International Journal of Technology Assessment in Health Care, **17:1** (2001), 17–26.

USING FULL PROBABILITY MODELS TO COMPUTE PROBABILITIES OF ACTUAL INTEREST TO DECISION MAKERS

Frank E. Harrell, Jr.
University of Virginia School of Medicine

Ya-Chen Tina Shih
MEDTAP International Inc.

Abstract

The objective of this paper is to illustrate the advantages of the Bayesian approach in quantifying, presenting, and reporting scientific evidence and in assisting decision making. Three basic components in the Bayesian framework are the prior distribution, likelihood function, and posterior distribution. The prior distribution describes analysts' belief *a priori*; the likelihood function captures how data modify the prior knowledge; and the posterior distribution synthesizes both prior and likelihood information. The Bayesian approach treats the parameters of interest as random variables, uses the entire posterior distribution to quantify the evidence, and reports evidence in a "probabilistic" manner. Two clinical examples are used to demonstrate the value of the Bayesian approach to decision makers. Using either an uninformative or a skeptical prior distribution, these examples show that the Bayesian methods allow calculations of probabilities that are usually of more interest to decision makers, e.g., the probability that treatment A is similar to treatment B, the probability that treatment A is at least 5% better than treatment B, and the probability that treatment A is not within the "similarity region" of treatment B, etc. In addition, the Bayesian approach can deal with multiple endpoints more easily than the classic approach. For example, if decision makers wish to examine mortality and cost jointly, the Bayesian method can report the probability that a treatment achieves at least 2% mortality reduction and less than $20,000 increase in costs. In conclusion, probabilities computed from the Bayesian approach provide more relevant information to decision makers and are easier to interpret.

Keywords: Bayesian statistics, Decision analysis, Multiple endpoints

Researchers in many fields present scientific evidence in a variety of ways to assist decision makers with their choices. In the healthcare marketplace, statistical evidence is commonly used by researchers to inform decision makers of the costs and consequences of alternatives available to them. Conventionally, the classical approach has been the standard way to report medical evidence. Under this approach, parameters of interest are treated as unknown fixed numbers, and estimates of these parameters are obtained from the observed data. Concepts such as hypothesis testing, power calculation, and p values are commonly used

The authors gratefully acknowledge the work of two referees whose critiques and suggestions resulted in significant improvements to this paper.

by classical statisticians (e.g., frequentists). Bayesian statisticians, on the other hand, think of the parameters as random variables, form a prior belief regarding these parameters, and use the observed data to update their belief in the posterior distribution. The Bayesian posterior distribution offers a lot more flexibility in the type of evidence one can report and produces results more transparent to interpret. Bayesian applications are especially useful in studies involving multiple endpoints, which are common in clinical studies. The purpose of this paper is to illustrate the advantages of the Bayesian approach in quantifying, presenting, and reporting scientific evidence and in assisting decision making.

QUANTIFICATION OF EVIDENCE AND DECISION MAKING

Scientific evidence can be quantified statistically in many ways. Healthcare researchers usually present quantitative evidence using one of the following six methods.

1. *Standard point estimates*: Estimates of evidence are considered to be random variables and summary statistics such as the mean are used to quantify the evidence. For example, researchers may report that a new pharmacologic agent has been shown to reduce blood pressure by 5 mm, or that the difference in cholesterol level between a new treatment and a conventional treatment is 30 mg/dl. To verify whether findings from these point estimates are statistically meaningful, researchers form and test hypotheses regarding the quantified evidence. The p value, defined as the probability of obtaining a result as extreme as or more extreme than the actual sample value obtained given that the null hypothesis is true, is then used to guide the decisions regarding the statistical significance of the quantitative evidence. Binary decisions (e.g., reject or do not reject the null hypothesis, usually a zero population difference) are made based on the p value. See Goodman (8) and Davidoff (4) for discussions of problems with p values.
2. *Interval estimates (confidence interval)*: Interval estimates quantify the evidence (usually summarized by a set of parameters) in a range of values that, if hypothesized to hold, would not be rejected at the selected level of significance, α. Alternatively, the confidence interval can be interpreted as the interval that gives the desired coverage probability for a parameter estimate. However, the concept of parameter estimate, instead of parameter, associated with confidence intervals often makes it difficult for beginners in statistics (and often even trained statisticians) to correctly interpret the meaning of a confidence interval. Confidence intervals also convey the false impression that all parameter values within the interval are equally supported by the data.
3. *Probability for a parameter conditional on current data*: Unlike the frequentists who tend to present evidence in a more deterministic manner (for example, point estimates or p values), the Bayesian statisticians quantify evidence in a more probabilistic manner. For example, instead of claiming that the estimated treatment effect is 7 mm Hg blood pressure reduction, researchers using the Bayesian approach often describe the evidence as "conditional on the currently available data, the probability that the blood pressure reduction exceeds 5 mm Hg is 0.8."
4. *The entire probability distribution for the parameter*: By treating the parameters summarizing the evidence as random variables (rather than fixed values), the Bayesian approach can use the entire probability distribution to describe the evidence of interest. For example, if blood pressure reduction is the targeted evidence, a distribution describing the different levels of population mean blood pressure reduction given the current data can be constructed using the Bayesian method. Once such a distribution is constructed, a variety of summary statistics and probabilistic descriptions can be used to quantify evidence (discussed below).
5. *Decision analysis to make an optimal binary decision*: The Bayesian reasoning is also used to assist optimal decision making. In trials as well as clinical practice settings, it is common for decision makers to adopt the strategy of "withhold/delay decision" or "wait for further data" in addition to a yes/no decision. This type of decision usually involves designs of decision models, specifications of prior beliefs, and selections of appropriate loss functions and data. An optimal decision will then be identified by minimizing expected losses or maximizing expected utilities while taking

into account the prior belief and decision structure. This method has been used to evaluate whether a clinical trial should be discontinued given the information collected (3;14).

6. *Relative evidence*: Evidence can also be quantified in a relative manner such as odds ratios, likelihood ratios, Bayes' factors, etc. Classical statisticians often estimate the efficacy of one treatment relative to another treatment. For example, a study comparing two treatments might estimate that patients under the new treatment were three times more likely to recover from a condition than those under the current treatment. Bayesian thinking can be used in that situation, but it can also provide a different kind of relative quantification of evidence: the researcher might conclude, after analyzing the data, that it is now 18 times more likely that the treatment is effective than was believed before the data became available. This relative quantification does not rely on a prior distribution when the null and alternative hypotheses each correspond to single parameter values (8;9).

Note that healthcare researchers should choose among the variety of statistical tools and methods with care. Depending on the cases to be examined, evidence produced by some methods may be irrelevant to actual decision making. For example, evidence regarding medical diagnosis reported under the classical approach is not sufficient for making clinical decisions.

When the frequentist approach to statistical inference is used in the medical diagnostic framework, frequentists use the probability of a test outcome given the disease state to quantify the evidence. Information such as sensitivity and specificity typifies this approach. Alternatively, the Bayesian approach focuses on the post-test probability of disease, which will help physicians decide whether a patient should be treated or selected for the next test. Taking a cholesterol test, for example, the frequentists will report the cholesterol distribution given that one has coronary disease. However, the reversed information (e.g., probability of coronary disease given the cholesterol level) is actually more relevant to clinicians.

In addition to the classical vs. Bayesian debate, another debate regarding the quantification of evidence is whether one should use probabilistic models versus classification models to report evidence (17). An example of the probabilistic model is logistic models whose outputs are predicted probabilities of a disease. On the other hand, discriminant analysis is an example of a classification model, where the final output is a classification such as dead versus alive or with versus without disease. In general, the former model provides richer and more flexible information to decision makers. Interestingly, some researchers use probabilistic models to obtain the value of the predicted probability, but then report the results in classifications; that is, if the predicted probability is greater than 0.5, the patient is classified as having the disease. When predicted probabilities are categorized by the analyst and not by the physician and patient at the time decisions are actually made, a significant amount of information is wasted and arbitrary utilities or loss functions are being imposed by the analyst or writer on the physician or patient.

Once the quantified evidence is obtained, researchers usually design a formal decision model to assist decision making (6;11;19). In formal decision models, a decision tree is structured to describe the alternatives and outcomes associated with each alternative. Probabilities are used to quantify the uncertainties associated with each outcome measure, which can be presented in a variety of ways, including utility scales and physical or monetary units. After the decision tree is completed, researchers can evaluate impacts (e.g., costs, life-years saved, quality of life, quality-adjusted life-years) of different strategies and derive the corresponding thresholds. The Bayesian technique provides an excellent tool in formal decision making. This point was well illustrated by Berry et al. (1). In their paper, the authors used a vaccine trial example to demonstrate how the prior information can be combined with an ongoing clinical trial to assist decisions such as when to stop the trial or whether to seek regulatory approval.

APPLICATION OF BAYESIAN ANALYSIS

Basics of Bayesian Analysis

Three basic elements in the Bayesian approach are prior, likelihood, and posterior. The prior distribution describes analysts' belief *a priori* on the parameters of interest, the likelihood function captures how data modify the prior knowledge of the parameters, and the posterior distribution synthesizes both prior and likelihood information in a probabilistic manner. Two distinct features of the Bayesian approach are the use of the prior distribution and the way θ, the parameter (or parameters) of interest, is conceptualized. The likelihood function serves as an intersection between frequentists and Bayesian statisticians; it is a common element in both approaches. Frequentists treat θ as a fixed unknown value and use the observed data to estimate θ. However, Bayesian statisticians treat θ as an unknown random variable. They have a prior belief regarding θ, update their prior belief with the data observed, and summarize this information in the posterior distribution. Frequentists view data as random variables while Bayesians view the unknown parameters as random because those quantities are unknowable and are the quantities about which belief statements are needed.

To operationalize Bayesian methods, suppose that we want to compare treatment 1 with treatment 2 where the population mean responses are μ_1 and μ_2, respectively. The parameter of interest is: $\theta = \mu_1 - \mu_2$. The information that researchers want to know is $\Pr(\theta > 0|\text{data})$ or $\Pr(\theta > \text{c}|\text{data})$, for example, where c is a clinically meaningful difference. The investigators form a prior distribution for θ, $\Pr(\theta)$, before data are collected. Once the data collection is completed (or at any stage of the study), the researchers apply formula (a) to obtain the posterior distribution:

$$\Pr(\theta|\text{data}) \propto \Pr(\text{data}|\theta)^* \Pr(\theta) \quad \textbf{(1)}$$

The posterior distribution obtained from formula (1) therefore contains, from left to right, information on both the prior belief of analysts and the likelihood for the observed data, obtained from the experiment or study. Depending on the specification of the prior distribution and the likelihood function, the posterior distribution may or may not be represented in a closed form. The limited number of closed-form posterior distributions has historically hindered the progress of Bayesian analysis. However, recent developments in computational capabilities allow for more complicated calculations; therefore, interest in this approach has been growing. Another appealing feature of the Bayesian approach is that it allows researchers to "factor expertise and prior knowledge into their computations" (15). Note that many researchers do not wish to incorporate prior beliefs or results from other studies into the analyses; those researchers can use "flat" or "noninformative" prior distributions.

Interesting Probabilities One Can Calculate in a Bayesian Approach

Posterior distributions calculated under the Bayesian approach results are a probability most clinicians think they are getting. Diamond and Forrester (5) asked 24 cardiologists what quantity they would most like to know, and posterior probabilities were rated highest by 19 of the 24 cardiologists. Probabilities of many interesting events can be computed from posterior distributions. For example, in the case of a single endpoint, the Bayesian approach can be used to calculate the probability that drug A is beneficial or that a mortality reduction from drug A is at least 5%. When comparing two treatments on a single endpoint, this approach can compute the probability that drug A is more than 5% better than drug B or that drug A is clinically similar to drug B. Posterior distributions can also be used to combine multiple endpoints simultaneously (e.g., the probability that a treatment leads to both mortality and cost reduction, that a treatment has more than $5,000 reduction in costs and either a mortality reduction or a higher than 0.02 morbidity reduction, or that the

incremental cost-effectiveness ratio is less than $30,000 per life-year saved). In addition, the Bayesian approach places emphasis on estimation and graphical presentation rather than hypothesis testing, making its results more straightforward to interpret.

Calculations of the aforementioned probabilities within the framework of the classical frequentist approaches, i.e., without a prior distribution for the unknown parameters, is not allowed in the classical approach. However, it is relatively easy using the Bayesian approach. One can easily closely approximate probabilities of complex events using simulations from posterior distributions. For example, if one is interested in the probability that an odds ratio (OR) is less than 0.9, one can sample 10,000 ORs from the posterior distribution of θ and then compute the fraction of these 10,000 realizations falling below 0.9. In fact, once a stream of realizations is drawn, one can calculate statistics such as the mean, median, mode, or credible interval (Bayesian confidence interval) of the posterior distribution of the ORs. As will be seen below, joint probabilities for multiple parameters are easily computed in this way.

It is its unique ability to deal with probabilities involving multiple parameters that makes the Bayesian approach especially appealing to healthcare researchers. Multiple endpoints such as mortality and morbidity are common in clinical studies. For example, a clinical study may involve three endpoints. Because it is very difficult to account for clinical significance in the classical approach, it is common to assess the evidence for treatment effects in terms of nonzero differences in the three endpoints between the two treatments. In the Bayesian approach it is simple to assess evidence for actual clinical targets for the endpoints. Suppose that target A_1 is achieved if the population mean blood pressure decreased 5 mm Hg or more, A_2 is achieved if the population exercise time increased by at least 1 minute, and A_3 is achieved if the population mean angina score decreased by at least one point. Suppose this study provided posterior probabilities of A_1, A_2, and A_3 of 0.97, 0.94, and 0.6, respectively. Thus, the probability that at least one study target was achieved is at least 0.97. Decision makers may be interested in probabilities such as having the majority of targets achieved.

Using the Bayesian approach, results from studies with multiple endpoints can be reported in any way that decision makers wish to present (18). Following the above example, researchers may assign different scores (weights) to different targeted endpoints, and evaluate the treatment alternatives based on the scoring system. For example, if one thinks that the most important endpoint is the mean angina score (e.g., A_3), one may assign 1, 2, and 3 points to target A_1, A_2, and A_3, respectively. Then one can examine the probability of reaching at least three points, which can be achieved by hitting only A_3, any two targets, or all three targets. Alternatively, one can estimate the expected number of efficacy endpoints for which the treatment is successful, which for the above example is: $2.51 = \Pr(A_1) + \Pr(A_2) + \Pr(A_3) = 0.97 + 0.94 + 0.6$.

Any other combinations of the multiple endpoints can be estimated from the joint Bayesian posterior distribution of all endpoints, e.g., draw 20,000 realizations of treatment effects on three responses and count the number of triplets meeting efficacy targets. The flexibility in the type of information one can report is an extremely attractive feature of the Bayesian approach.

EXAMPLES FROM CLINICAL STUDIES

The GUSTO Study

The Global Utilization of Streptokinase and t-PA (tissue plasminogen activator) for Occluded Coronary Arteries (GUSTO-I) trial was an international trial to examine the effects of four thrombolytic strategies on the mortality of patients with acute myocardial infarction (AMI): a) streptokinase with subcutaneous heparin, streptokinase with intravenous

heparin, accelerated t-PA with intravenous heparin, and streptokinase with t-PA and intravenous heparin (10). A total of 41,021 patients from 1,081 hospitals in 15 counties were enrolled in the GUSTO trial. Using the combined endpoints of death and nonfatal disabling stroke, the GUSTO investigators found significant reductions in risk when comparing accelerated t-PA with two streptokinase strategies together (OR = 0.88; two-tailed p value = .006). They concluded that accelerated t-PA combined with intravenous heparin is the best thrombolytic strategy for patients with AMI.

Brophy and Joseph (2) challenged the above result when they reanalyzed GUSTO-I using a Bayesian approach. They used varying mixtures of two prior distributions: a) a "lump" prior specifying that there is a nonzero probability that the difference between t-PA and streptokinase is exactly zero (when instead we view the unknown treatment effect to have a continuous distribution); and b) a continuous prior from two previous international trials of nonaccelerated dosing of t-PA, GISSI-2 (12) and ISIS-3 (13).[1] We used two different smooth prior distributions to represent two types of beliefs regarding the odds ratio of t-PA versus streptokinase: uninformative and skeptical. An uninformative prior describes the situation where investigators believe that their previous experiences or knowledge prior to the trial did not provide much information regarding the odds ratio. We use a Gaussian distribution with large variance, 10^6, for log(OR) to represent an uninformative prior. Another prior distribution, a skeptical prior, represents the case where the investigator thought it relatively most likely that there is no effect of t-PA, that a very large effect (positive or negative) is impossible, and that a large effect is unlikely. This skeptical prior was summarized by a truncated Gaussian distribution for log(OR):Pr(OR > 4 or OR < 1/4) = 0 & Pr(OR > 4/3 or OR < 3/4) = 0.05.

Figure 1 depicts the posterior distribution for the log odds ratio for death or disabling stroke, using an uninformative prior and the GUSTO trial data. It shows that the posterior

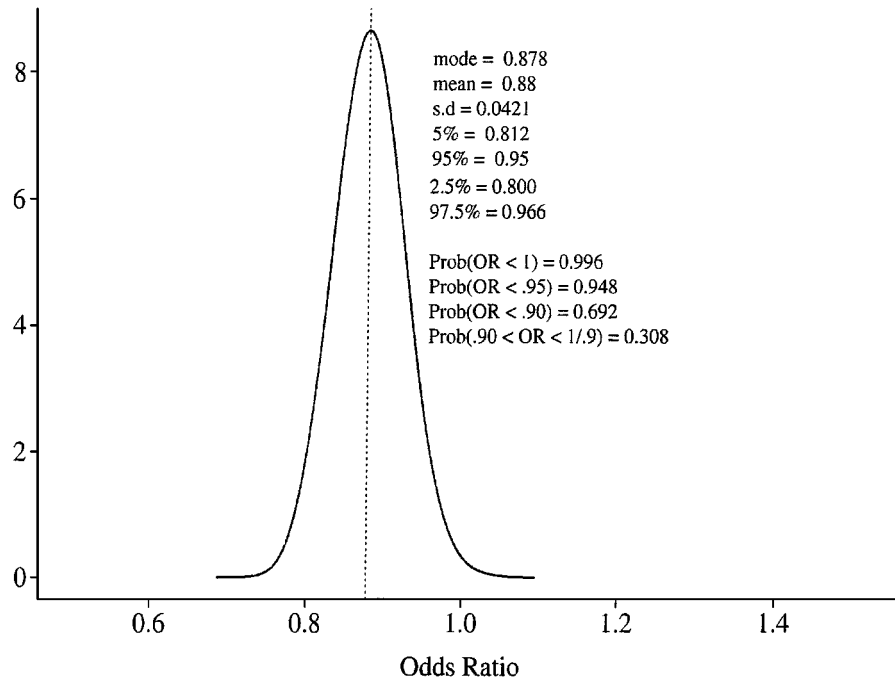

Figure 1. Posterior probability density for accelerated t-PA compared with streptokinase, using a flat prior for log OR.

mean OR is 0.88, which is the same as that reported in the GUSTO trial. The identical OR is reported in the Bayesian example with uninformative prior and the classical approach-based GUSTO trial because in the case of the uninformative prior, investigators rely on the likelihood function (e.g., the trial data) as the only source of information, which is then reflected in the posterior distribution. As discussed earlier, the advantage of a Bayesian approach is its flexibility in reporting evidence. Such an advantage is illustrated in Figure 1. The posterior distribution can answer the efficacy question for different definitions of efficacy. For example, if efficacy is defined as lower combined risk (i.e., OR < 1), then the probability that t-PA is better than streptokinase is 0.996, a strong evidence supporting t-PA. If efficacy is defined as risk reduction of at least 5%, the probability reduces to 0.948. For a more aggressive efficacy measure such as risk reduction of at least 10%, such probability further reduces to 0.692. Note that unlike the basic classical approach where efficacy has to be defined in the trial designing stage so as to "power" the study and the trial cannot be stopped until reaching a prespecified sample size, the Bayesian approach allows intermittent analysis of a trial (with no penalty, as Bayesian inference does not involve type I errors) to determine whether adding more patients is likely to change the outcome (15).

Efficacy can also be defined using the concept of similarity region, which examines how similar (or different) one treatment is from the other. For example, if risk reduction is within 10% from the treatment alternative (e.g., OR $\in [0.9, 1/0.9]$), then the two treatments are defined as similar. Figure 1 shows that the probability that t-PA and streptokinase fall into the similarity region is 0.308.

If the second skeptical prior is used to construct the posterior distribution, Figure 2 shows that the probability of OR in the similarity region rises to 0.396. In addition, the

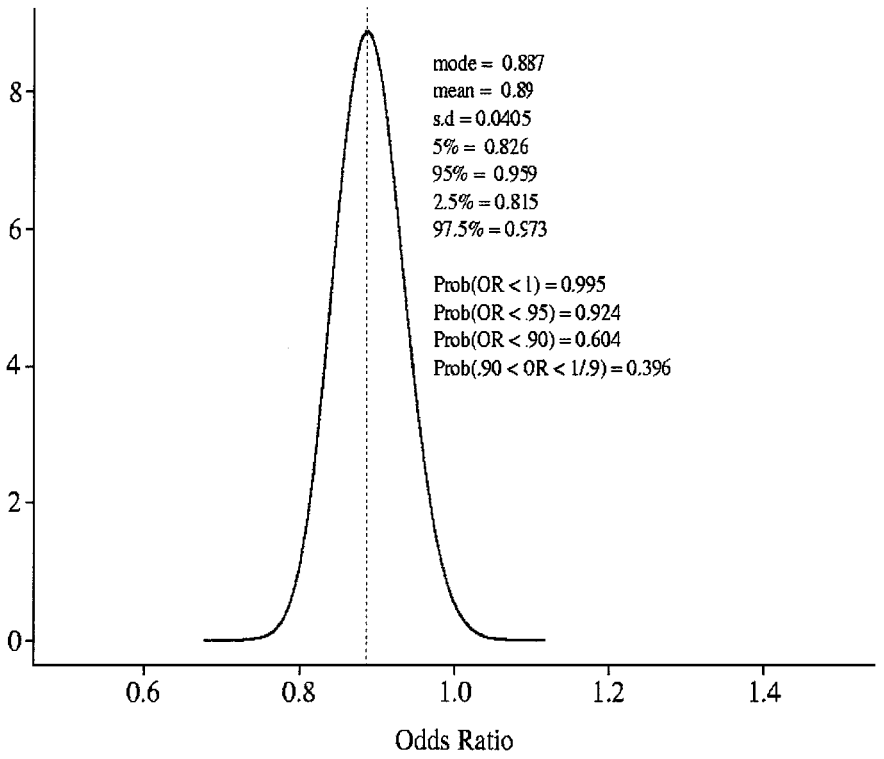

Figure 2. Posterior probability density for accelerated t-PA compared with streptokinase, using a prior distribution that assumed that $\Pr(\text{OR} > 1\frac{1}{3}) = \Pr(\text{OR} < \frac{3}{4}) = 0.025$.

probability of achieving at least 5% risk reduction (i.e., Prob[OR < 0.95]) is 0.924, and the probability of at least 10% risk reduction (i.e., Prob[OR < 0.90]) is 0.604; that is, researchers' "skepticism" *a priori* leads to slightly stronger support of similarity but slightly weaker support of having substantial risk reduction.

Short-acting Nifedipine and Dose-related Mortality

Furberg et al. (7) conducted a meta-analysis from 16 randomized secondary-prevention trials to assess the effect of the dose of short-acting nifedipine on mortality in patients with coronary heart disease. A dose-related increase in mortality was supported by this study. Opie and Messerli (16) later challenged the conclusion of the study by Furberg et al. (7) and argued that "clinicians could continue to use calcium antagonists for approved indications...."

We took a Bayesian approach to revisit the above issue. We assumed a linear dose-response (this was checked by other analyses) and took the quantity of interest to be the odds ratio of 100 mg nifedipine versus placebo (e.g., 0 mg) for all-cause mortality. As in the above example, we used a truncated Gaussian distribution of log(OR) to represent our prior distribution. Specifically, this skeptical prior distribution is described as: Pr(OR > 4 or OR $< 1/4) = 0$ & Pr(OR > 2 or OR $< 1/2) = 0.05$ (Figure 3). Using the skeptical prior tilted toward no mortality effect, the posterior distribution presented the following evidence. First, by defining the similarity zone (e.g., equivalence) as having the log OR between -0.05 and 0.05, our posterior distribution showed that the probability that 100 mg and

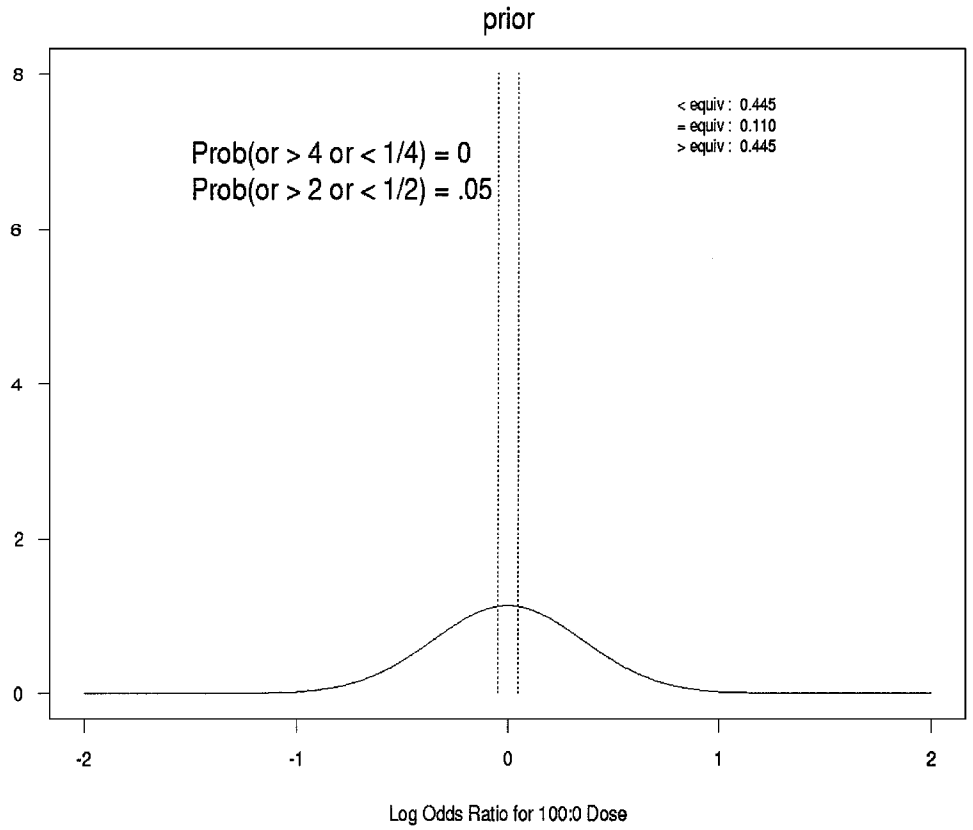

Figure 3. Skeptical prior density for log OR; similarity ("equivalence") zone is log odds $\in (-0.05, 0.05)$.

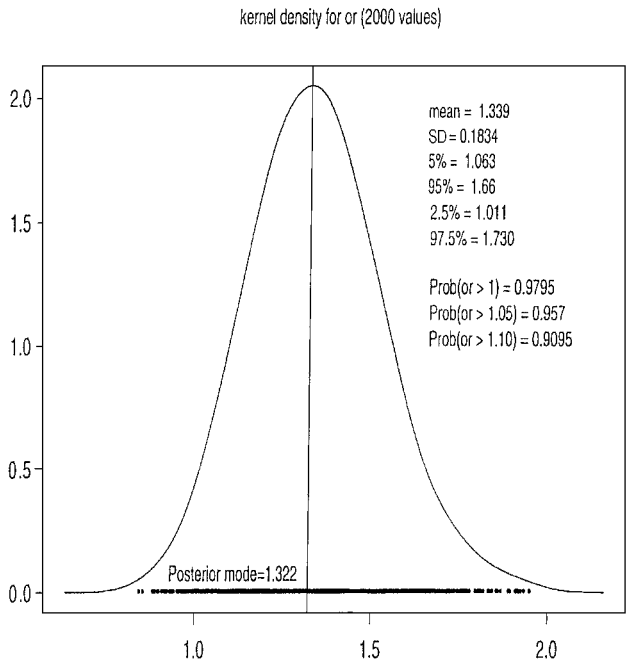

Figure 4. Posterior density for pooled 100 mg : 0 mg nifedipine OR using a skeptical prior tilted toward no mortality effect.

0 mg nifedepine had similar risk of mortality is 0.110. Second, the posterior distribution indicated that the probability of any harm (i.e., Prob(OR > 1) is 0.9795 (Figure 4). We found that the probability that 100 mg and 0 mg nifedepine have a similar mortality risk is quite low, even with a skeptical prior. However, the probability of any harm from 100 mg nifedepine, compared with the placebo group, is quite high. There is also a high probability of at least a 10% increase in the odds of death. Therefore, our findings suggest that there is strong evidence supporting an increase in mortality.

SUMMARY

Probabilities computed using the Bayesian approach are useful and easy to interpret. Using this approach, one can calculate the probability of any efficacy, of clinically important efficacy, of similarity, noninferiority, or of any harm. The Bayesian approach can compute probabilities related to multiple endpoints that are of greater interest to clinicians, patients, and policy makers than probabilities associated with type I error given no treatment effect.

NOTE

[1] It is our opinion that the prior distributions used by Brophy and Joseph are problematic. The "lump" prior was used in an attempt to mimic the frequentist single-point null hypothesis testing approach. Discontinuous prior beliefs do not seem realistic. The smooth prior distribution component they used has a different problem, being based on results of prior studies of nonaccelerated dosed t-PA, which GUSTO-I demonstrated was significantly less effective than was the accelerated dosing of t-PA used in GUSTO-I. Hence the previous studies are not "poolable" with GUSTO-I.

REFERENCES

1. Berry DA, Wolff MC, Sack D. Decision making during a phase III randomized controlled trial. *Controlled Clin Trials.* 1994;15:360-378.

2. Brophy JM, Joseph L. Placing trials in context using Bayesian analysis. *JAMA.* 1995;273:871-875.
3. Claxton K, Posnett J. An economic approach to clinical trial design and research priority-setting. *Health Econ.* 1996;5:513-524.
4. Davidoff F. Standing statistics right side up. *Ann Intern Med.* 1999;130:1019-1021.
5. Diamond GA, Forrester JS. Clinical trials and statistical verdicts: Probable grounds for appeal. *Ann Intern Med.* 1983;98:385-394.
6. Drummond MF, O'Brien B, Stoddart GL, Torrance GW. *Methods for the economic evaluation of health care programmes*. 2nd ed. New York: Oxford University Press; 1997:Chapter 8.
7. Furberg CD, Psaty BM, Meyer JV. Nifedipine: Dose-related increase in mortality in patients with coronary heart disease. *Circulation.* 1995;92:1326-1331.
8. Goodman S. Toward evidence-based medical statistics, I: The P value fallacy. *Ann Intern Med.* 1999;130:995-1004.
9. Goodman S. Towards evidence-based medical statistics, II: The Bayes factor. *Ann Intern Med.* 1999;130:1005-1013.
10. The GUSTO Investigators. An international randomized trial comparing four thrombolytic strategies for acute myocardial infarction. *N Engl J Med.* 1993;329:673-682.
11. Haddix AC, Teutsch SM, Shaffer PA, Dunet DO. *Prevention effectiveness: A guide to decision analysis and economic evaluation*. New York: Oxford University Press; 1996:Chapter 3.
12. The International Study Group. In-hospital mortality and clinical course of 20,891 patients with suspected acute myocardial infarction randomised between alteplase and streptokinase with or without heparin. *Lancet.* 1990;336:71-75.
13. ISIS-3 (Third International Study of Infarct Survival) Collaborative Group. ISIS-3: A randomised comparison of streptokinase vs tissue plasminogen activator vs anistreplase and of aspirin plus heparin vs aspirin alone among 4,299 cases of suspected acute myocardial infarction. *Lancet.* 1993;339:753-770.
14. Lacey LF, Walker SG, Claxton K. Selecting treatments: A decision theoretic approach. *J Royal Stat Soc A.* 2000;163:211-225.
15. Malakoff D. Bayes offers a "new" way to make sense of numbers. *Science.* 1999;286:1460-1464.
16. Opie LH, Messerli FH. Nifedipine and mortality: Grave defects in the dossier. *Circulation.* 1995;92:1068-1073.
17. Spiegelhalter DJ. Probabilistic prediction in patient management and clinical trials. *Stat Med.* 1986;5:421-433.
18. Thall PF, Sung H. Some extensions and applications of a Bayesian stragety for monitoring multiple outcomes in clinical trials. *Stat Med*. 1998;17:1563-1580.
19. Weinstein M, Fineberg HV. *Clinical decision analysis*. Philadelphia: W.B. Saunders Company; 1980.