Benefit/Risk Estimates in Clinical Trials

Adil E. Shamoo and Bilal M. Ayyub

Drug Information Journal 2011 45: 669
DOI: 10.1177/009286151104500617

Benefit/Risk Estimates in Clinical Trials

Adil E. Shamoo
University of Maryland School of Medicine, Baltimore, Maryland

Bilal M. Ayyub
Center for Technology and Systems Management, University of Maryland, College Park, Maryland

Most current benefit/risk analyses of drugs in clinical trials are qualitative and subjective. This study details the development of scientific and objective equations to numerically estimate the benefit/risk (B/R) ratio. The methodology in this study is to consider different severity levels of adverse events and improvement of symptoms by assigning different weights to them.

The set of equations to calculate the B/R ratio take into consideration the clinical data's inherent uncertainties due to sample size, adverse event reporting, and many other factors. Furthermore, the uncertainties of clinical trial data are taken into account by treating the benefit and risk as random variables, utilizing the reliability index, and performing the first-order analysis to evaluate the benefit and risk of the treatment. The methodology discussed can be used on quantitative assessments of not only drugs in clinical trials but also drug treatments of patients. To test the B/R ratio calculation, we applied the method to two large clinical trials obtained from published literature. This experimental approach could assist clinical investigators and regulatory oversight groups in comparing and determining the suitability of various clinical trials and drug treatments by providing a quantitative analysis beyond the existing subjective methods.

Key Words
Benefit/risk; Clinical trials; Decision making; Human subjects; Ethics; Uncertainty

Correspondence Address
Adil E. Shamoo, University of Maryland School of Medicine, 108 N. Greene St., Baltimore, MD 21201 (email: ashamoo@umaryland.edu).

INTRODUCTION

Part of the moral duty of a society is to reduce pain and suffering, and prolong life by the development of new medical products. Conducting clinical trials on human subjects is a crucial final step in this endeavor. However, medical products not only provide great benefits to the public but also can cause harm. The safety of a medical product does not mean that there are no risks associated with the product. Instead, a safe medical product is one having reasonable risks with certain expected benefits. To improve the efficacy of drug or other medical treatments, the balance between the benefits and risks of products and procedures must be determined. All parts of the healthcare delivery chain must strive to maintain the benefit/risk balance in a way that maximizes benefit and minimizes risk (1). The benefit/risk analysis of drugs has been widely performed; however, the assessments are limited to qualitative approaches rather than quantitative practice. Moreover, qualitative methods do not provide a proper treatment of the inherent uncertainties in benefits and risks of the data. In addition, subjective methods do not lend themselves to the evolution of the scientific assessment of benefit/risk. For example, Holden (2) introduced a new and better attempt at assessing and managing drug risks during treatment. In 2000 and 2006, the Institute of Medicine issued two reports (3,4) on medical errors. The 2006 report addressed the issue of inadequate reporting of adverse events during clinical trials and, more specifically, in the postmarketing stage. There is a clear indication that the current reporting requirements and methods of reporting, with one exception for vaccines, may require improvement.

A national survey of outpatient emergency department visits (5) indicates that adverse drug events (ADEs) are sources of significant morbidity, especially for patients aged 65 years or older. The data showed that over 700,000 patients were treated in emergency departments for ADEs annually between 2004 and 2005. These ADEs were for the most common drugs, such as insulin, ibuprofen, and naproxen. An article from industry authors (6) called on industry to take up the serious challenge of ADEs and their underreporting.

RESEARCH WITH HUMAN SUBJECTS

Research with human subjects is vast and encompasses a wide range of research with several different regulatory systems (7,8). There are two large components of research with human subjects. One type of research is regulated by the

Submitted for publication: September 10, 2010
Accepted for publication: January 11, 2011

FDA and the other is regulated by the Office for Human Research Protections (OHRP). Most clinical investigators are familiar with research with human subjects to collect data for submission to the FDA for a license to market a drug in the United States. The FDA-regulated research is conducted with chemicals and drugs that have gone through rigorous testing in animals and some in humans. Moreover, once the drug is approved for use, postmarketing surveillance attempts to gather data on ADEs through the FDA AERs database or MedWatch. The OHRP-regulated research is funded by the federal government and is conducted on chemicals and drugs. The number of human subjects in research (oversight by both OHRP and FDA) exceeds 20 million (7,9,10).

ASSESSMENT OF BENEFIT/RISK

During a clinical trial, a portion of human subject volunteers will experience adverse events. These adverse events will vary from the very serious though rare, such as death, to the benign, such as a headache. The reporting of adverse events is therefore important. Loke and Derry (11) in their literature review of drug treatment confirmed that there are many deficiencies in reporting adverse events. In reviewing 185 trials, they found that 14% made no mention of adverse events and 32% could not be fully evaluated because of missing information. Ioannidis and Lau (8) found that only 39% of 192 randomized controlled clinical trials were adequate in reporting adverse events.

What is also important is how we manage and analyze the reported adverse events. This is important not just ethically, to show respect for subjects, but also for the safety of human subjects during clinical trials. Moreover, it is important for the safety of millions of patients once the drug is approved for marketing. The number of adverse events reported to OHRP between 1990 and 2000 was minimal. However, the FDA data indicate that for all newly approved drugs, about 18,000 adverse events with few deaths are reported annually. The issue of reporting adverse events is so important that the National Bioethics Advisory Commission (12) recommended that "government should create a uniform system for reporting and evaluating adverse events." Our proposal goes to the heart of evaluating the reported adverse events to encourage the development of a more uniform method of reporting to be used as a vital component of the B/R ratio. The National Cancer Institute has developed common terminology criteria for adverse events reporting (13).

A QUANTITATIVE ASSESSMENT OF BENEFIT AND RISK

The seminal work of Chuang-Stein and Mohberg (14) used data from clinical trials to propose categorizing the relationship of side effects and efficacy of treatments of interest, with such distinctions as the following:

Category 1: Efficacy without serious side effects
Category 2: Efficacy with serious side effects
Category 3: No efficacy and no serious side effects
Category 4: No efficacy but serious side effects
Category 5: Unacceptable serious side effects leading to withdrawal

Risks and benefits are evaluated by weighting each category and multiplying by corresponding probabilities. The weight for each parameter can be estimated using formal expert opinion elicitation. The probability of each category can be obtained from clinical data, such as those shown in Table 1. In describing the work of Chuang-Stein (15) and Holden (2), we use their terminology, although we do not think these are the best terms. While describing a clinical trial, they use the word "drug," when they could be referring to a chemical not yet approved or simply a "test article," or they use the term "patient" when they really mean "human subject." Only the term "test article" is defined under sections 351 or 354–360F of the Public Health Service Act.

In addition, a benefit-less-risk measure is performed by Chuang-Stein (15–17) to compare two different drugs. The assessment took place with 10 classes of events and four safety grades from 0 to 3. The classes of events represented

TABLE 1

Adverse Events Occurring in Tolvaptan and Placebo Groups

	No. (%) of Patients									
	Trial A					Trial B				
	Tolvaptan (n = 1,015)		Placebo (n = 1,027)			Tolvaptan (n = 1,048)		Placebo (n = 1,028)		
Adverse Events	No.	%	No.	%	*P* Value	No.	%	No.	%	*P* Value
Dry mouth	43	4.2	7	0.7	<0.001	63	6.0	7	0.7	<0.001
Thirst	79	7.8	5	0.5	<0.001	118	11.3	10	1.0	<0.001
Pollakiuria	13	1.3	3	0.3	0.01	10	1.0	4	0.4	0.18
Polyuria	6	0.6	2	0.2	0.18	35	3.3	5	0.5	<0.001
Hypernatremia	14	1.4	0	0.0	<0.001	5	0.5	0	0.0	0.06
Ventricular extrasystoles	5	0.5	5	0.5	>0.99	11	1.0	2	0.2	0.02
Constipation	35	3.4	20	1.9	0.04	38	3.6	49	4.8	0.23
Atrial fibrillation	3	0.3	5	0.5	0.75	15	1.4	11	1.1	0.56
Ventricular tachycardia	21	2.1	16	1.6	0.41	18	1.7	19	1.9	0.87
Cardiac failure	10	1.0	22	2.1	0.04	17	1.6	16	1.6	0.90
Hypotension	44	4.3	30	2.9	0.10	30	2.9	34	3.3	0.61
Hyponatremia	4	0.4	5	0.5	>0.99	4	0.4	5	0.5	0.75
Hypokalemia	23	2.3	28	2.7	0.57	25	2.4	37	3.6	0.12
Hypomagnesemia	3	0.3	2	0.2	0.69	5	0.5	10	1.0	0.20
Renal failure	21	2.1	20	1.9	0.72	29	2.8	25	2.4	0.63

Source: Gheorghiade et al. (21).

different side effects or adverse reactions to drugs, while the safety grades indicated the relative seriousness. Safety grade 0 represents no safety concern, while grade 3 represents the highest safety concern. Last, the safety score of each class and each individual in the treatment group were summarized to compare the safety of different drugs.

Besides those measures mentioned above, Holden (2) proposed two techniques to quantitatively synthesize a drug's benefits and risks. These two techniques are the relative-value-adjusted number needed to treat (RV-NNT) and its extension, minimum clinical efficacy (MCE) analysis. Evaluations of these techniques rely upon efficacy, or effectiveness data, adverse event data, and utility data from patients describing their preferences for an outcome given potential risks (2). In an RV-NNT analysis, if the number needed to treat is less than the relative-value-adjusted number needed to harm, the patients are willing to take a treatment with a certain benefit knowing that they may be at risk of potential adverse events. In an MCE analysis, the worth of a new treatment in relation to an old treatment is evaluated by giving not only the benefits that might be obtained and risks of potential adverse events, but also by considering the risk of disease with taking no treatment. Nevertheless, the primary shortcomings of these techniques are the lack of definitions of what is being quantified and the lack of treatment of inherent uncertainties. A clear definition of the quantified outcomes is not demon-

strated, such as the number of fatalities, the probability of nonperformance, and the losses represented in dollars. In addition, the inherent uncertainties of clinical data are not considered in these approaches. Uncertainties, such as sample sizes, the estimation of patients' responses to adverse events and improvements from different medical personnel, the environmental effects on patients' responses and on medical personnel's treatment and estimation, and so on, influence the outcomes of the treatment and cause variability that must be accounted for in any analytical quantitative framework.

Current benefit/risk analyses of drugs in clinical trials and treatment are qualitative and subjective. To reduce the subjectivity and lead to objective decisions, an advanced assessment, such as the methodology developed in this article, is needed. Moreover, the methodology introduced here can be utilized through web-based software at a very small cost.

The framework of the methodology developed here provides several significant characteristics, including being analytic, transparent, defensible, quantitative, probabilistic, and consistent. First, it is an analytic methodology that offers a systematic approach for assessing benefits and risks by decomposing them into their respective basic elements. Second, as an analytic methodology, it requires an analyst to lay bare underlying assumptions and data sources at each step of the process. All assumptions and analytical steps in the analysis are clearly defined; therefore, the methodology is transparent. Third, the framework of this methodology is defensible, since values for each parameter are supported by available data, including knowledge from previous studies or credible judgment using expert opinion elicitation. Moreover, the methodology assesses benefits and risks in meaningful and consistent units (eg, weights for each different adverse event similar to dollars and fatalities in other fields) that facilitate comparisons across drugs, usage, and so on, and it provides a basis for performing trade-offs and benefit-cost analysis. It also leverages the mathematics of probability theory for capturing uncertainty in all model parameters and assessing the likelihoods. In addition, the methodology examines this ratio as a time-dependent variable and the final outcome as a time-invariant quantity. The uncertainties in this ratio are also studied by characterizing their probability distribution. Therefore, the methodology developed in this study is quantitative and probabilistic. Last, it is consistent with the established and accepted practices of probabilistic risk assessment used in many other fields; hence it provides a characteristic of consistency.

ETHICAL CONSIDERATIONS

The informed consent documents and procedures communicate the qualitative information available for the drug or study—specific risks and benefits to human subjects. Honigman et al. (18), in assessing electronic medical records in ambulatory settings, illustrated that the current system of reporting ADEs is not adequate when they found that a more thorough computer search for ADEs in medical records dramatically enhanced the detection of the reported adverse events. The authors speculated that many more ADEs go unreported. The crucial importance of thorough and accurate reporting was highlighted by a commentary by Mucklow (19), when he stated, "at the time a new drug is launched, clinical trial data provide the only opportunity for clinicians to assess the frequency and severity of its adverse effects, and the only basis for relating benefit to risk." The approach in this article will contribute to improving and quantifying adverse event reporting (10,20). Such advances, if reliably confirmed widely, could increase safety of drugs and lessen the subjectivity of the informed consent document. These advances in drug development will enhance the ethics as well as the perceptions of clinical trials. Furthermore, the communication of improved and objective information regarding adverse events will enhance participation in clinical trials and increase public confidence in the system. Finally, the quantitative assessment

of benefit/risk could aid physicians and their patients in selecting a more appropriate treatment therapy.

CLINICAL AND HEALTH SIGNIFICANCE

Quantitative approaches of performing benefit/risk analyses of chemicals or drugs are necessary not only to reduce subjectivity but also to provide a clear and transparent definition of potential hazards and benefits. Some techniques proposed by Chuang-Stein and Mohberg (14), Chuang-Stein (15), and Holden (2) are partly quantitative approaches to benefit/risk analysis of chemicals or drugs. However, clinical data have inherent uncertainties, such as data sample size, medical personnel's estimation of patients' responses to adverse events and improvements, the environmental effects on patients' responses, the environmental effects on medical personnel's treatment and estimation, and so on.

The methodology in this article considers different severity levels of adverse events and improvement of symptoms by assigning different weights. The weight assignment can be made by eliciting a formal expert opinion or, ideally, an expert panel. The degree of severity is then consistent between adverse events and improvements. The performance margin of the chemical or drug of interest will be estimated using the B/R ratio concept. Furthermore, the uncertainties of clinical trial data are taken into account by treating the benefits and the risks as random variables, utilizing the reliability index (β), and performing first-order analysis to evaluate the benefits and risks of the treatment.

CALCULATION OF B/R RATIO FOR PUBLISHED CLINICAL TRIAL DATA

Generally, the safety of a chemical or drug is evaluated by comparing the reported adverse events by the human subjects in the control and treatment groups. The efficacy of a study is evaluated by comparing the improvements or relief of symptoms obtained by the subjects in the control and experimental groups. These evaluations are usually conducted by counting the number of subjects who experienced adverse events and sickness relief, and then dividing by the total number of subjects in the control and experimental groups. The safety and efficacy of the chemical or drug are evaluated by comparing the responses of subjects in the control and treatment groups. The method developed here can be used in all research with human subjects, such as clinical trials, as well as in the drug treatment stage.

Furthermore, *P* values are used to evaluate the difference of the performance of the control and the experimental groups. *P* value is the probability of recording a difference between two group means as large as or larger than the difference that was observed. The smaller the *P* value, the stronger the evidence against the null hypothesis. If comparing the treatment and the control group, for instance, the smaller *P* value indicates the larger effects of the treatment. In practice, a *P* value less than 0.001 is interpreted as having strong evidence against the null hypothesis, which means that the treatment has significant effect.

An example of the presentation of clinical trials is shown in Tables 1 and 2. The two examples were chosen to illustrate B/R ratio calculations in clinical trials with numerous adverse events. Two trials, A and B, were conducted by Gheorghiade et al. (21) to evaluate the short-term clinical effects of tolvaptan on patient/subjects (these subjects started as patients) hospitalized for heart failure. There are two groups of subjects receiving tolvaptan and placebo, respectively, in each trial. The safety, or risk, of tolvaptan is shown in Table 1 by comparing the adverse events occurring in the study group (those taking tolvaptan) and the control group (those taking placebo) in each trial. Some adverse events of interest, such as dry mouth, pollakiuria, hypotension, renal failure, cardiac failure, and so on, were observed. The efficacy, or benefit, assessed by comparing the improvements obtained in the control group and the study group is shown in Table 2. Some symptoms of interest, such as dyspnea, orthopnea,

TABLE 2

The Efficacy Obtained in Tolvaptan and Placebo Groups

		Tolvaptan		Placebo		
Condition	**Day**	**No.**	**%**	**No.**	**%**	***P* Value**
Dyspnea	1	933	51.6	853	47.1	0.006
	2	1,244	68.2	1,160	63.7	0.001
	3	1,374	75.2	1,330	73.0	0.02
	4	1,456	79.7	1,431	78.5	0.04
Orthopnea	1	668	63.1	631	59.2	0.01
	2	840	78.5	793	74.1	0.006
	3	893	83.4	863	80.4	0.03
	4	913	85.3	915	85.0	0.06
Fatigue	1	673	40.7	644	38.8	0.19
	2	923	55.3	886	53.1	0.07
	3	1,074	64.3	1,007	60.2	0.02
	4	1,147	68.6	1,107	66.1	0.03
Jugular venous distention	1	698	48.6	631	43.8	0.03
	2	923	63.7	855	59.1	0.01
	3	1,030	71.0	953	65.6	0.002
	4	1,077	74.3	1,014	69.8	0.06
Rales	1	744	45.8	711	43.7	0.03
	2	1,067	65.3	1,041	63.6	0.07
	3	1,201	73.4	1,166	71.1	0.006
	4	1,274	77.9	1,259	76.7	0.02
Edema	1	913	57.6	832	52.6	<0.001
	2	1,229	76.9	1,169	73.5	0.002
	3	1,340	83.8	1,296	81.4	<0.001
	4	1,381	88.4	1,371	86.0	0.004

Source: Gheorghiade et al. (21).

fatigue, and so on, are compared to estimate the efficacy.

Further, some clinical trial data show more detailed responses to side effects, such as those shown in Table 3. These data are from a randomized clinical trial conducted by Treanor et al. (22). The objective of this study was to determine the dose-related safety, immunogenicity, and protective efficacy of an experimental trivalent influenza virus hemagglutinin vaccine. The trial was conducted with three groups of patients, the control group receiving placebo and the treatment groups receiving 75 μg and 135 μg vaccines. The risk is represented as the side effects experienced in each group, and the efficacy is assessed as the antibody response to vaccination in each group. Observations of side effects of this trial not only recorded the num-

TABLE 3

The Side Effects Observed in Placebo and Vaccine Groups

	Placebo (n = 154)						75 µg Vaccine (n = 151)					
	Mild		Moderate		Severe		Mild		Moderate		Severe	
Symptoms	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Arthralgias	7	5	1	1	0	0	7	5	2	1	0	0
Chills	3	2	0	0	0	0	2	1	2	1	0	0
Fatigue	21	14	7	5	0	0	21	14	7	5	0	0
Headache	48	31	15	10	0	0	48	32	4	3	0	0
Myalgias	16	10	3	2	0	0	24	16	2	1	0	0
Nausea	9	6	1	1	0	0	5	3	2	1	0	0
Pain	24	16	1	1	0	0	67	44	0	0	0	0
Sweats	6	4	1	1	0	0	6	4	1	1	0	0
Tenderness	3	2	0	0	0	0	12	8	0	0	0	0

Source: Kirkwood and Sterne (25).

ber and percentage of patients who suffered the side effects but also described the level of response, such as mild, moderate, and severe, as shown in Table 3.

DEVELOPMENT OF PROBABILISTIC METHODOLOGY FOR ASSESSING PERFORMANCE MARGIN FOR B/R ANALYSIS

In general, decision-making situations involve various potential scenarios leading to an adverse event. The risk can be quantified as the probability distribution of the loss and can be illustratively estimated in terms of its expected value as the summation of the product of pairs of probabilities and the corresponding consequences (23–26). When the risk is not acceptable, actions are considered. Based upon the concept of benefit/risk analysis, the costs are these mitigation actions, and the benefits can be represented as the difference between the risk before and after an action, or as a function of the unmitigated risk minus the mitigated risk, as follows:

$$\text{Benefit} = \text{Unmitigated risk} - \text{Mitigated risk}.$$

The benefit/cost ratio can be expressed as:

$$\text{Benefit/cost ratio } (B/C) = \frac{\text{Benefit}}{\text{Cost}} = \frac{\text{Unmitigated risk} - \text{Mitigated risk}}{\text{Cost of mitigation action}}$$

where cost is based on the valuation of the side effects, injuries, and death. Ratios greater than 1 are desirable. Generally, the larger the ratio, the better the mitigation action (26).

Similarly, the performance margin of a chemical or drug can also be described based on the benefit/cost ratio concept. Instead of cost, risk expresses its equivalent, since the cost of having a treatment is the risk of experiencing potential adverse events. The quantitative measure of the performance margin of a treatment can be expressed as a performance margin function:

$$\text{Margin of performance} = \text{Benefit} - \text{Risk}.$$

A direct comparison between benefit and risk is only possible when they are measured on the same scale. Since they are not, valuation analysis must be used to convert them to the same units on the same scales so that simple arithme-

tic makes sense. The performance margin function can also be defined as a ratio:

$$\text{Margin of performance} = \text{Benefit/Risk}$$

or

$$Z = B - R$$
$$Z = B/R$$

where B represents the benefit and R represents the risk. The nonperformance probability can be defined as the probability of $Z = B - R < 0$. The inherent uncertainties of data can be considered by treating the benefit and risk as random variables. If both benefit and risk are treated as random variables, the performance of a drug can be measured using a reliability index (β) as follows:

$$\beta = \frac{\mu_B - \mu_R}{\sqrt{\sigma_B{}^2 + \sigma_R{}^2}}$$

where μ_B = mean value of benefit (B), μ_R = mean value of the risk (R), σ_B = standard deviation of benefit (B), σ_R = standard deviation of the risk (R). The nonperformance probability (P_n) with normal distributed benefit and risk variables has the following relationship with the reliability index (β) :

$$P_n = 1 - \Phi(\beta)$$

where $\Phi(.)$ = cumulative probability distribution function of standard normal distribution. The relationship of the nonperformance probability (P_n) and the reliability index (β) mentioned above assumes that all the random variables in the performance margin function have normal probability distribution and the performance margin function is linear.

Chuang-Stein and Mohberg (14) provide a review of measures for evaluating benefits and risks using categorical data from clinical trials, including the use of a statistic to test the null hypothesis of equal benefit/risk measures (by comparing with a normal distribution to test for significance under the null hypothesis). Two approaches are explored in this study: (a) a time-dependent B/R ratio, and (b) a final-outcome B/R ratio, as described in subsequent sections.

DEVELOPMENT OF TIME-DEPENDENT B/R RATIO

We present here the metrics given in terms of percentiles of time to event distributions, where an event can be interpreted as death or recurrence of disease. The life of a subject with a disease who is not taking a treatment (T_b) for months can be expressed by the following summation:

$$T_b = \sum_{i=1}^{m} iq(i) \qquad (1)$$

where m = number of months of observation, and $q(i)$ = life quality modifier during month i. The life quality modifier should possess the following attributes:

$$0 \le q(i) \le 1$$

where $q(i) = 1$ means life quality in line with age, that is, complete recovery; and $q(i) = 0$ means death.

Similarly, the life of a subject with a disease who is taking a treatment (T_a) for months can be expressed by the following summation:

$$T_a = \sum_{i=1}^{m} iq(i)s(i) \qquad (2)$$

where $s(i)$ = side effect modifier to life quality during month i. The side effect modifier should possess the following attributes:

$$0 \le s(i) \le 1$$

where $s(i) = 1$ means no side effect, and $s(i) = 0$ means death.

The benefit/cost ratio can be defined in this case as follows:

$$\frac{\text{Benefit}}{\text{Risk}} = \frac{T_a - T_b}{T_b} = \frac{T_a}{T_b} - 1. \qquad (3)$$

Most probably, we are going to apply the median as the percentile (ie, $P = 50\%$), but in some studies, lower percentages like 5% and 10% might be appropriate as well. Confidence intervals will be drawn on these percentiles to account for sample sizes.

According to Chuang-Stein et al. (17), it is very desirable that benefit/risk evaluation

should take into account the time dependence because both benefit and risk could be functions of time. It is clear that the suggested metrics are time dependent, and varying the percentage makes them very self-expressing as such.

The approach proposed herein will be compared with the work of Gail and Pfeiffer (27) on absolute risk (π), defined as the probability that an individual who is free of a given disease at an initial age, *a*, will develop that disease in the subsequent interval (*a*, *t*] as provided by:

$$\pi = \int_a^t h_1(u)\exp\left[-\int_a^u \{h_1(v)+h_2(v)dv\}\right]du.$$

The function π is the probability that a subject who is free of the disease of interest at age *a* will be diagnosed with that disease in a subsequent age interval (a, *t*]; $h_1(u)$ is the cause-specific hazard at age *u* for the disease of interest; and $h_2(u)$ is the hazard of mortality from other causes.

FINAL OUTCOME B/R RATIO

The benefit of a treatment can be computed as the inverse of the number needed to treat (NNT) as follows:

$$\text{Benefit} = \frac{1}{\text{NNT}} = p_1 - p_2$$

where p_1 = the proportion, or occurrence probability, of the disease of interest in the control group, and p_2 = the proportion, or occurrence probability, of the disease of interest in the treatment group (2). In other words, the benefit function shown above describes the absolute risk reduction and shows the probability of preventing the disease of interest through treatment. The benefit can also be expressed as the improvements obtained from the treatment, in which p_1 = the proportion, or occurrence probability, of an improvement of interest in the treatment group, and p_2 = the proportion, or occurrence probability, of an improvement of interest in the control group. The latter expression is used in the example section.

The risk of a treatment can be computed as the inverse of the number needed to harm (NNH) as follows:

$$\text{Risk} = \frac{1}{\text{NNH}} = q_1 - q_2$$

where q_1 = the proportion, or occurrence probability, of an adverse event of interest in the control group, and q_2 = the proportion, or occurrence probability, of an adverse event of interest in the treatment group. The risk shows the probability of having additional adverse events of interest due to the treatment.

To take into account the different severities of different adverse events and improvements, different weights (w_i) are assigned to each adverse event in which there is the number of adverse events or side effects and improvement symptoms. In addition, for those data with more descriptions of the improvement and adverse event levels, such as mild, moderate, and severe, as shown in Table 3, different weights (l_j) are also needed, such as 1 for mild, 2 for moderate, and 3 for severe, where *j* is the number of each level. These weights can differ from protocol to protocol. The current choice is arbitrary for the sake of illustration. The weight assigned to each datum can be expressed as $w_i l_j$. For keeping consistency between severities of adverse events and improvements, each weight assigned is divided by the total weight as follows:

$$f_{ij} = \frac{w_i l_j}{\sum_i w_i \cdot \max(l_j)}$$

in which f_{ij} is the adjusted weight for each datum. It should be noted that f_{ij} is a value between 0 and 1, and the summation of f_{ij} for each datum is 1.

$$\Sigma f_{ij} = 1.$$

Therefore, the adjusted benefit can be computed using the benefit function mentioned above, with adjusted p_1 and p_2, which are computed by multiplying the occurrence probabilities of each improvement of interest by the corresponding weights f_{ij}. Similarly, the adjusted risk can be computed by multiplying the occurrence probabilities of each adverse event or side effect of interest by the corresponding weights f_{ij} as adjusted q_1 and q_2.

TABLE 4

Weighting of Improvement of Symptoms and Adverse Events

i		Severity [1, 4] w_i	f_i
	Improvement		
1	Dyspnea	4	0.077
2	Orthopnea	3	0.058
3	Fatigue	1	0.019
4	Jugular venous distention	3	0.058
5	Rales	2	0.038
6	Edema	2	0.038
	Subtotal	15	0.288
	Adverse Events		
7	Dry mouth	1	0.019
8	Thirst	1	0.019
9	Pollakiuria	1	0.019
10	Polyuria	1	0.019
11	Hypernatremia	3	0.058
12	Ventricular extrasystoles	3	0.058
13	Constipation	1	0.019
14	Atrial fibrillation	3	0.058
15	Ventricular tachycardia	3	0.058
16	Cardiac failure	4	0.077
17	Hypotension	3	0.058
18	Hyponatremia	3	0.058
19	Hypokalemia	3	0.058
20	Hypomagnesemia	3	0.058
21	Renal failure	4	0.077
	Subtotal	37	0.712
Total		52	1.000

The methodology described above is to evaluate the performance of the tested chemicals or drugs in a clinical trial based upon the B/R ratio concept. Overall, the risk will be quantified in terms of the probability distribution of adverse events valued in monetary terms and can be illustratively estimated as the summation of the product of the probabilities of experiencing potential adverse events of interest obtained from the collected data and the corresponding severities f_{ij}. The difference of the risk in the treatment group and the control group describes the risk due to the treatment. On the other hand, the benefit is expressed as the difference of the risk before and after the treatment, or as the improvement due to the treatment. Similarly, the benefit can be quantified in terms of the probability distribution of reduction in adverse events valued in monetary terms and can be illustratively quantified as the summation of the product of the probabilities of having improvements of interest obtained from the collected data and the corresponding severities f_{ij}. Furthermore, the higher the B/R ratio, the better the performance of the tested chemicals or drugs in a clinical trial. When the performance margin function is expressed as benefit less risk, the higher the outcome and the better the treatment. In addition, the nonperformance probability of the risk outweighing the benefit is evaluated using a reliability index (β) (28,29).

EXAMPLE USING THE NEW APPROACH

This example demonstrates only the computation of the B/R ratio as a final outcome.

The clinical trial data collected by Gheorghiade et al. (21) are used in this study as an example to illustrate the methodology mentioned above. The safety/risk data are shown in Table 1 by comparing the adverse events occurring in the treatment group (taking tolvaptan) and the control group (taking placebo) in each trial. Adverse events of interest, such as dry mouth, pollakiuria, hypotension, renal failure, cardiac failure, and so on, were observed to evaluate the safety of the treatment. The efficacy or benefit is assessed by comparing the improvements obtained in the control group with those obtained in the treatment group, as shown in Table 2. Symptoms, such as dyspnea, orthopnea, fatigue, and so on, are compared to estimate the efficacy, and the data of the fourth day of improvements are used in this example. As mentioned

TABLE 5

Risk Estimation of Tolvaptan

	Tolvaptan					Placebo				
Adverse Events	f_i	n	No.	q_T	$R_T = f_i q_T$	n	No.	q_C	$R_C = f_i q_C$	$R_T - R_C$
Dry mouth	0.019	2063	106	0.0514	0.0010	2055	14	0.0068	0.0001	0.0009
Thirst	0.019	2063	197	0.0955	0.0018	2055	15	0.0073	0.0001	0.0017
Pollakiuria	0.019	2063	23	0.0111	0.0002	2055	7	0.0034	0.0001	0.0001
Polyuria	0.019	2063	41	0.0199	0.0004	2055	7	0.0034	0.0001	0.0003
Hypernatremia	0.058	2063	19	0.0092	0.0005	2055	0	0.0000	0.0000	0.0005
Ventricular extrasystoles	0.058	2063	16	0.0078	0.0004	2055	7	0.0034	0.0002	0.0003
Constipation	0.019	2063	73	0.0354	0.0007	2055	69	0.0336	0.0006	0.0000
Atrial fibrillation	0.058	2063	18	0.0087	0.0005	2055	16	0.0078	0.0004	0.0001
Ventricular tachycardia	0.058	2063	39	0.0189	0.0011	2055	35	0.0170	0.0010	0.0001
Cardiac failure	0.077	2063	27	0.0131	0.0010	2055	38	0.0185	0.0014	–0.0004
Hypotension	0.058	2063	74	0.0359	0.0021	2055	64	0.0311	0.0018	0.0003
Hyponatremia	0.058	2063	8	0.0039	0.0002	2055	10	0.0049	0.0003	–0.0001
Hypokalemia	0.058	2063	48	0.0233	0.0013	2055	65	0.0316	0.0018	–0.0005
Hypomagnesemia	0.058	2063	8	0.0039	0.0002	2055	12	0.0058	0.0003	–0.0001
Renal failure	0.077	2063	50	0.0242	0.0019	2055	45	0.0219	0.0017	0.0002
Total	0.712				0.0134				0.0100	0.0034

Mean, 0.00023; standard deviation, 0.00053.

above, weights are needed to show the severities of different adverse events and improvements. A measure from 1 to 4 is used on each adverse event and improvement, as shown in Table 4. Since the levels of adverse events and improvements are not available from the clinical data shown in Tables 1 and 2, l_j is not needed in this case or may be assigned as a constant value, say 1. The weights of each symptom shown in Table 4 demonstrate the consistency of the severities between adverse events and improvements. As we mentioned earlier, the weights used here are arbitrary. However, in an actual protocol the investigators involved should choose these numbers prior to the commencement of the study. For example, assuming that patients are willing to have dry mouth to obtain the improvement in fatigue, the weights of fatigue (w_3) and dry mouth (w_7) are assigned the same value as 1 to represent that the severities of fatigue and dry mouth are considered as the same level. By dividing the weight, 1, by the total weight of adverse events and improvements, 52, both fatigue and dry mouth symptoms have weights as follows:

$$f_3 = f_7 = 1/52 = 0.019.$$

In addition, the summation of each weight is computed as

$$\sum_{i=1}^{21} f_i = 1.$$

Risk is estimated in Table 5, in which q_T is denoted as the occurrence probabilities of adverse events of interest in the treatment group; q_C is denoted as the occurrence probabilities of adverse events of interest in the control group; R_T and R_C are the risk experienced in the treatment and control groups, respectively; and

TABLE 6

Benefit Estimation of Tolvaptan

	Tolvaptan					Placebo				
Improvements	f_i	n	No.	p_T	$B_T = f_i p_T$	n	No.	p_C	$B_C = f_i p_C$	$B_T - B_C$
Dyspnea	0.077	1826	1456	0.7974	0.0613	1822	1431	0.7854	0.0604	0.0009
Orthopnea	0.058	1070	913	0.8533	0.0492	1075	915	0.8512	0.0491	0.0001
Fatigue	0.019	1670	1147	0.6868	0.0132	1673	1107	0.6617	0.0127	0.0005
Jugular venous distention	0.058	1450	1077	0.7428	0.0429	1452	1014	0.6983	0.0403	0.0026
Rales	0.038	1635	1274	0.7792	0.0300	1640	1259	0.7677	0.0295	0.0004
Edema	0.038	1598	1381	0.8642	0.0332	1594	1371	0.8601	0.0331	0.0002
Total	0.288				0.2298				0.2251	0.0047

Mean, 0.00078; standard deviation, 0.00092.

$R_T - R_C$ is the risk difference between the treatment and control groups. In other words, if 17 per 10,000 patients experience renal failure, for example, in the control group, there might be 19 per 10,000 patients experiencing renal failure in the treatment group. The mean value of $R_T - R_C$ indicates the risk of taking the treatment for each adverse event on average. Similarly, the benefit can be computed and is shown in Table 6, where p_T is the occurrence probability of improvements of interest in the treatment group; p_C is the occurrence probability of improvements of interest in the control group; B_T and B_C are the improvements experienced in the treatment and control groups, respectively; and $B_T - B_C$ is the benefit difference between the treatment and control groups of improvement symptoms of interest.

The performance of tolvaptan can be quantitatively described based on the benefit/cost ratio concept utilizing the estimates in Tables 5 and 6. Instead of cost, risk is used to express the treatment performance, since the risk of experiencing potential adverse events is the cost of having a treatment. The quantitative assessment of performance can be expressed as a performance-margin function as follows:

$$\begin{aligned}\text{Margin of performance } (Z) &= \text{Benefit } (B) - \text{Risk } (R)\\ Z &= B - R\\ &= \mu_B - \mu_R\\ &= 0.00078 - 0.00023\\ &= 0.00055\end{aligned}$$

where μ_B is the average of the benefit (B) and μ_R is the average of the risk (R). The margin of performance (Z) as computed is larger than 0, which indicates that the benefit of taking tolvaptan is larger than the possible risk experienced. The estimation also implies that if there are 10,000 patients in the trial, there would be approximately 8 people obtaining benefit and 2 people experiencing risk resulting from the treatment.

The margin of performance function can also be presented as the following form:

$$\begin{aligned}\text{Margin of performance } (Z) &= \text{Benefit } (B)/\text{Risk } (R)\\ Z &= B/R\\ &= \mu_B/\mu_R\\ &= 0.00078 \,/\, 0.00023\\ &= 3.4.\end{aligned}$$

The margin of performance (Z) as computed is larger than 1, which is desirable, indicating that the benefit of taking tolvaptan is larger than the possible risk experienced.

The evaluation of benefit and risk, however, does not consider the statistical uncertainties associated with the clinical data. The clinical data uncertainties are those that can lead to various results in the evaluation of the same drug. Possible uncertainties are the sample

size, different benefits and risks of different symptoms, estimations from medical personnel, the environmental effects, and so on. Moreover, the human subjects at a given trial may not have received education and training on what to report as adverse events. A larger sample size can lead to relatively more reliable outcomes of the benefit and risk estimation, although a large sample size may be somehow unavailable. Moreover, the estimation of the improvements and side effects may be different from different medical personnel. For instance, one medical staff member may rate a symptom as an improvement, while the other may not. Furthermore, the hospital environment may affect the behavior of medical personnel and patients. The inherent uncertainties of clinical data influence the objectivity of evaluation outcomes.

To take into account the data uncertainties, benefit and risk are treated as random variables. Since both benefit and risk are treated as random variables, the performance of a drug can be measured using a reliability index (β) as follows:

$$\beta = \frac{\mu_B - \mu_R}{\sqrt{\sigma_B^2 + \sigma_R^2}}$$

where μ_B = mean value of benefit B, μ_R = mean value of the risk R, σ_B = standard deviation of benefit B, and σ_R = standard deviation of the risk R. The nonperformance probability (P_n) with normally distributed benefit and risk variables has the following relationship with the reliability index (β):

$$P_n = 1 - \Phi(\beta)$$

where $\Phi(.)$ = cumulative probability distribution function of standard normal distribution. The relationship between the nonperformance probability and the reliability index mentioned above assumes that all the random variables in the performance margin function have normal probability distributions and the performance margin function is linear. According to the estimation of the benefits and risks of tolvaptan, as shown in Tables 5 and 6, μ_B, the mean value of the benefit (B), is computed as 0.00078; μ_R, the mean value of the risk (R), is 0.00023; σ_B, the standard deviation of the benefit (B), is 0.00092; and σ_R, the standard deviation of the risk (R), is 0.00053. The reliability index (β) can be computed as:

$$\beta = \frac{\mu_B - \mu_R}{\sqrt{\sigma_B^2 + \sigma_R^2}} = \frac{0.00078 - 0.00023}{\sqrt{0.00092^2 + 0.00053^2}} = 0.5256.$$

And the nonperformance probability (P_n) can be computed as:

$$P_n = 1 - \Phi(\beta) = 0.2996.$$

The outcome obtained above indicates that there is a nonperformance probability of 0.2996 for taking tolvaptan. In other words, there is an almost 30% chance of nonperformance when taking tolvaptan. This evaluation not only shows that the benefit is higher than the risk of taking tolvaptan, but also indicates that the probability of nonperformance takes into account the uncertainties of clinical trial data.

Including death as an adverse event, the assessment is redone as follows. The weight of death in this case is assigned as 6 (of course, it can be a much higher number), and the revised weighting of improvement symptoms and adverse events is shown in Table 7. The risk is reestimated to include death as an adverse event, as shown in Table 8. When performing the project tasks, the concept of valuation based on willingness to pay (28) will be used instead of weight factors. We will evaluate the sensitivity of the conclusion to the choice of the weights since the choice of the weights has a subjective component.

The performance of tolvaptan including death as an adverse event is evaluated using a B/R ratio as follows:

$$\begin{aligned} \text{Margin of performance } (Z) &= \text{Benefit } (B)/\text{Risk } (R) \\ Z &= B/R \\ &= \mu_B/\mu_R \\ &= 0.00078/0.00019 \\ &= 4.1 \end{aligned}$$

TABLE 7

Weighting of Improvement of Symptoms and Adverse Events Including Death as an Adverse Event

i		Severity [1, 6] w_i	f_i
	Improvement		
1	Dyspnea	4	0.069
2	Orthopnea	3	0.052
3	Fatigue	1	0.017
4	Jugular venous distention	3	0.052
5	Rales	2	0.034
6	Edema	2	0.034
	Subtotal	15	0.259
	Adverse Events		
7	Dry mouth	1	0.017
8	Thirst	1	0.017
9	Pollakiuria	1	0.017
10	Polyuria	1	0.017
11	Hypernatremia	3	0.052
12	Ventricular extrasystoles	3	0.052
13	Constipation	1	0.017
14	Atrial fibrillation	3	0.052
15	Ventricular tachycardia	3	0.052
16	Cardiac failure	4	0.069
17	Hypotension	3	0.052
18	Hyponatremia	3	0.052
19	Hypokalemia	3	0.052
20	Hypomagnesemia	3	0.052
21	Renal failure	4	0.069
22	Death	6	0.103
	Subtotal	43	0.741
Total		58	1.000

where μ_B is the average of the benefit (B), and μ_R is the average of the risk (R). The margin of performance (Z) as computed is larger than 1, which indicates that the benefit of taking tolvaptan is larger than the possible risk.

The margin of performance of the treatment can also be evaluated as follows:

$$\begin{aligned}\text{Margin of performance } (Z) &= \text{Benefit } (B) - \text{Risk } (R)\\ Z &= B - R\\ &= \mu_B - \mu_R\\ &= 0.00078 - 0.00019\\ &= 0.00059.\end{aligned}$$

The margin of performance (Z) as computed is larger than 0, which indicates that the benefit of taking tolvaptan outweighs the possible risk experienced.

The nonperformance probability can be estimated using a reliability index calculated as follows:

$$\beta = \frac{\mu_B - \mu_R}{\sqrt{\sigma_B^{\,2} + \sigma_R^{\,2}}} = \frac{0.00078 - 0.00019}{\sqrt{0.00092^2 + 0.00046^2}} = 0.5791.$$

The nonperformance probability (P_n) then is computed as:

$$P_n = 1 - \Phi(\beta) = 0.2812.$$

The results obtained previously excluding death as an adverse event indicate that the nonperformance probability of tolvaptan is 0.2996, and the B/R ratio is 3.466. Comparing the results of the assessments excluding and including death as an adverse event, the assessment outcome including death as an adverse event gives a lower nonperformance probability, 0.2812, and a higher B/R ratio, 4.166. This is because the numbers of deaths in treatment and control groups are both 17, which gives a credit to the treatment.

In this illustrative example, the risk of the treatment is evaluated as the summation of the product of the probabilities of the adverse events of interest obtained from the collected data and the corresponding severities. The benefit of the treatment is analyzed as the summation of the product of the probabilities of the improvements of interest obtained from the collected data and the corresponding severities. The severity of each adverse event and improvement (w_i) is measured as a score from 1 to 6, in which 1 represents the lightest severity and 6 represents the worst severity. To be con-

TABLE 8

Risk Estimation of Tolvaptan Including Death as an Adverse Event

	Tolvaptan					Placebo				
Adverse Events	f_i	n	No.	q_T	$R_T = f_i q_T$	n	No.	q_C	$R_C = f_i q_C$	$R_T - R_C$
Dry mouth	0.017	2072	106	0.0512	0.0009	2061	14	0.0068	0.0001	0.0008
Thirst	0.017	2072	197	0.0951	0.0016	2061	15	0.0073	0.0001	0.0015
Pollakiuria	0.017	2072	23	0.0111	0.0002	2061	7	0.0034	0.0001	0.0001
Polyuria	0.017	2072	41	0.0198	0.0003	2061	7	0.0034	0.0001	0.0003
Hypernatremia	0.052	2072	19	0.0092	0.0005	2061	0	0.0000	0.0000	0.0005
Ventricular extrasystoles	0.052	2072	16	0.0077	0.0004	2061	7	0.0034	0.0002	0.0002
Constipation	0.017	2072	73	0.0352	0.0006	2061	69	0.0335	0.0006	0.0000
Atrial fibrillation	0.052	2072	18	0.0087	0.0004	2061	16	0.0078	0.0004	0.0000
Ventricular tachycardia	0.052	2072	39	0.0188	0.0010	2061	35	0.0170	0.0009	0.0001
Cardiac failure	0.069	2072	27	0.0130	0.0009	2061	38	0.0184	0.0013	−0.0004
Hypotension	0.052	2072	74	0.0357	0.0018	2061	64	0.0311	0.0016	0.0002
Hyponatremia	0.052	2072	8	0.0039	0.0002	2061	10	0.0049	0.0003	−0.0001
Hypokalemia	0.052	2072	48	0.0232	0.0012	2061	65	0.0315	0.0016	−0.0004
Hypomagnesemia	0.052	2072	8	0.0039	0.0002	2061	12	0.0058	0.0003	−0.0001
Renal failure	0.069	2072	50	0.0241	0.0017	2061	45	0.0218	0.0015	0.0002
Death	0.103	2072	17	0.0082	0.0008	2061	17	0.0082	0.0009	0.0000
Total	0.741			0.3605	0.0120			0.1960	0.0090	0.0030

Mean, 0.00019; standard deviation, 0.00046.

sistent, the same grade system is utilized to describe both adverse events and improvements. Besides, some data also provide different levels of each adverse event and improvement, such as mild, moderate, and severe. For those data that have additional descriptions, another score (l_j) is used to capture the information. To normalize the weights and make the summation of all weights assigned to each adverse event and improvement equal 1, the weight of each adverse event and improvement is computed as the score (w_i) divided by the summation of all scores (w_i) and multiplied by another score (l_j), which was captured using additional information from the data, then divided by the maximum l_j. Therefore, the risk and benefit of each adverse event and improvement is evaluated as the product of corresponding probability and weight. The average risk due to the treatment is estimated as the average of the difference of the risk in the treatment and control groups for each adverse event. Similarly, the average benefit of the treatment is estimated as the average of the difference of the benefit in the treatment and control groups for each improvement. The margin of performance is evaluated as the benefit less risk, and the outcome, which is larger than 1, represents that the benefit outweighs the risk. The probability of nonperformance is estimated using a reliability index (β), which is computed from the average and standard deviation of the risk and the benefit. The standard deviation of the risk due to the treatment is computed as the standard deviation of the difference of the risk in the treatment and control groups for each adverse event. Similarly, the standard deviation of the benefit is computed the same way. The

nonperformance probability is estimated as 1 less the cumulative probability distribution function of the standard normal distribution Φ(.) of the reliability index (β).

CONCLUSION

The approach discussed in this article could help to provide a broader foundation in the development of benefit/risk estimation and its use during the conduct of clinical trials. We are cognizant of some of the pitfalls of this relatively new area. The B/R ratios may be so varied that they may not show any discernable pattern at first glance. But the two clinical trial examples we used indicate otherwise. Careful examination of patterns consistent with current practices should be contemplated in the future. Current practices used to discern the ethics of these decisions is, in most cases, rational and reasonable, but not quantitative. This approach could assist in determining the suitability of various research designs with human subjects in clinical trials by providing a quantitative analysis beyond the existing parameters. It is unclear to us, at this time, how much resistance this approach will face from management even though the cost increase of drug development would be very small.

The approach we illustrated is practicable and beneficial. In the long run, hopefully, the outcome of this approach would be used by researchers, institutional review boards, oversight agencies, and the community at large. We can then delineate in detail how review boards and clinical investigators would utilize the knowledge and calculate the parameters we just described.

REFERENCES

1. Food and Drug Administration. *Managing the Risks from Medical Product Use: Creating a Risk Management Framework*. Washington, DC: US Department of Health and Human Services, Food and Drug Administration; 1999. http://www.fda.gov/oc/tfrm/1999report.html.
2. Holden WL. Benefit-risk analysis: a brief review and proposed quantitative approaches. *Drug Saf*. 2003;26:853–862.
3. Institute of Medicine. *To Err Is Human: Building a Safer Health System*. Washington, DC: National Academy of Science; 2000.
4. Institute of Medicine. *Preventing Medication Errors*. Washington, DC: National Academy of Science; 2006.
5. Budnitz DS, Pollock DA, Weidenback KN, et al. National surveillance of emergency department visits for outpatient adverse drug events. *JAMA*. 2006;296:1858–1866.
6. Meinert L, Ward H. Press one for safety. *Appl Clin Trials*. 2007;(September):46–50.
7. Shamoo AE, Schwartz J. Universal and uniform protections of human subjects in research. *Am J Bioethics*. 2007;7(12):7–9.
8. Ioannidis JPA, Lau J. Completeness of safety reporting in randomized trials: an evaluation of 7 medical areas. *JAMA*. 2001;285:437–443.
9. Shamoo AE. Adverse events reporting: the tip of an iceberg. *Accountability Res*. 2001;8:197–218.
10. Shamoo AE. The need for better reporting of adverse events for human subjects in research. *DIA Today*. 2005;5(2):18, 20.
11. Loke YK, Derry S. Reporting of adverse drug reactions in randomized controlled trials: a systematic survey. *BMC Clin Pharmacol*. 2001. http://www.biomedcentral.com/1472-6904/1/3 (accessed August 29, 2007).
12. National Bioethics Advisory Commission. *Ethical and Policy Issues in Research Involving Human Participants*. Bethesda, MD: NBAC; 2001.
13. National Cancer Institute. *Common Terminology Criteria for Adverse Events (CTCAE)*, Version 3.0. 2003. http://ctep.cancer.gov/.
14. Chuang-Stein C, Mohberg NR. Three measures for simultaneously evaluating benefits and risks using categorical data from clinical trials. *Stat Med*. 1991;10:1349–1359.
15. Chuang-Stein C. The role and analysis of benefits and risks in the development of new drugs. In: Korting HC, Schafer-Korting M, eds. *The Benefit/Risk Ratio: A Handbook for the Rational Use of Potentially Hazardous Drugs*. Boca Raton, FL: CRC Press; 1999.
16. Chuang-Stein C. A new proposal for benefit-less-risk analysis in clinical trials. *Control Clin Trials*. 1994;15:30–43.

17. Chuang-Stein C, Entsuuh R, Pritchett Y. Measures for conducting comparative benefit:risk assessment. *Drug Inf J*. 2008;42:223–233.
18. Honigman B, Lee J, Rothschild J, et al. Using computerized data to identify adverse drug events in outpatients. *J Am Med Inform Assoc*. 2001;8:254–266.
19. Mucklow JC. Reporting drug safety in clinical trials: getting the right emphasis right. *Lancet*. 2001;357:1384.
20. Steinbrook R. Improving protection for research subjects. *New Engl J Med*. 2002;346:1425–1430.
21. Gheorghiade M, Konstam MA, Burnett JC Jr, et al. Short-term clinical effects of Tolvaptan, an oral vasopressin antagonist, in patients hospitalized for heart failure: the EVEREST clinical status trials. *JAMA*. 2007;297:1332–1343.
22. Treanor JJ, Schiff GM, Hayden FG, et al. Safety and immunogenicity of a baculovirus-expressed hemagglutinin influenza vaccine. *JAMA*. 2007; 297:1577–1582.
23. Ang AH-S, Tang W. *Probability Concepts in Engineering Planning and Design*, Vol. 2. New York: Wiley; 1984.
24. Ayyub BM. *Risk Analysis in Engineering and Economics*. Boca Raton, FL: Chapman & Hall/CRC Press; 2003.
25. Kirkwood BR, Sterne JAC. *Medical Statistics*. Boston: Blackwell Science; 2003.
26. Korting HC, Schäfer-Korting M, eds. *The Benefit/Risk Ratio: A Handbook for the Rational Use of Potentially Hazardous Drugs*. Boca Raton, FL: CRC Press; 1999.
27. Gail M, Pfeiffer RM. On criteria for evaluating models of absolute risk. *Biostatistics*. 2005;6(2): 227–239.
28. Ayyub BM. *Elicitation of Expert Opinions for Uncertainty and Risks*. Boca Raton, FL: CRC Press; 2001.
29. Ayyub BM, McCuen RH. *Probability, Statistics and Reliability for Engineers and Scientists*. Boca Raton, FL: CRC Press; 2003.

Adil E. Shamoo has disclosed that he is a stock shareholder of Human Research Technology. Bilal Ayyub reports no relevant relationships to disclose.