between individuals. There may be a question regarding the validity of these assessments because the methods used to present the relative weights may affect the outcome. The conversion from a qualitative assessment into a quantitative value may also be subject to bias. The advantage of BLRA is that it provides a structure for combining risks and benefits into a single measure.

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

Another risk–benefit measure is Time Without Symptoms and Toxicity (TWiST). Time lost due to ADEs or toxicity is subtracted from time gained through treatment. A quality-adjusted version of TWiST, known as Q-TWiST, converts time into QALYs; the QALYs lost due to ADEs or toxicity are subtracted from QALYs gained from treatment [26,49–51]. The benefit is measured by drug-attributed gain in QALYs. Cumulative risks of toxicities and disease progression are calculated to obtain drug-attributed loss of QALYs (risk). Q-TWiST compares the relative therapeutic value of treatments based on the patient’s experience within the context of clinical outcomes related to a disease and its treatment. This method assumes that patients progress through a series of health states with varying quality of life. This RBA method has been used for risk–benefit analysis of cancer treatments [50,52,53], as well as other therapies and transplantation procedures [54–56]. By weighting the durations of health states by their respective utility values, a patient arrives at a single end point reflecting the duration of survival and the quality of life. Using survival analysis techniques, the mean duration of each
Table I  Summary of twelve quantitative risk–benefit assessment techniques

<table>
<thead>
<tr>
<th>Seq. No.</th>
<th>Title and quantitative approach</th>
<th>Parameters for assessment</th>
<th>Theoretical model and key features</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Quantitative Framework for Risk and Benefit Assessment (QFRA) [27,28,38–40,42–46]</td>
<td>Risk and benefit are defined and quantified separately; Risk focuses on adverse events or outcomes, defined as relative risk, attributable risk, population attributable risk. Benefit focuses on risk differences (i.e., relative risk reduction, absolute risk reduction).</td>
<td>Theoretically sound quantitative method; Probability driven assessment for risk of adverse events and benefits of improved outcomes; Relatively simple calculation; Often used for drug safety surveillance in industry and regulatory agencies.</td>
</tr>
<tr>
<td>2</td>
<td>Benefit-risk analysis (BLRA) [47,48]</td>
<td>Intensity scores are used to compare severity and frequency of adverse drug events (ADEs) and assigned for each patient. Data on observed benefit from the treatment are required. Proportionality constant determines how much penalty the ADEs offset benefit measure.</td>
<td>Simple empirical method with sound theoretical basis; Differences between treatments can be statistically analyzed (t-test or ANOVA); Requires subjective rankings of ADE intensity scores; Patient preferences are incorporated using a discounting process; Subjective ranking for ADEs and proportionality a potential threat to internal validity.</td>
</tr>
<tr>
<td>3</td>
<td>Quality-adjusted Time Without Symptoms and Toxicity (Q-TWiST) [49–57]</td>
<td>Benefit measured as drug-attributed gain in quality-adjusted life-years (QALY); Benefit measured as drug-attributed loss of QALY. Compare gain versus loss of QALY.</td>
<td>Statistical method can be conducted to compare alternative treatments; QALY incorporates patients preference measurement changes over time; Validity of QALY measurements and differences between techniques are potential concerns; Widely used risk–benefit assessment technique in oncology.</td>
</tr>
<tr>
<td>4</td>
<td>Number needed to treat (NNT) and number needed to harm (NNH) [27,28,58–60]</td>
<td>Benefit measure: number of persons treated (NNT) to avoid one person developing disease of interest (absolute risk reduction, Relative risk reduction); Risk measure: number of persons treated when one person experiences ADE (NNH). Ratio of NNT and NNH.</td>
<td>Well-defined quantitative framework; Simple calculation for RBA used for comparing treatment and control groups.</td>
</tr>
<tr>
<td>5</td>
<td>Relative risk adjusted number needed to treat (RV-NNT) [27,28]</td>
<td>Expands NNT to incorporate relative utility values (RV) based upon patient preferences; RV-NNHR can be also be determined; Ratio of RV-NNHR and RV-NNH.</td>
<td>Lack of strong statistical properties; Risk-benefit relation can be compared directly by NNT to NNH ratio; NNH should be lower than RRR for the drug to be valuable in term of risk benefit ratio; Risk factors are used for RBA in different therapy areas; Difficult to incorporate more than one outcome (A and benefit reduction) simultaneously;</td>
</tr>
<tr>
<td>6</td>
<td>Minimum clinical efficacy (MCE) [27,28,61]</td>
<td>Benefit measure: efficacy difference between new treatment and conventional treatment or placebo; Harm measure: probability of AE (risk) in patients receiving new treatment vs. conventional treatment or placebo; Relative utility values may be considered.</td>
<td>Relative value of ADE versus benefit not considered.</td>
</tr>
<tr>
<td>7</td>
<td>Incremental net health benefit (INHB) [17,62]</td>
<td>Risk measured as decrease in QALY; INHB as relative gain or loss of QALY due to treatment versus usual care or placebo.</td>
<td>Statistical method can be conducted to compare alternative treatments; QALY incorporates patients preference measurement changes over time; Validity of QALY measurements and differences between techniques are potential concerns;</td>
</tr>
<tr>
<td>8</td>
<td>Risk–benefit plane (RBP) and risk–benefit acceptability threshold (RBAT) [18,64]</td>
<td>Risk measured as relative probability of AEs between treatment and control groups; Benefit measured as relative probability response between treatment and control groups; Visual application of risk–benefit comparisons.</td>
<td>Two-dimensional plot with benefit measurement on x-axis, risk measurement on y-axis; An acceptable threshold of relative risk–benefit ratio can be plotted to visually compare with other treatments; Often used for explaining the phenomenon of drug safety surveillance.</td>
</tr>
<tr>
<td>9</td>
<td>Probabilistic simulation methods (PSM) and Monte Carlo simulation (MCS) [18,62–64]</td>
<td>Average difference in the probability of risk and benefit for the new therapy relative to conventional therapy; Incremental risk–benefit ratio (IRBR).</td>
<td>Framework applies incremental risk–benefit ratio; Joint density of benefits and risk scatter plot can be presented with a risk–benefit acceptability curve;</td>
</tr>
<tr>
<td>10</td>
<td>Multicriteria decision analysis (MCDA) [45,46]</td>
<td>Benefit measured as clinically relevant endpoints from clinical trials; Risk measured as incidence of ADE, discontinuation rate, drug interactions, and other risk criteria; Decision tree model is developed to incorporate all key risks and benefits; Relative weights are of risks and benefits are assigned.</td>
<td>Decision tree describes clinical outcomes in a hierarchical manner; Decision-making tool incorporates evaluation of both drug’s risks and benefits; Relative scores for alternative treatments can be calculated based on modeling; Data extraction from clinical trials are critical for internal validity; Missing data and uncertainty can be addressed.</td>
</tr>
<tr>
<td>11</td>
<td>Risk–benefit Contour (RBC) [46]</td>
<td>Probability of potential benefit of treatment such as an increased survival rate; Probability of potential risk due to severe ADE or drug toxicity.</td>
<td>A decision tree describes clinical outcomes in a hierarchical manner; Decision-making tool incorporates evaluation of both drug’s risks and benefits; Relative scores for alternative treatments can be calculated based on modeling;</td>
</tr>
<tr>
<td>12</td>
<td>Stated preference method (SPM) or maximum acceptable risk (MAR) [67–77]</td>
<td>Relative utility values for therapeutic treatment alternatives; Vector of attribute levels for treatment options; Benefit–risk trade-off preferences estimated by probability of severe AEs versus benefit in terms of treatment success; Patient surveys required to provide data regarding value of benefit versus negative impact of risk.</td>
<td>Subjective ranking for ADEs and proportionality a potential threat to internal validity.</td>
</tr>
</tbody>
</table>
disease state and the mean quality-adjusted survival time can be estimated [52–57].

As for BLRA, QTwiST provides a summary measure incorporating both benefits and risks over time. Quality-adjusted survival is an understandable concept and can help individual patients make informed treatment decisions. As with any QALY-based methodology, concerns regarding the validity and application of QALYs in a broad decision-making context are inherent with the method. Additionally, the type of QALY-measurement technique may influence the results.

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

NNT is defined as the number of patients who need to be treated in order to prevent one more occurrence of a disease [27,28,39,58–60]. NNH can be calculated by taking the reciprocal of the ARR, and, as such, is dependent upon the incidence of the disease for calculating ARR. Treatment is warranted if NNT is low enough so that benefits accrue at a sufficiently low number of patients before ADEs occur [27,28,58–60].

NNH is calculated similarly for ADEs related to a specific therapy, thus describing how many patients will be treated before a patient will experience an ADE [27,28,58–60]. Both NNT and NNH have been used widely for RBA across different therapeutic areas [29,30,33,34,58,59]. 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 to achieve benefit than will be treated to have one additional occurrence of an ADE [27,28]. Rather than compare NNT and NNH explicitly, NNT can be adjusted to account for ADE potential. **Adverse-event-adjusted NNT** is defined as the number of patients who need to be treated to prevent one additional disease outcome after adjusting for the potential risks (e.g., ADEs) of the treatment, that is, after allowing for the probability of incurring a treatment-related adverse event [27,28].

Both NNT and NNH are widely used in clinical practice because they are simple decision-making tools for clinicians. Indeed, these tools currently have a ubiquitous role in medical decision-making. However, NNT and NNH by themselves do not account for the patients’ value of the benefit or the value of the harm, respectively. In addition, it is cognitively complex to incorporate several benefits and harms simultaneously, which commonly exists in clinical situations.

**Relative-Value-Adjusted- (RV) NNT and RV-NNH**

In an attempt to account for patient preferences, both the NNT and NNH measures have been revised to incorporate patients’ relative utility values, in this way correcting for one of the problems mentioned above. Adding a patient’s relative utility values to the risk–benefit evaluation should make this RBA method more robust [27,28]. Relative utility values are obtained using either the standard-gamble method or the time-trade-off approach. A metric, RV, is calculated from a numeric representation of patients’ preferences for specific outcomes. Specifically, $RV = (1 – utility of adverse event)/(utility of improvement using a specific treatment)$, where the value of 1 represents the value of perfect health. RV can then be interpreted as a quantification of the overall value of patients’ preferences for avoiding an ADE relative to avoiding the disease of interest or target event.

NNH adjusted for relative value can also be calculated similarly to RV-NNT. Thus, a **risk–benefit ratio** (RV-NNH/NNT) can be also calculated between treatment and control groups. A favorable risk–benefit outcome is obtained when RV-NNH/NNT > 1. This technique overcomes one of the weaknesses of NNH and NNT. It may also be possible to incorporate several benefits and risks simultaneously using this technique.

**Minimum Clinical Efficacy (MCE)**

The MCE of a treatment is defined as the minimal clinical efficacy required by the treatment in comparison with a standard treatment, after taking into account the efficacy of the standard treatment, the adverse event profiles associated with both the standard treatment and the treatment under consideration, and the risk of the disease of interest associated with no treatment [27,28,61]. A broader definition of benefit as the efficacy difference between two treatments (the new treatment versus the standard treatment) is calculated for RRR for the disease of interest. Risk of an adverse event among persons treated with the new treatment versus the standard treatment can also be calculated [27,28,61]. A new treatment is warranted over the standard treatment if the difference in risk is less than the efficacy difference. MCE seeks to improve clinical care through a quantitative comparison of the potential benefit against the potential risk of a particular treatment. 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 an 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 the acceptance of a new treatment alternative. The details required to balance the ADE profiles as well as efficacy impact can be extensive. This method is similar to the RV-NNT model [27,28], which cannot easily handle uncertainty in the measurement of benefits or risks.

**Incremental Net Health Benefit (INHB)**

Risk–benefit differentials can be expressed as either ratios or differences, although the latter are more mathematically manageable if the units of measurement are the same. The INHB accomplishes this through the application of QALYs. The INHB of a treatment in comparison with standard treatments is calculated as the incremental difference between effectiveness and risk changes; i.e., \( INHB = (E_2 - E_1) - (R_2 - R_1) \), where \( E \) denotes effectiveness, and \( R \) is risk [17]. Both risk and benefit must be measured using QALYs. A favorable risk–benefit balance exists when \((E_2 - E_1) > (R_2 - R_1)\). That is, the expected QALY gains as a result of efficacy must exceed the expected losses from risk (ADEs) in order for the net health benefit to be positive [17,62]. Some of the literature mentioned that the benefits and risks for INHB could be measured using value-adjusted life-years as opposed to QALYs [63].

The INHB approach is a theoretically sound modeling method with strong potential for usefulness in clinical and regulatory decision-making. As with QTwiST, the potential problems with this technique rely upon the inherent difficulties with the use of QALYs as outcome measures. Furthermore, these measurements must be accrued over time to gather sufficient data. There may be benefits or risks that occur after the time of original data collection.

**Risk–Benefit Plane (RBP) and Risk–Benefit Acceptability Threshold (RBAT)**

For this method, the benefit–risk ratio can be interpreted as the increase in the expected number of patients who will benefit for
each additional ADE that occurs—from new treatment rather than usual care [64]. \( R = \frac{(P_E - P_C)}{(q_E - q_C)} \), where the probabilities of benefit from the experimental treatment and in the control arm are \( P_E \) and \( P_C \), respectively, and the probabilities of risk from the experimental treatment and in the control arm are \( q_E \) and \( q_C \), respectively [17,63].

Figure 2 shows a hypothetical model of the RBP, which is a two-dimensional plot with benefit and risk on the two axes, including four quadrants NE, SE, NW, and SW [17,64]. The risk measurement can be incidence of ADEs or frequency of ADE. If the risk is on the 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 the y-axis increases from bottom to top. Hence, risk–benefit ratios in NW depicts that experimental therapy dominates because of 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 risk–benefit 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 RBP plot, which is indicated by a slope of line that crosses over the SW and NE quadrants [15,62].

This method is a well-defined hypothetical model and offers a visual tool to make complex comparisons between risk and benefit. Some ambiguity is associated with collapsing benefits and risks into single measures. Because of two-dimensional model for RBP, it is unclear how one might incorporate multiple dimensions of risks and benefits. It may be useful to use these types of graphical representations in patient decision-making regarding alternative treatments.

### Probabilistic Simulation Methods (PSMs) 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 (\( \Delta B \)), and the average difference in the probability of risk for the new therapy can be plotted on the y-axis (\( \Delta R \)). Both axes therefore range from -1 to 1, with 0 at the origin. Then, four quadrants are labeled with points of the compass NE, SE, NW, and SW [17,62,64]. Because the benefit increases from left to right along the x-axis, positive values (to the right of the vertical axis) represent greater benefits with the new treatment. Similarly, positive Y-coordinates indicate a greater probability of the risk for the new treatment.

Figure 3 demonstrates a hypothetical model of PSM. Using differences in the probability of achieving benefit and risk, the incremental risk–benefit ratio related to the new therapy can be defined as the incremental probability of an ADE (\( \Delta R \)) with a
new therapy relative to conventional treatment divided by the incremental probability of a beneficial effect (ΔB) [18]. The y-axis represents the relative difference in the probability of risk for the new therapy versus conventional treatment.

There are two potential methods for quantifying the joint density of the uncertainty around the risks and benefits, depending on the availability of patient-level data [18]. If data are available, a nonparametric bootstrap sample of data can be selected repeatedly; if original data are not available, a simulation can be run using information on the distributions fit to the data. From the simulations or bootstrap estimates, the incremental risk–benefit pairs can be plotted on the RBP risk–benefit. For example, in a randomized, double-blind, controlled clinical trial, low-dose unfractionated heparin was compared with a low molecular weight heparin, enoxaparin, for the prophylaxis of venous thromboembolism following major trauma [63]. In a previous clinical study, a MCS was applied to compare the efficacy and safety of administering anticoagulants to trauma patients who are already at an elevated risk of bleeding [18].

RBP permits the estimation of the joint density of risks and benefits with their associated uncertainty and facilitates estimation of the probability that a therapy is net-beneficial. MCS can be used to compare drugs for both efficacy and safety. These tools rely upon more complex statistical techniques to describe relationships between treatments similar to RBP. The same limitations exist, but the statistical modeling provides a greater level of confidence in the findings, and provides the data used to generate the results are valid and are collected reliably.

**Multicriteria 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 [45,46]. The benefit involves clinically 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 end points, relevant subgroups, and relevant interactions [45,46]. Benefit criteria can be split into one or more pivotal trials and other benefit criteria. Each pivotal trial may include efficacy related to primary end points in overall subjects and in relevant subgroups, as well as efficacy nonprimary end points. Other benefit criteria may include efficacy from nonpivotal trials and anticipated patient compliance in clinical practice. Risk criteria are split into adverse effects, other risk criteria, and potential for nondemonstrated additional risks. Each adverse effect involves its incidence and discontinuation rate. Other risk criteria involve safety in subgroups, interaction with other drugs and food, and potential for off-label use leading to safety hazards. Potential for nondemonstrated additional risk involves long-term exposure safety profile, safety issues observed in preclinical studies, and safety issues observed in the same pharmacological class [45,46].

MCDA can handle missing data and uncertainty using appropriate modeling and relative weights. There are several key steps for MCDA. First, the decision context (tree) needs to be established with defined options. Clear objectives and criteria should be identified for assessing the consequences for each option with high-level and low-level rates of occurrence in the hierarchy. Then, the expected performance scores of each option against the criteria of consequences should be assessed. Finally, weighted scores at each level in the hierarchy and overall weighted scores can be calculated. The result can be examined with sensitivity analysis [45,46].

MCDA is a relatively new model that takes in account multiple criteria, judgment, and uncertainty. Although the MCDA model can be customized by adding or changing benefit and risk criteria, the process may be too burdensome for limited evaluations. Data extraction from clinical trials is critical for the internal validity assessment of the MCDA technique.

**Risk–Benefit Contour (RBC)**

The RBC is another relatively new method that provides 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) [66]. 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 RBCrisk–benefits can be determined. For example, the RBC can show whether a patient is willing to accept a 40% probability of acute toxicity as long as there is a 5% improvement in the chance of survival [66].

The RBC approach is a way of formalizing the risk–benefit trade-off for patients and, as such, can be a useful tool for patients and physicians. It can take information from clinical trials in order to present the risk–benefit trade-off visually. However, as unique to each patient, the, RBC method does not have an intrinsic risk–benefit threshold and does not allow for multiple risks and multiple benefits from a drug therapy.

**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 $U_i = V_i + \varepsilon_i$ with $V_i = X_i\beta$ [67], where $V_i$ is the determine part of the utility function for treatment $i$, $X_i$ is a vector of attribute levels for treatment $i$, $\beta$ is a vector of attribute parameters, and $\varepsilon_i$ is a random error. Benefit–risk trade-off preferences can be estimated based on consumer experience or probability of adverse events (AEs). Patients’ 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. This RBA method has been applied to several therapeutic areas such as Crohn, HIV, type 2 diabetes, multiple sclerosis treatments, and vasomotor symptom relief [68–77].

SPM or MAR is theoretically sound and uses similar techniques to contingent valuation. It requires the collection of patients’ treatment preference, which allows the evaluation of benefit–risk trade-offs. The method can demonstrate the patient’s willingness to accept risks to the benefits of controlling disease symptoms.

**Discussion**

Multiple scientific methods that quantify benefits and risks of medical treatments are available and can be used to reduce the
current subjective onus on regulatory agencies, thus guiding them to more objective and transparent decisions regarding drug efficacy and safety. Each assessment method requires different data and has its unique characteristics and features, as well as strengths and weaknesses. The reviewed methodologies for RBA are representative of technical advances in the field of quantitative analysis. So far, they have been used primarily for research demonstrations and to showcase their scientific and mathematical feasibility. They have not yet been systematically adopted by regulatory agencies or by the pharmaceutical industry. The reviewed RBA methods might be helpful for making decisions at the population level. In addition, they are clearly valuable in providing information for individual patients and their physicians in disease treatment decisions.

In trying to develop a consistent RBA strategy, Lynd (Lynd L, unpublished presentation) suggested that desirable features include 1) universal applicability across medications; 2) inclusiveness (taking account of all risks and benefits); 3) patient sensitivity (person-specific preference weightings); 4) straightforward interpretability; 5) flexibility; 6) adaptability; 7) ability to incorporate uncertainty; 8) integrability with economic evaluations of medications; and 9) ability to be consistent with a quantitative risk–benefit threshold. Some methods like INHB, RBP, PSM, MCDA, SPM, and RBC seem to be flexible and easier to adopt for certain clinical situations than others. Although some have argued that INHB is the best RBA because it meets all nine criteria, it is too early to conclude that any single method is suitable for all purposes [2,8–10,21,66,78]. From ongoing academic activities at both the FDA and the EMEA, three RBA approaches have received the most attention: INHB using health-outcomes modeling, MCDA, and SPM (comments from an expert reviewer which is based on communication with the FDA and the EMEA).

Because RBP and PSM provide health-care decision-makers with a functional relationship between patient preference and probability of net benefit, they may be two promising methods for forming a standardized approach to modeling risk–benefit relationships. Significant research, however, is required to determine probability distributions for incremental risks and benefits. Further validation using multiple data sets and various simulation techniques may be required to generate comparative results. On the other hand, BLRA may be most appropriate for comparisons between treatments in the same therapeutic class because it takes both benefit and risk into consideration as long as the risks and benefits are available in similar qualitative measurements. Nevertheless, BLRA does not clearly delineate benefit and risk, and interpretation is very complicated.

We note that both RV-NNT and MCE analysis rely heavily upon efficacy or effectiveness data, adverse event rates, and patient preference (utility) data. Patients must specify their relative preferences for an outcome given potential risks. These data may be difficult to obtain. Furthermore, NNT has been criticized for its lack of strong statistical properties. When the ARR value gets closer to 0, NNT will be nearly infinity and therefore hard to interpret [27,28]. MCE analysis allows for determining the worth of a new treatment relative to an older one, given not only the potential risks of adverse events and benefits that may be gained, but also taking into account the risk of disease without any treatment. However, statistical properties for MCE are unstudied, and it is not clear whether a valid confidence interval can be created around an MCE value [27,28].

The QALY is often used to measure disease burden and provides the foundation for cost-effectiveness analysis despite the difficulties in measuring QALYs lost or gained through adverse events or successful therapy, respectively. QALYs are used for both INHB and Q-TWiST; benefit is measured by expected health improvements (QALY gains), whereas risk is measured by adverse health impacts (QALY losses). QALY estimates for known potential side effects may be inferred from knowledge of the product’s mechanism of action or signals from clinical trials. Information about rare events may be nonexistent when a drug is first marketed before wider exposure to the drug by patients. Presumably, one could rely on broad historical experience in order to include an explicit but very small probability for a rare event [17,62]. In the absence of obtainable utilities from the literature or other studies, primary data collection may be required to determine these values.

We are not suggesting that any or all of the quantitative RBA approaches reviewed here should replace the existing decision-making process of either clinicians or regulatory agencies; rather, they should serve as supplemental tools to inform such decision-making. Quantitative RBA methodologies may be particularly beneficial in an era of limited health-care resources for making appropriate decisions about pharmaceutical products. All reviewed RBA methods have limitations, predominantly in their data requirements, statistical properties, and the availability of valid patient preference measures (utilities). In many situations, we are limited to retrospective analysis of historical data and existing databases.

**Limitations and Future Research**

The present study was limited to review of existing, published scientific RBA techniques by using a Web-based literature review of peer-reviewed journal articles found through PubMed/MEDLINE, the Cochrane Library, the FDA repository, and other Internet-based search engines. Many other risk–benefit approaches based on specific clinical parameters or specific perspectives were not included [79–84]. There may also be other quantitative approaches that regulators have not formalized or published. Further research is warranted to ensure the robustness of these latter measures and to define the appropriate context in which they should be applied. Our review of RBA methods is consistent with the observation that technical development of risk–benefit analysis is far from what is needed in the contemporary regulatory environment of pharmaceutical product development.

**Conclusion**

Several scientifically sound, quantitative RBA methodologies can be used to reduce the current subjective onus on the regulatory agencies and help guide them toward making more objective, transparent, and evidence-based decisions regarding drug risks and benefits. Each quantitative method has its unique advantages and disadvantages based on data requirements and statistical properties. Whereas some methods rely on subjective weighting schemes or simple statistical assessments (e.g., NNT and NNH), other methods are used for obtaining joint distributions of benefit and risk (e.g., RBP, PSM, MCDA, SPM, and RBC) and have a sound mathematical basis. Methods that incorporate the patient’s risk tolerance and preference for health states may represent a promising channel of study in this area. Numerous RBA methodologies have been proposed, but there are a limited number of empirical applications of these techniques, and there is no consensus among regulators necessary for defining a clear gold standard. More testing, and particularly head-to-head comparisons, is needed if these approaches are to receive serious consideration. When evaluating any new health-care technology, we recommend the use of multiple RBA approaches across...
different therapeutic indications and treatment populations to bound the risk–benefit profile.

Acknowledgements

The authors would like to thank both Pamela C. Heaton and Azhar Choudhry for their suggestions and advices for the early research report. We would like to thank Dr. Christina ML. Kelton for her language edits. We recognize the efforts of Ashish Parekh, Shital Kamble, Anthony Lockett, Kara Suter, and many committee members from the International Society for Pharma-coeconomics and Outcomes Research (ISPOR) Risk–Benefit Management Working Group for their helpful comments.

Source of financial support: The authors did not receive any funding for this study. This study was presented as a research workshop at the ISPOR Annual Meeting, Orlando, FL, USA, May 2009.

References

39. Schechtman E. Odds ratio, relative risk, absolute risk reduction, and the number needed to treat—which of these should we use? Value Health 2002;5:431–6.
postmenopausal women–principal results from the Women's Health Initiative randomized controlled trial. JAMA 2002;288:321–33.


