DOCUMENT RESUME

ED 325 480 TM 015 657

AUTHOR Berry, Donald A.
TITLE A Bayesian Approach to Multicenter Trials and Metaanalysis.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE 89
CONTRACT DMS-8911548
NOTE 12p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Bayesian Statistics; Drug Use; Mathematical Models; *Meta Analysis; Pharmacology; *Probability; *Research Methodology
IDENTIFIERS *Clinical Trials; *Multicenter Trials

ABSTRACT

The use of a Bayesian approach in evaluating data from clinical trials with many treatment centers and from many studies is discussed. The main distinction between a metaanalysis and an analysis of a multicenter trial is that different studies may have very different designs, while the centers in a multicenter trial usually follow the same protocol. In particular, different studies in a metaanalysis may involve different treatment comparisons, while centers within the same trial usually consider the same treatments. The Bayesian statistical approach focuses on the probability distribution of any unknowns given the available information. An advantage of Bayesian methods is that they allow the use of all available information. In the case of a multicenter clinical trial of the effects of a particular drug, Bayesian methods require assessment of the information available before the trial as a probability distribution. A comparison of the Bayesian and frequentist approaches indicates that Bayesian methods have greater flexibility. Results from nine studies of an antidepressant drug illustrate the Bayesian hierarchical approach. Bayesian updating for multiple treatment studies is the same as for a single treatment. Three tables present study data and 15 figures illustrate the analysis. (SLD)

A BAYESIAN APPROACH TO MULTICENTER TRIALS AND METAANALYSIS

Donald A. Berry
University of Minnesota

A BAYESIAN APPROACH TO MULTICENTER TRIALS AND METAANALYSIS

Donald A. Berry*, School of Statistics
270 Vincent Hall, Univ. of Minnesota, Minneapolis MN 55455

Abstract

Bayesian inference focuses on questions of interest to clinicians: In view of the available information, is the therapy effective? how effective? How does the response depend on the type of patient? on the treating physician? The Bayesian approach requires that the statistician use all available information in drawing conclusions. This makes the approach ideal for analyzing data from many centers and for metaanalyses. I will describe a hierarchical Bayes approach to analyzing such data.

Key Words: Assessing prior probabilities; Hierarchical Bayesian analyses; Mixtures; Center effects.

1. Introduction

An analysis of data from more than one study is a metaanalysis. The main distinction between a metaanalysis and an analysis of a multicenter trial is that different studies may have very different designs while the centers in a multicenter trial usually follow the same protocol. In particular, different studies in a metaanalysis may involve different treatment comparisons while centers within the same trial usually consider the same treatments.

Different studies in a metaanalysis often deal with different types of patients. So it is not too surprising that they frequently show different treatment effects. Multicenter trials are similar in the sense that different centers may well have different patient populations. This is in part because investigators at the individual centers may interpret the trial's patient inclusion/exclusion criteria differently and so end up with different types of patients. Sometimes it is possible to account for such a difference using measurable covariates, and sometimes not.

2. Bayesian Approach

The focus of the Bayesian approach is the probability distribution of any unknowns given the available information. In particular, Bayesian inference deals with probabilities of hypotheses and probability distributions of parameters. A conclusion of a hypothesis test of equality of treatment means, say, is the probability that the means are equal in view of the data. And one can calculate the probability that the true mean difference is contained in any interval.

Probabilities of hypotheses that are conditioned on data from an experiment are *posterior probabilities*. Calculating a posterior probability requires Bayes' theorem:

$$P(H|\text{data}) \propto P(\text{data}|H)P(H),$$

where H is any hypothesis and P(Hldata) is the *likelihood function* evaluated at H. So Bayes' theorem relates the conditional probability of a hypothesis given data to its unconditional probability. The latter depends on information present *before* the experiment, and so is called a *prior probability*.

The designations "prior" and "posterior" refer to a particular experiment. Probabilities between experiments are posterior to the previous experiment and prior to the next one—in the words of the Bard, "What's past is prologue." So perhaps it would be better to use "current" in place of both "prior" and "posterior".

Bayesian inference is not merely data analysis to be applied to a particular trial. Rather, Bayes' theorem provides a formalism for learning: "That's what I thought before, this is what I've just seen, so here's what I now think; and I may learn something more tomorrow."

An advantage of Bayesian methods is that they allow for using all available information. This characteristic makes such methods ideal for analyzing data from clinical trials with many centers and from many studies, though a Bayesian analysis in such cases may not be easy.

Consider a drug whose effect in some population is not completely known. A multicenter clinical trial is contemplated. Bayesian methods require assessing the information available before the trial (along with its associated uncertainty) as a probability distribution. This prior distribution depends on the person doing the assessing, and so is

* Research supported in part by the National Science Foundation under grant DMS 8911548.

subjective. Since posterior probabilities depend on prior probabilities, posterior probabilities are also subjective. A common (though not unanimous) view among Bayesians is that *all* probabilities are subjective. An advantage of the subjective view is that probabilities apply in any setting in which a person has an opinion. Counting ignorance as an opinion, though obviously a very weak one, this includes every setting.

3. Bayesian vs. Frequentist Approach

In the frequentist Neyman-Pearson approach, analysis and design are tied together, the design dictating the analysis. Strictly speaking, a frequentist analysis is not possible when the design is not known—this can lead to nonsense (Berry 1987). And tying the analysis to the design means that when several separate experiments are conducted to address the same question, each has to be analyzed separately. The consumer is left with the task of combining them, and with little guidance from frequentist methods.

Bayesian methods are more flexible. Data from clinical trials affect Bayesian inferences only through the likelihood function. Any multiplicative constants in the likelihood function are irrelevant—see Bayes' theorem. This means that the stopping rule and other such characteristics of the design are also irrelevant. So Bayesian methods allow continual or periodic data analysis without penalties such as those imposed by classical inferences (Berry 1987). In particular, available data can be taken at face value when deciding whether to stop or continue a trial, or otherwise change its design. For example, a decision to continue a multicenter trial may include stopping patient accrual in some centers while continuing it in others.

There is an attitude among some frequentist statisticians that says that centers in a multicenter trial can be pooled in a single analysis only if their results are similar. If the data cannot be pooled then the sample in each center has to be large enough to "stand on its own." I hope and believe that this is a minority view. All the data contain evidence about the safety and efficacy of a drug, so the results from the individual centers must be combined in some way and by someone to come to a single conclusion. This is difficult but not impossible using frequentist ideas; it is required in the Bayesian paradigm.

Small trials have small power: they provide the ability to reject the null hypothesis of no treatment difference with high probability only if one treatment is much better than the other. Some frequentist statisticians complain about trials that are too small to have a reasonable "chance of detecting differences of therapeutic value." (Mosteller, Gilbert, and McPeek 1983). Some contend that many actual trials are too small to be worthwhile, and recommend large trials (Peto, Pike, Armitage, et al. 1976). Some even suggest that it is unethical to conduct a small trial since some of the patients will be exposed to inferior treatment with little hope of rejecting a false null hypothesis. This is true, but it's not the point. Many researchers view science differently from the way frequentists view science. If a study turns out to be too small to be conclusive then the researchers can conduct another study (or studies) and combine results—the first study is never wasted (as long as it was honestly conducted). A piecemeal approach allows researchers to digest information as it becomes available and decide whether further investigation is appropriate. If the experimental treatment turns out to be clearly bad or clearly good, they can stop. And if the data are equivocal, then it may be reasonable to continue experimenting. Such an approach has the additional advantages of revealing any variation over time and of showing reproducibility of results.

There are many small trials in medicine today because the associated flexibility is important to clinicians. They bypass statistics as they know it, with its obscure P values, in favor of addressing the important questions: Is this treatment effective? Is drug A better than drug B for Ms. Smith? They answer these questions in an informal, subjective and usually private way, using all the information at their disposal. The Bayesian approach provides a formalism for addressing these questions that is not unlike the informal way that clinicians are forced to do it now.

4. Hierarchical Approach to Assessing Center Effect

I have indicated that Bayesian updating can take place at any time. Such updating requires a likelihood function: the conditional probability of the current data given the unknown parameters.

I will describe a Bayesian hierarchical approach (Lindley and Smith 1972; Berger 1986). Individual centers have unknown characteristics that set them apart from the

other centers. Like all unknowns in the Bayesian approach, these are random. So the Bayesian approach gives rise to a random effects model.

Think of each center as having a particular distribution of patient responses for each therapy. Selecting a center means selecting one of these distributions. If the distribution of the selected center were to be revealed this would give us direct information about how the center distributions are themselves distributed, and we would have a standard statistics problem, one that could be addressed by either Bayesian or frequentist methods. But since each center contributes only a finite number of patients, the individual center distributions are not revealed. Instead, we observe only a sample from each center's distribution. This gives indirect information about the distribution of center distributions.

Consider a simple analogy. A bag contains several thousand coins. The coins may have different probabilities of heads. We'd like to know something about the distribution of probabilities of heads among these coins. So we select ten of them, and toss each of the ten a total of 30 times. Our data consist of ten sample proportions of heads (along with the ten sample sizes). If these proportions are wildly different then the coins in the bag must have different probabilities of heads, though perhaps not as different as the sample proportions among the ten coins tossed. And if the sample proportions are quite similar then the coins in the bag may have similar probabilities of heads. In any case, the sample proportions give information about the distribution of probabilities of heads among the coins in the bag.

5. Example with Dichotomous Responses

Table 1 gives results for nine studies (Janicak, Lipinski, Davis, et al. 1988) involving the anti-depressant drug S-adenosylmethionine (SAMe). The number of patients in study i is n_i and the number successes is x_i. These data were part of a metaanalysis, but you may also think of them as coming from a multicenter trial. Suppose that the patients in center i are exchangeable in the sense that all had the same probability p_i of success. (For a Bayesian analysis in the presence of differing prognoses, see Berry (1989).)

TABLE 1: Successes observed on the antidepressant drug S-adenosylmethionine

i	x_i	n_i	$\hat{p}_i = x_i/n_i$
1	20	20	1.00
2	4	10	0.40
3	11	16	0.69
4	10	19	0.53
5	5	14	0.36
6	36	46	0.78
7	9	10	0.90
8	7	9	0 78
9	4	6	0.67
Totals	106	150	0.71

The likelihood function of $(p_1, p_2, \ldots, p_9)$ is

$$\Pi_{i=1}^{9} p_i^{x_i}(1-p_i)^{n_i-x_i}.$$

A combined analysis assumes that all 150 patients are exchangeable, so that the nine p_i's are equal (with common value p, say). The likelihood function of p is then

$$p^{106}(1-p)^{44},$$

which is shown in Figure 1.

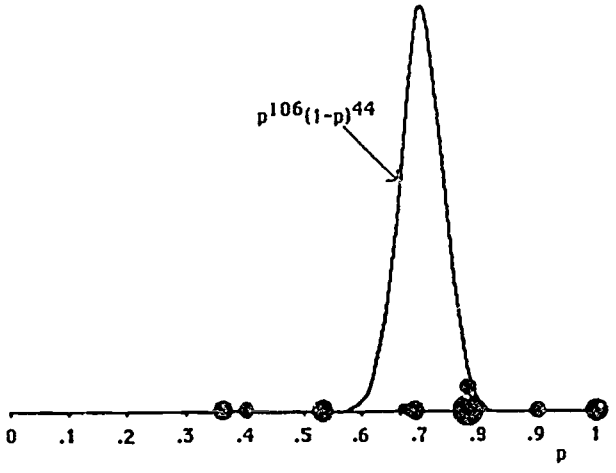

Figure 1. Likelihood function of p for the data in Table 1 assuming the 150 patients in the nine studies are exchangeable. (The nine dots correspond to the observed proportions for the nine studies, with dot areas proportional to sample sizes.)

This figure shows that p is very likely to be between 60 and 80%. This conclusion is

somewhat curious since, as shown by the dots on the p-axis in Figure 1, the observed success proportions in five of the nine studies are outside this range. While sampling variability accounts for some differences, the variability in Table 1 is greater than would be expected from sampling alone. This suggests that the p_i's may not be equal.

Separate analyses of the nine studies is even less satisfactory than the above combined analysis. The effect of the drug is not well addressed by giving nine different likelihood functions, or by giving nine different confidence intervals. Suppose one would like to know the probability of success if the drug were given to a patient in a tenth center. How should the results in these nine centers be weighed? And how should the other eight centers be weighed when considering the next patient to be treated at one of these nine centers?

The Bayesian hierarchical perspective is that each center's success proportion is selected from some population. To quantify information about the population requires a probability distribution of population distributions.

Suppose $p_1, \ldots, p_9$ is a random sample from population distribution F which is itself random. Assume F is a beta distribution with parameters a and b, where a and b are unknown. An observation p from F has density

$$(1) \qquad B(a,b)\, p^{a-1}(1-p)^{b-1},$$

where $a > 0$ and $b > 0$ and

$$B^{-1}(a,b) = \int_0^1 p^{a-1}(1-p)^{b-1}\,dp.$$

The variance of such an observation is

$$\frac{ab}{(a+b)^2(a+b+1)}.$$

So if a + b is large, the distribution of the p's is highly concentrated and consequently there is little center effect. While if a + b is small, the p's will tend to be spread out, and there is a large center effect.

Like all unknowns in the Bayesian approach, the user must assess a probability distribution for a and b; call it $\pi(a,b)$. If the user has information suggesting that there is little center effect then much of the π-probability should be concentrated on large values of a and b, and if information suggests the possibility of substantial center differences then much of the π-probability can be placed on small values of a and b. The prior distribution π can be discrete or continuous, though in the examples below I will assume it is discrete.

Consider a generic sample, say p, on F. Suppose it were possible to observe p. Call $\pi'(a,b|p)$ the posterior distribution of (a,b) given p. From Bayes' theorem,

$$\pi'(a,b|p) \propto B(a,b)\, p^{a-1}(1-p)^{b-1}\, \pi(a,b).$$

Extending this to observing a sample $p_1, \ldots, p_9$,

$$(2) \qquad \pi'(a,b|p_1, \ldots, p_9) \propto \prod_{i=1}^{9} \{ B(a,b) p_i^{a-1}(1-p_i)^{b-1} \}\, \pi(a,b).$$

In Section 7 I will assume $p_1 = \ldots = p_9 = 1/2$ and evaluate (2) for the two special cases of $\pi(a,b)$ given in Section 6.

Now consider an observation on x, a binomial variable with parameters p and n. Such an x contains only indirect information about F. Call $\pi^*(a,b|x)$ the posterior probability distribution of a and b given x and n. From Bayes' theorem,

$$\pi^*(a,b|x) \propto f(x|a,b)\, \pi(a,b),$$

where

$$f(x|a,b) = \int \binom{n}{x} p^x(1-p)^{n-x}\, B(a,b)\, p^{a-1}(1-p)^{b-1}\,dp = \binom{n}{x} \frac{B(a,b)}{B(a+x,b+n-x)}.$$

Therefore,

$$\pi^*(a,b|x) \propto \frac{B(a,b)}{B(a+x,b+n-x)}\, \pi(a,b).$$

Upon observing a sample $x_1, \ldots, x_9$, where the x_i are binomial variables with parameters p_i and n_i and $p_1, \ldots, p_9$ is a random sample from F,

$$(3) \qquad \pi^*(a,b|x_1, \ldots, x_9) \propto \prod_{i=1}^{9} \left\{ \frac{B(a,b)}{B(a+x_i,b+n_i-x_i)} \right\} \pi(a,b).$$

As $n_i \longrightarrow \infty$, the limit of this expression is the expression in (2), with x_1/n_1 set equal to p_i.

Consider the response of an as yet untreated patient. First suppose the patient is treated with SAMe at one of the centers

considered in Table 1. Given the results in Table 1, the probability of success for a patient treated at center i, for i = 1, 2, . . . , 9, is

$$(4)\quad E(p_i|x_1, \ldots, x_9) = E\left\{\frac{a+x_i}{a+b+n_i}\,|x_1, \ldots, x_9\right\}.$$

This expectation is with respect to distribution (3). On the other hand, if the patient is treated with SAMe at a new center—call it center 10—then

$$(5)\quad E(p_{10}|x_1, \ldots, x_9) = E\left\{\frac{a}{a+b}\,|x_1, \ldots, x_9\right\}.$$

This is just the expected posterior mean.

In Section 7 I will evaluate (3), (4), and (5) for the data in Table 1 assuming the two different forms for $\pi(a,b)$ given in the next section. I will also evaluate the expected posterior density of the p's.

6. Assessing Priors for Beta Parameters

There are two general attitudes toward selecting prior probabilities $\pi(a,b)$. One is "subjective" and the other I will call "objective", for want of a better word. The subjective approach assumes a particular assessor. I will give an example in which an assessor is quite confident that there is substantial variability among centers and so assigns most of the prior probability with small values of a + b. It is not appropriate for employees of a pharmaceutical company to use their prior probabilities in filing a new drug application to the Food and Drug Administration, say. But it is appropriate to consider various types of assessors and show how the available information may be used to update each assessor's opinions. An alternative is to use various types of objective prior distributions.

I believe that every inference is subjective, and that prior probability assignments cannot be objective. However, "objective" is sometimes used to describe "uninformative priors", which are uniform in some parametrization. When prior probabilities are uniform, the posterior probabilities are proportional to the likelihood function on those points where the prior probabilities are positive.

Consider the example of the previous section. Suppose an assessor's best estimate of the effectiveness of SAMe over all centers is 50%. Moreover, the assessor's tentative opinion is that the distribution F of success proportions over centers is that they are uniformly spread on the interval (0,1)—the beta (1,1) distribution shown in Figure 2. This is not only the assessor's prior estimate of F, it has about 40% of the assessor's probability: $\pi(1,1) = 0.40$. Figure 3 shows that 15% of the assessor's probability is associated with each of the two densities with a + b = 3: $\pi(2,1) = \pi(1,2) = 0.15$. It happens that the average of these two densities is assessor's prior estimate: beta (1,1). Figure 4 shows that 5% of the assessor's probability is associated with each of the three densities with a + b = 4: $\pi(3,1) = \pi(2,2) = \pi(1,3) = 0.05$, and again the average of these densities is also assessor's prior estimate: beta (1,1). And so on.

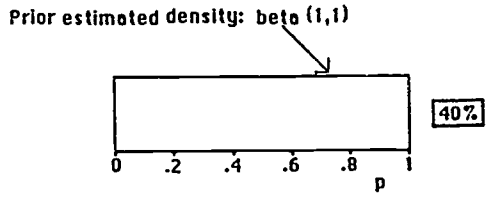

Figure 2. Assessor's prior estimate of population distribution of success proportions. This is a mixture of beta (a,b) distributions that happens to be itself a beta density: a = b = 1. The beta (1,1) density has 40% of the prior probability.

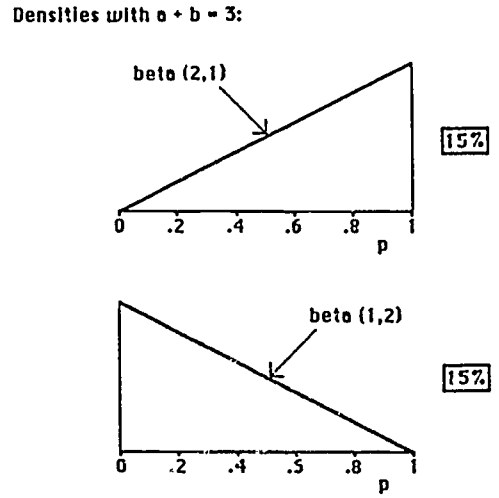

Figure 3. The assessor's prior probability of each of these two densities with a + b = 3 is about 15%. The average of the beta (2,1) and beta (1,2) densities happens to equal the prior estimate, the beta (1,1) density.

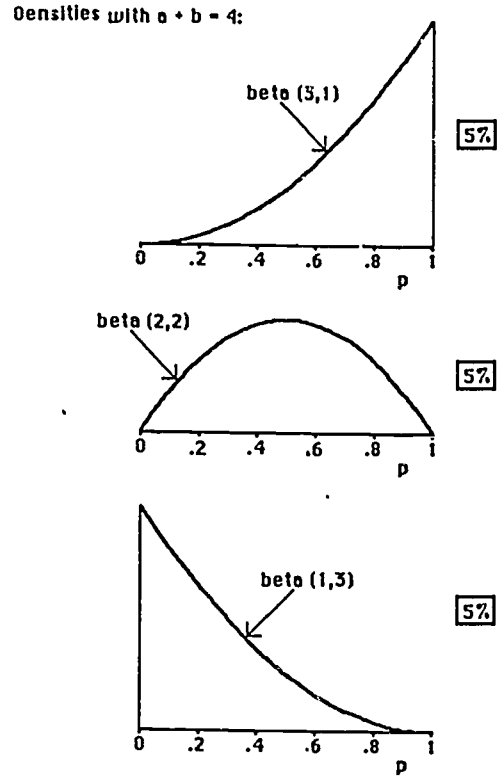

Figure 4. The assessor's prior probability of each of these three densities with a + b = 4 is about 5%. The average of the beta (3,1), beta (2,2), and beta (1,3) densities happens to equal the prior estimate, the beta (1,1) density.

The joint distribution $\pi(a,b)$ implicit in the previous paragraph is the product of two independent geometric variables:

(6) $\pi(a,b) \propto \exp\{-a-b\}$ for a, b = 1, 2, ...

This distribution is pictured in Figure 5.

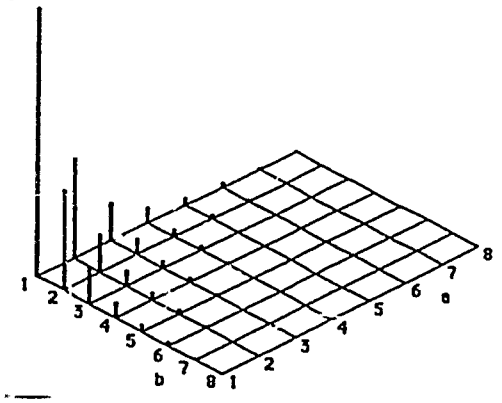

Figure 5. Independent geometric distributions on a and b—formula (6); the points with the six largest probabilities correspond to the densities shown in Figures 2, 3, and 4.

The other distribution of prior probabilities considered in the next section is the product of two independent uniform variables:

(7) $\pi(a,b) \propto 1$ for a, b = 1, 2, . . . , 10.

(An alternative is the uniform distribution on a and b with their sum restricted, say $a + b \leq 20$.) Prior distribution (7) is pictured in Figure 6. This distribution associates a reasonably large probability with a + b large and also with a + b small. Distribution (7) gives substantial probability to a and b nearly equal, and corresponds to a stronger opinion that the p's will tend to be near 1/2 than under distribution (6). As indicated above, for any uniform distribution the posterior probabilities are proportional to the likelihood function on the lattice points where the prior probabilities are positive.

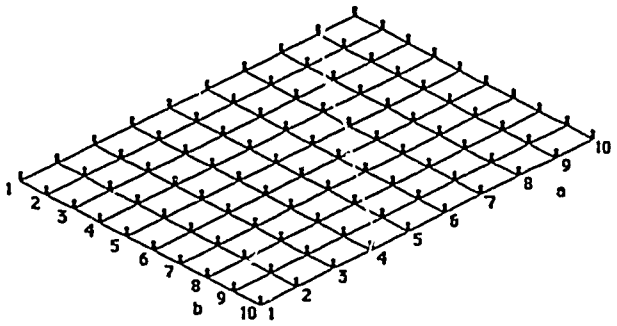

Figure 6. Uniform probabilities for beta parameters (a,b) for a = 1, . . . ,10 and b = 1, . . . ,10—formula (7).

7. Calculations When Observing Each $p_i = 1/2$

Suppose it were possible to observe actual population success proportions, and that for nine studies they all equal 1/2. The posterior probabilities of (a,b) are then calculated from (2) with $p_1 = \ldots = p_9 = 1/2$. (As indicated in Section 5, this hypothetical circumstance is approximated by $x_1/n_1 = \ldots = x_9/n_9 = 1/2$ with $x_i/n_i \longrightarrow \infty$.)

Figures 7 and 8 show the posterior probabilities assuming $\pi(a,b)$ given by (6) and (7), respectively. Figure 8 reflects the fact that the data are consistent with study homogeneity, and that most of the success proportions are close to 1/2. Figure 7 shows that the prior distribution given in (6) is so heavily weighted in favor of small a + b—heterogeneity among the studies— that in

the face of substantial contrary evidence, small values of a + b continue to weigh heavily.

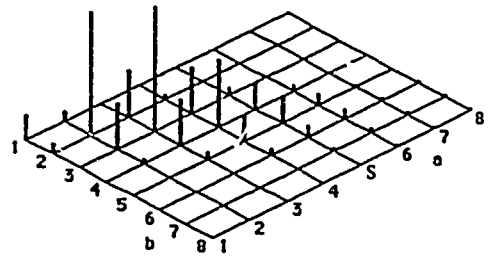

Figure 7. Posterior probabilities $\pi'(a,b|p_1, \ldots, p_9)$ calculated from (2), assuming $\pi(a,b)$ given by (6) and shown in Figure 5, and with $p_1 = \ldots = p_9 = 1/2$.

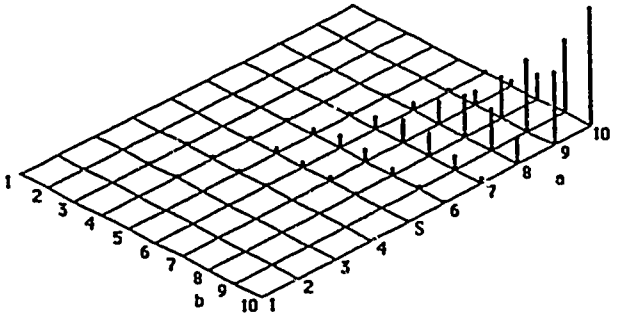

Figure 8. Posterior probabilities $\pi'(a,b|p_1, \ldots, p_9)$ calculated from (2), assuming $\pi(a,b)$ given by (7) and shown in Figure 6, and with $p_1 = \ldots = p_9 = 1/2$. (Compare Figure 7.) When observing these values of the p_i, the likelihood function increases as a and b increase, with a = b. On this restricted set, the maximum likelihood estimate of (a,b) occurs at (10,10), the point with highest posterior probability assuming a uniform prior distribution.

Figures 9 and 10 show the posterior density estimates assuming $p_1 = \ldots = p_9 = 1/2$ and the prior distributions $\pi(a,b)$ given by (6) and (7), respectively. These are averages of beta densities, where the respective averages are with respect to the distributions shown in Figures 7 and 8. Again, it is evident from these two figures that geometric prior (6) is more resistant to data suggesting that the studies are homogeneous than is uniform prior (7).

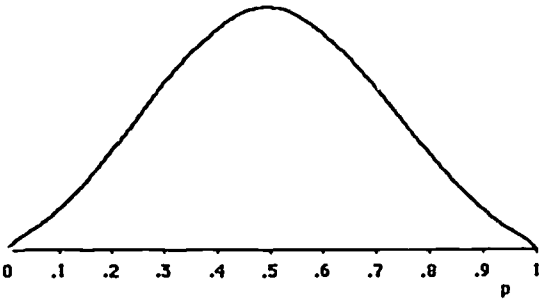

Figure 9. Posterior density estimate of population success proportions for the posterior distribution of beta parameter given in Figure 7, which assumes $p_1 = \ldots = p_9 = 1/2$. This is not itself a beta density but is a mixture of beta densities. It is similar to the beta (2,2) and beta (3,3) densities because, as is evident from Figure 7, there is substantial posterior probability on these (a,b).

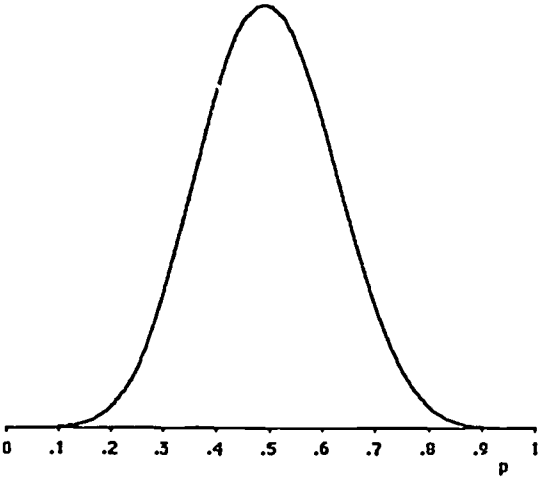

Figure 10. Posterior density estimate of population success proportions for the posterior distribution of beta parameters given in Figure 8, which assumes $p_1 = \ldots = p_9 = 1/2$. (Compare Figure 9.)

8. Estimating the Effectiveness of SAMe

Consider the data in Table 1. The posterior probabilities of (a,b) can be calculated from (3). Figures 11 and 12 show these probabilities assuming prior distributions $\pi(a,b)$ given by (6) and (7), respectively.

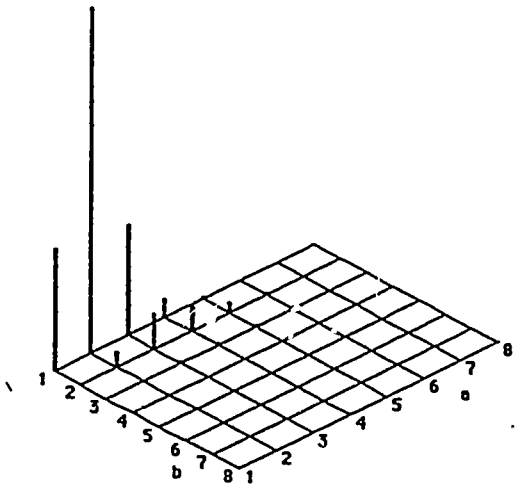

Figure 11. Posterior probabilities $\pi^*(a,b|x_1,\ldots,x_9)$ calculated from (3), assuming $\pi(a,b)$ given by (6) and shown in Figure 5, and conditioning on the results of the nine studies shown in Table 1.

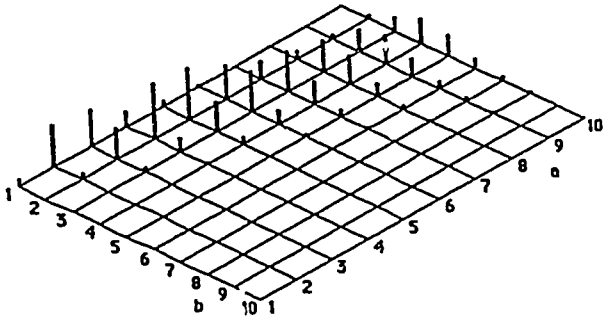

Figure 12. Posterior probabilities $\pi^*(a,b|x_1,\ldots,x_9)$ calculated from (3), assuming $\pi(a,b)$ given by (7) and shown in Figure 6, and conditioning on the results of the nine studies shown in Table 1. (Compare Figure 11.) The maximum likelihood estimate of (a,b) occurs at (4,2), the point with highest posterior probability assuming a uniform prior distribution.

Figures 13 and 14 show the posterior density estimates assuming the data in Table 1 and the prior distributions $\pi(a,b)$ given by (6) and (7), respectively. These are averages of beta densities, where the respective averages are with respect to the distributions shown in Figures 11 and 12. Again, it is evident from these two figures that geometric prior (6) is more heavily weighted toward substantial study heterogeneity than is uniform prior (7). The means of the densities in Figures 13 and 14 are 0.65 and 0.68, which can be evaluated using (5).

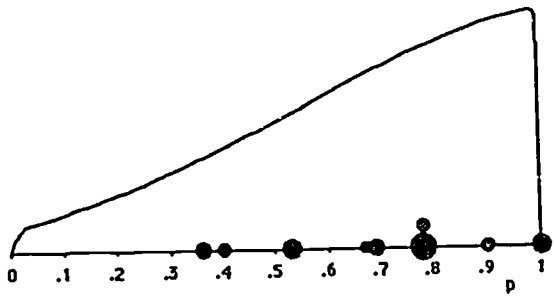

Figure 13. Posterior density estimate of population success proportions for geometric prior (6) and the data in Table 1. The average is with respect to the probability distribution of (a,b) shown in Figure 11. Due to the large posterior probability on a=2, b=1, this estimate is similar to the beta (2,1) density. This estimate should be compared with the likelihood function pictured in Figure 1, which assumes that all 150 patients from these centers can be treated as though they come from a single center. (As in Figure 1, the nine dots correspond to the observed proportions for the nine studies.)

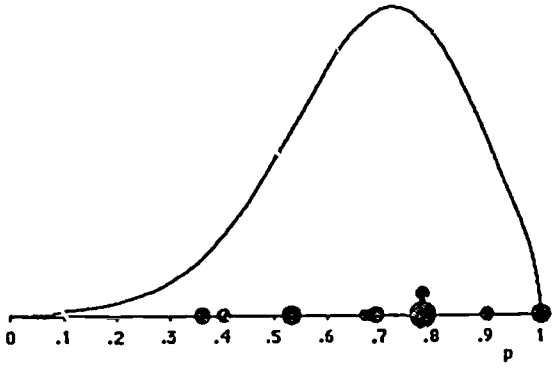

Figure 14. Posterior density estimate of population success proportions for uniform prior (7) and the data of Table 1. This is the average density with respect to the posterior distribution of (a,b) shown in Figure 12. (As in Figure 13, the dots correspond to the observed proportions.)

Table 2 repeats Table 1 and also shows the probability of success for the next patient at each of the nine constituent centers (4) and for a patient at a tenth center (5)—the latter is the overall mean and is shown as the column total. The column headed (6) assumes geometric prior (6) and the one headed (7) assumes

uniform prior (7). The individual center probabilities are shrunk toward the overall mean. This shrinkage is less for the geometric prior because it associates more credence with study heterogeneity than does the uniform prior. Also, as is reasonable, shrinkage to the overall mean is greater for a smaller study.

TABLE 2: Successes observed on SAMe and predictive probabilities of success by center

i	x_i	n_i	$\hat{p}_i = x_i/n_i$	Pred. prob. from (6)	Pred. prob. from (7)
1	20	20	1.00	0.95	0.90
2	4	10	0.40	0.46	0.53
3	11	16	0.69	0.68	0.69
4	10	19	0.53	0.55	0.57
5	5	14	0.36	0.41	0.48
6	36	46	0.78	0.77	0.77
7	9	10	0.90	0.84	0.80
8	7	9	0.78	0.75	0.73
9	4	6	0.67	0.66	0.68
Totals	106	150	0.71	0.65	0.68

9. Comparing Treatments

So far I have addressed a single treatment. Multicenter trials frequently involve two or more treatments. Bayesian updating for multiple treatments is the same as described above for a single treatment. For example, suppose there are two treatments, A and B, and responses are dichotomous. Then F is a bivariate distribution of two success proportions p_A and p_B, and is again random. The calculations are now more complicated. In particular, to allow for center effect it is necessary to include a covariance between p_A and p_B.

I will not extend the analysis of the previous sections but will instead simply refer to DuMouchel (1989), who analyzes the study (Janicak et al. 1988) that reported the data given in Table 1. The full data set is given in Table 3, and in a dot diagram in Figure 15. SAMe was compared in nine randomized trials with either placebo or standard therapy or both. I'll call these treatments A, B, and C, respectively.

TABLE 3: Results from nine clinical trials.

	SAMe (A)		Placebo (B)		Standard (C)	
i	n_{Ai}	x_{Ai}	n_{Bi}	x_{Bi}	n_{Ci}	x_{Ci}
1	20	20	10	1		
2	10	4	10	0		
3	16	11			15	9
4	19	10			10	8
5	14	5			14	4
6	46	36			41	30
7	10	9			10	9
8	9	7			9	6
9	6	4	5	0	4	3
	150	106 (71%)	25	1 (4%)	103	69 (67%)

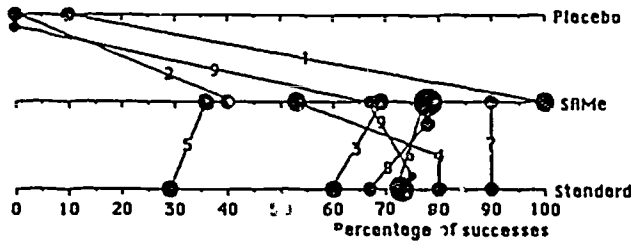

Figure 15. Dot diagram version of data in Table 3. Areas of dots are approximately proportional to sample sizes. Lines connecting dots are labeled with study numbers. The dots for SAMe are the same as in Figures 1, 13, and 14.

Let (p_{Ai}, p_{Bi}, p_{Ci}) stand for the success probabilities of A, B, and C that apply in study i, for $i = 1, 2, \ldots, 9$. As in the case of a single treatment, the nine triples (p_{Ai}, p_{Bi}, p_{Ci}) are not observed. Rather, the data consist of (n_{Ai}, n_{Bi}, n_{Ci}) and (x_{Ai}, x_{Bi}, x_{Ci}), where x_{Ai} is distributed as binomial (n_{Ai}, p_{Ai}), x_{Bi} is binomial (n_{Bi}, p_{Bi}), and x_{Ci} is binomial (n_{Ci}, p_{Ci}). For example, the first row of Table 3 gives $n_{A1} = 20$, $x_{A1} = 20$, $n_{B1} = 10$, $x_{B1} = 1$, $n_{C1} = 0$, and $x_{C1} = 0$.

DuMouchel (1989) considers differences between p's, uses the normal approximation to the binomial, and assumes uniform (hence "improper") priors. He finds that the posterior mean and standard deviation of $p_A - p_B$ are 0.70 and 0.12, those of $p_A - p_C$ are 0.00 and 0.09, and those of $p_C - p_B$ are 0.70 and 0.14. He also describes incorporating historical data and subjective information into the prior distribution.

REFERENCES

Berger, J.O. (1986), *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. Springer-Verlag, New York.

Berry, D. A. (1987). Interim analysis in clinical trials: The role of the likelihood principle. *The American Statistician* 41:117-122.

Berry, D. A. (1989). Monitoring accumulating data in a clinical trial. *Biometrics* 45:1197-1211.

DuMouchel, W. (1989), "Bayesian Metaanalysis," *Statistical Methodology in the Pharmaceutical Sciences*, Marcel Dekker, New York, 509-529. (Ed. by D.A. Berry.)

Janicak, P.G., J. Lipinski, J.M. Davis, J.E. Comaty, C. Waternaux, B. Cohen S. Altman, and R.P. Sharma (1988), "S-Adenosyl-Methionine (SAMe) in Depression: a Literature Review and Preliminary Data Report," *Alabama Journal of Medical Sciences* 25, 306-312.

Lindley, D.V., and A.F.M. Smith (1972), "Bayes Estimates for the Linear Model" (with discussion), *Journal of the Royal Statistical Association, Series B* 34, 1-41.

Mosteller, F., J.P. Gilbert, and B. McPeek (1983), "Controversies in Design and Analysis of Clinical Trials," *Clinical Trials*, Marcel Dekker, New York, 13-64. (Ed. by S.H. Shapiro and T.A. Louis.)

Peto, R., M.C. Pike, P. Armitage, N.E. Breslow, D.R. Cox, S.V. Howard, N. Mantel, K. McPherson, J. Peto, and P.G. Smith (1976), "Design and Analysis of Randomized Clinical Trials Requiring Prolonged Observation of Each Patient: I. Introduction and Design," *British Journal of Cancer* 34, 585-612.

Appendix 16

END

U.S. Dept. of Education

Office of Education
Research and
Improvement (OERI)

ERIC

Date Filmed

March 29, 1991