# Large-Sample Bayesian Posterior Distributions for Probabilistic Sensitivity Analysis

*Gordon B. Hazen, PhD, Min Huang, MS*

*In probabilistic sensitivity analyses, analysts assign probability distributions to uncertain model parameters and use Monte Carlo simulation to estimate the sensitivity of model results to parameter uncertainty. The authors present Bayesian methods for constructing large-sample approximate posterior distributions for probabilities, rates, and relative effect parameters, for both controlled and uncontrolled studies, and discuss how to use these posterior distributions in a probabilistic sensitivity analysis. These results draw on and extend procedures from the literature on large-sample Bayesian posterior distributions and Bayesian random effects meta-analysis. They improve on standard approaches to probabilistic sensitivity analysis by allowing a proper accounting for heterogeneity across studies as well as dependence between control and treatment parameters, while still being simple enough to be carried out on a spreadsheet. The authors apply these methods to conduct a probabilistic sensitivity analysis for a recently published analysis of zidovudine prophylaxis following rapid HIV testing in labor to prevent vertical HIV transmission in pregnant women.* **Key words:** *decision analysis; cost-effectiveness analysis; probabilistic sensitivity analysis; Bayesian methods; random effects meta-analysis; expected value of perfect information; HIV transmission; zidovudine prophylaxis.* ***(Med Decis Making 2006;26:512–534)***

Sensitivity analysis is today a crucial element in any practical decision analysis. Analysts have long recognized the dimensionality limitations of graphically based sensitivity analysis in portraying the robustness of a decision analysis to variations in underlying parameter estimates: If graphical methods allow at most 2- or 3-way sensitivity analyses, how can one be sure that a decision analysis is robust to the simultaneous variation of 10 to 20 parameters? Probabilistic sensitivity analysis was introduced to address this issue.[1,2] In a probabilistic sensitivity analysis, the analyst assigns distributions to uncertain parameters and can thereby compute as a measure of robustness the probability of a change in the optimal alternative due to variation in an arbitrary set of parameters, or alternately,[3,4] the expected value of perfect information regarding any set of parameters. This computation is most frequently done via Monte Carlo simulation.

The task of fitting distributions to uncertain parameters prior to a probabilistic sensitivity analysis has been approached in several standard ways. Traditionally, distributions of unobservable parameters (such as probabilities or rates) are fitted with a combination of mean and confidence interval estimated from data. Distributions typically used are the beta distribution[2,5,6] the logistic-normal distribution,[1,7-9] the uniform distribution, and the normal distribution.[10] Analysts also use bootstrap methods to obtain sampling distributions.[9,11] For observable parameters (such as costs), these methods are also applicable. However, in this case, it may be more convenient to simply fit a theoretical distribution to the empirical distribution of observations. For example, in Goodman and colleagues,[10] observations showed a peaked distribution, and a triangular distribution was used. Lord and Asante[11] used a piecewise linear approximation to the empirical distribution.

Received 2 September 2004 from IEMS Department, Northwestern University, Evanston, IL. Revision accepted for publication 13 December 2005.

Address correspondence to Gordon B. Hazen, IEMS Dept, McCormick School, Evanston, IL 60208-3119; phone: (847) 491-5673; e-mail: gbh305@lulu.it.northwestern.edu.

DOI: 10.1177/0272989X06290487

Data relevant for estimating parameters or parameter distributions typically come from one or more studies, so it is natural to attempt to use techniques of meta-analysis to combine evidence from relevant studies. However, none of the standard approaches to distribution fitting just mentioned can address issues—such as heterogeneity across studies and correlations between control and treatment parameters—that arise in a meta-analytic setting. Techniques of Bayesian meta-analysis[12-14] can address these issues and provide as well posterior distributions for parameters, exactly what is needed for a probabilistic sensitivity analysis. However, among complex cost-effectiveness models that draw on several studies to estimate parameters, we are aware of only a few[12,15] that use Bayesian meta-analytic posteriors in this way. No doubt the barrier to bridging these 2 areas is computational and theoretical complexity: Meta-analytic posterior distributions typically have no closed form and must be estimated via Markov-chain Monte Carlo (MCMC) methods, hardly worth the effort from the modeler's perspective.

Fortunately, for large study sizes, meta-analytic Bayesian posteriors do have closed forms not requiring MCMC estimation. The contribution of this article is to provide a summary of meta-analytic large-sample posterior distributions for probabilities, rates, and relative effect parameters that can easily be used to feed probabilistic sensitivity analyses. These techniques are simple enough to be carried out on a spreadsheet, but they still allow data from multiple studies to be combined responsibly, properly accounting for study heterogeneity as well as correlation between control and treatment parameters. As an example, we apply these methods to conduct a probabilistic sensitivity analysis for a recently published analysis by Mrus and Tsevat of zidovudine prophylaxis following rapid HIV testing in labor to prevent vertical HIV transmission in pregnant women.[16]

The methods we present for obtaining approximate normal posteriors for (possibly transformed) parameters (Table 1 through Table 4) are standard in the Bayesian literature and have been summarized in many texts.[17,18] One exception is our normal posterior results for heterogeneous controlled studies, where we have not been able to find literature precedents and have derived the results ourselves (see Appendix B). We have also derived the material presented here on estimating the cross-study covariance matrix and sample size requirements for logit and log transformation accuracy.

Ades, Lu, and Claxton[19] have drawn on like results for computing approximate expected values of *sample information*, a topic we do not address here. The methods we discuss overlap with Ades and colleagues in part, but differ in the use of random effects models for combining heterogeneous studies; where Ades and colleagues calculate a point estimate for the overall population mean, but we obtain an approximate normal posterior distribution. The latter more accurately reflects population-wide parameter variation. Our material also augments Ades and colleagues by addressing the issue of combining data from controlled and uncontrolled studies, and accounting for correlation between baseline and treatment parameters.

## OVERVIEW OF MODEL TYPES

The influence diagram Figure 1(a) shows the common situation in which observed data $y_i$ from each of $n$ studies $i = 1, \ldots, n$ are influenced by an unknown parameter $\xi$. Typically, $\xi$ is a probability or rate, and $y_i$ is a count of observed critical events in some subject population $i$. The observations $y_1, \ldots, y_n$ inform the choice of a decision or policy whose cost or utility for an individual or group is influenced by the not-yet-observed count $y$ of critical events for that patient or group. The not-yet-observed count $y$ is also influenced by the unknown $\xi$. Analysts can use the observations $y_1, \ldots, y_n$ to make statistical inferences about the unknown $\xi$ and can use these inferences to make predictions about the critical count $y$ and to make recommendations concerning the optimal decision or policy. We will be interested in Bayesian procedures in which the analyst calculates a posterior probability distribution for $\xi$ and uses it to form a predictive distribution for $y$, from which an expected-utility maximizing decision or policy can be computed.

Many studies compare a treatment with a control, and for these, the underlying parameter $\xi$ must be taken to be a vector $\xi = (\xi^0, \varepsilon)$, where $\xi^0$ is the unknown probability or rate for the control group and $\varepsilon$ is the unknown *efficacy*, some measure of the effectiveness of the treatment. The unknown probability or rate for the treatment group may be $\xi^1 = \xi^0 + \varepsilon$, or $\xi^1 = \xi^0 \cdot \varepsilon$, or some other function of $\xi^0$ and $\varepsilon$, depending on model specifics. In this situation, the observations $y_i$ are also vectors $y_i = (y_i^0, y_i^1)$, or perhaps $y_i = (y_i^0, y_i^1 - y_i^0)$, one component for control and one for treatment or treatment effect, and the decision or policy in question might be whether the treatment is cost-effective or beneficial in terms of expected utility.

Figure 1(a) assumes that the unknown parameter $\xi$ has the same value for each population $i$, as well as for the population from which future observations are to arise—that is, the populations are *homogenous*.

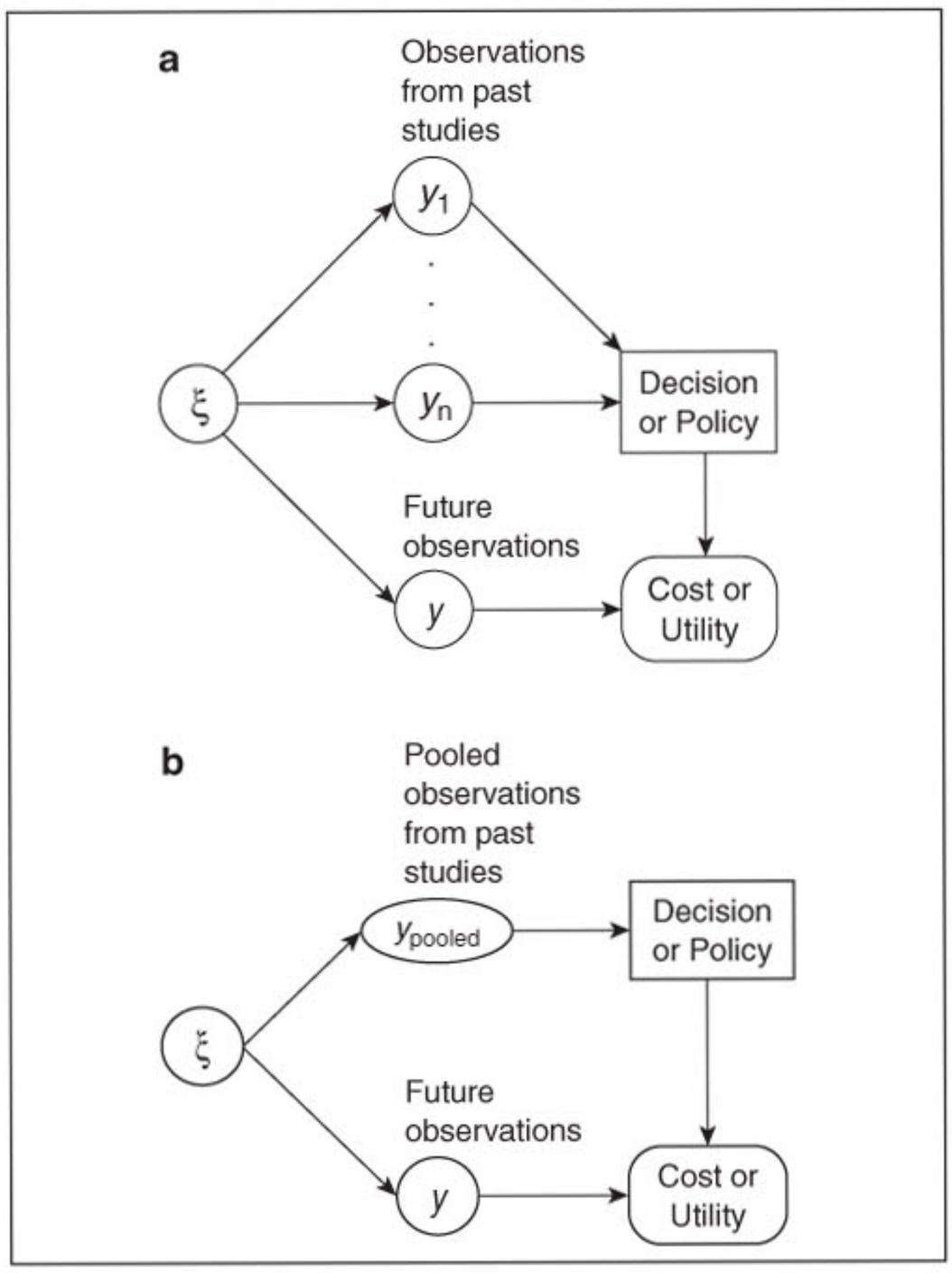


*Figure 1 (a) An influence diagram depicting the cost-effectiveness or decision-analytic setting when data from homogenous studies is available to estimate a parameter ξ. In a Bayesian approach, the posterior distribution on ξ given $y_1, \ldots, y_n$ can be used to determine the optimal decision or policy given the observations $y_1, \ldots, y_n$. In this situation, observations are typically pooled as if they come from a single study, as shown in (b).*

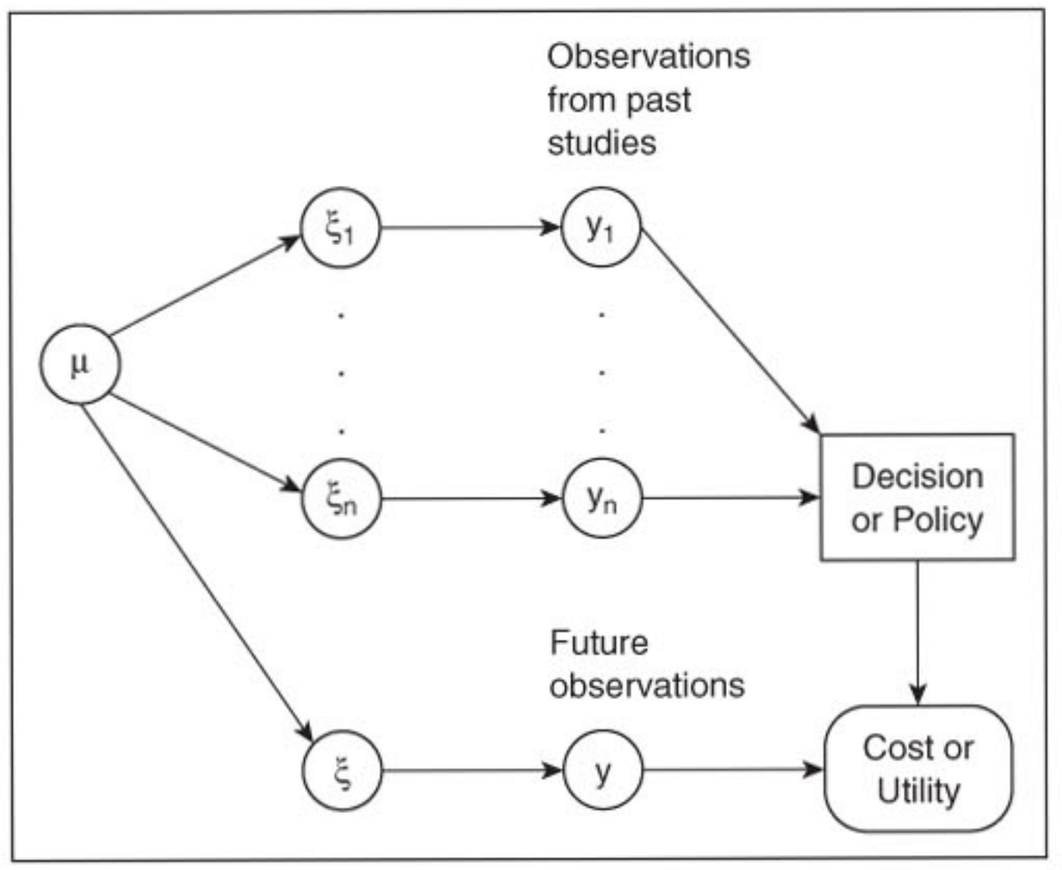


*Figure 2 An influence diagram depicting the cost-effectiveness or decision-analytic setting when data from heterogeneous studies is available to estimate a parameter ξ. Past studies aimed at estimating ξ are portrayed using a random effects model in which the corresponding parameter $\xi_i$ for the ith study is sampled from a population with mean μ. In a Bayesian approach, the posterior distribution on μ given $y_1, \ldots, y_n$ is used to infer a predictive distribution on ξ. This predictive distribution can be used to determine the optimal decision or policy given the observations $y_1, \ldots, y_n$.*

One may test this assumption statistically, or alternately, estimate the degree of heterogeneity, as we discuss below. If homogeneity is confirmed, it is common to proceed with analyses that effectively pool the data as though they are from the same source, as shown in Figure 1(b).

If homogeneity fails, then the populations are *heterogeneous*, with a different value $\xi_i$ of the unknown parameter for each population $i$ and the population of future observations. This situation is depicted in Figure 2, which shows a *random effects* model of heterogeneity, in which it is assumed that the values $\xi_i$ are independent draws from a population of $\xi$ values with mean $\mu$. Once again the observations $y_1, \ldots, y_n$ inform the choice of a decision or policy whose cost or utility is influenced by the not-yet-observed count $y$ of critical events for a patient or group. However, now $y$ is influenced by an unknown $\xi$ drawn from the same population of $\xi$ values with mean $\mu$. The analyst aims to use the observations $y_1, \ldots, y_n$ to make inferences about $\mu$ and in turn about $\xi$, from which predictions of the critical count $y$ can be made. A Bayesian analyst would seek a posterior distribution on $\mu$, from which could be inferred a predictive distribution on $\xi$. This in turn would allow the computation of a utility-maximizing decision or policy given the observations $y_1, \ldots, y_n$.

This situation also allows for controlled studies. Again the observations $y_i = (y_i^0, y_i^1)$ or perhaps $y_i = (y_i^0, y_i^1 - y_i^0)$ would be for control and treatment groups, the unknown $\xi_i = (\xi_i^0, \varepsilon_i)$ would consist of a component $\xi_i^0$ for the $i$th control group, and an efficacy parameter $\varepsilon_i$ for the $i$th group, and $\mu = (\mu^0, \mu^\varepsilon)$ would also have components for control and efficacy.

In the sequel, we examine in turn each combination of these 2 dichotomies (homogenous/heterogeneous, uncontrolled/controlled), summarizing how to obtain large-sample approximate Bayesian posterior distributions for $\xi$ (in the homogeneous case) or $\mu$ (in the heterogeneous case), and how to use these posterior distributions to conduct a probabilistic sensitivity

**Table 1.1** Homogeneous-Study Models Using Conjugate Prior Distributions

| Parameters | Observations | Prior Distributions | Posterior Distribution |
|---|---|---|---|
| Probability $p$ | $k$ events in $n$ independent trials; $k \sim$ binomial($n$, $p$) | $p \sim$ beta($\alpha$ ,$\beta$) | $p \sim$ beta($\alpha + k$, $\beta + n - k$) |
| | | Noninformative uniform prior on $p$ ($\alpha = 1$, $\beta = 1$) | $p \sim$ beta($k + 1$, $n - k + 1$) |
| Rate $\lambda$ | $k$ events in duration $\Delta t$; $k \sim$ Poisson($\lambda \Delta t$) | $\lambda \sim$ gamma($r$,$\theta$) | $\lambda \sim$ gamma($r + k$, $\theta + \Delta t$) |
| | | Noninformative prior on $\lambda$ ($r = 0$, $\theta = 0$) | $\lambda \sim$ gamma($k$, $\Delta t$) |
| Mean $\xi$ | $y \sim$ normal($\xi$, $\sigma^2$) $\sigma^2$ known or estimated | $\xi \sim$ normal($\xi_0$,$\sigma_0^2$) | $\xi \sim$ normal($\xi_1$, $\sigma_1^2$) $\xi_1 = \frac{\sigma_0^{-2}\xi_0 + \sigma^{-2}y}{\sigma_0^{-2} + \sigma^{-2}}$, $\sigma_1^{-2} = \sigma_0^{-2} + \sigma^{-2}$ |
| | | noninformative uniform prior on $\xi$ ($\sigma_0^{-2} = 0$) | $\xi \sim$ normal($y$, $\sigma^2$) |

analysis using Monte Carlo simulation. In each situation, we illustrate our results with data cited by Mrus and Tsevat[16] in their analysis of zidovudine prophylaxis following rapid HIV testing in labor. We conclude, as mentioned above, with a complete probabilistic sensitivity analysis of all probability and efficacy parameters in the Mrus and Tsevat model.

## OBSERVATIONS FROM HOMOGENEOUS STUDIES

If homogeneity holds, as in Figure 1, then the posterior distribution of the unknown parameter $\xi$ can be calculated exactly using standard conjugate Bayesian methods, for example, DeGroot.[20] These are summarized in Table 1.1. Exact posterior distributions may be obtained when $\xi$ is a probability $p$, a rate $\lambda$, or a mean. The beta, gamma, and normal distributions mentioned may be found in standard probability texts. The table also illustrates how posterior distributions simplify when the prior is noninformative, reflecting the paucity of prior information relative to the available data.

If sample size is large, then standard normal approximations to the binomial and Poisson distributions may be invoked. This in combination with the normal conjugate results in Table 1.1 gives the "no transformation" results in Table 1.2. Large sample size implies that $\sigma_0^{-2}$ is so small relative to $\sigma^{-2}$ that it may be effectively assumed zero, so the noninformative case from Table 1.1 applies.

A one-to-one transformation of probability or rate parameters may, however, be desirable to maintain logical correctness or consistency with statistical analyses. For instance, without transformation, a normal posterior distribution on a probability parameter $p$ is logically incorrect, as it allows an infinite range of values for a variable that is restricted to the interval [0,1]. If a normal posterior is used in a Monte Carlo simulation, randomly generated values for $p$ falling outside the interval [0,1] may lead to misleading results. Truncating or rounding the normal approximation to avoid this results in biased estimates.

These difficulties are negligible for normal posteriors tightly focused within [0,1], but they may be avoided altogether by transforming $p$ to a parameter $\xi$ having infinite or semi-infinite range. Common transformations for probability parameters $p$ are the log transformation $\xi = \ln p$, or the logit transformation $\xi = \text{logit}(p) = \ln (p/(1-p))$, and for rate parameters $\lambda$, the log transformation $\xi = \ln \lambda$. The observations $k$ can be transformed similarly to observations $y$, and the well-known Delta method (see Lemma 2 in Appendix D) then implies that for large samples, approximate normality is retained. The conjugate normal posterior from Table 1.1 can then be invoked, and this yields the transformed entries in Table 1.2.

The final column of Table 1.2 indicates how a random parameter value may be generated for Monte Carlo simulation: First generate a random $\xi$ from its approximate posterior normal($y$, $\sigma^2$) distribution, then transform $\xi$ back to obtain a random value of the

**Table 1.2** Large-Sample Approximate Homogeneous-Study Models for Probabilities and Rates

| Parameters | Observations | Transformation to an Approximate Normal Model (Table 1.1) | Generating New Parameter Values for MC Simulation: First Generate $\xi \sim \text{normal}(y, \sigma^2)$, Then: |
|---|---|---|---|
| Probability $p$ | $k$ events in $n$ independent trials<br>$k \sim \text{binomial}(n, p)$ | $\xi = p$<br>(no transformation)<br>$y = k/n$<br>$\sigma^2 = (k/n)(1 - k/n)/n$ | Set $p = \xi$. |
| | | $\xi = \text{logit}(p)$<br>$y = \text{logit}(k/n)$<br>$\sigma^2 = 1/k + 1/(n - k)$ | Calculate $p = \dfrac{e^{\xi}}{1 + e^{\xi}}$. |
| | | $\xi = \log(p)$<br>$y = \log(k/n)$<br>$\sigma^2 = 1/k - 1/n$ | Calculate $p = e^{\xi}$. |
| Rate $\lambda$ | $k$ events in duration $\Delta t$<br>$k \sim \text{Poisson}(\lambda \Delta t)$ | $\xi = \lambda$<br>(no transformation)<br>$y = k/\Delta t$<br>$\sigma^2 = k/\Delta t^2$ | Set $\lambda = \xi$. |
| | | $\xi = \ln(\lambda)$<br>$y = \ln(k/\Delta t)$<br>$\sigma^2 = 1/k$ | Calculate $\lambda = e^{\xi}$. |

MC = Monte Carlo.

original parameter $p$ or $\lambda$. For pooled observations from homogeneous studies, these large-sample approximate normal posteriors are unnecessary, as the conjugate beta or gamma posteriors from Table 1.1 are exactly correct and easily computed. However, the large-sample approximations generalize to more complicated situations that we will discuss below, where conjugate prior models are no longer tractable. We present them here to illustrate our general approach in a simple situation.

### Homogeneity of Past and Future Observations

Calculation of the posterior distributions above relies on the assumption of homogeneity of past studies. But one cannot use the posterior on $\xi$ as the prior on the parameter $\xi_f$ for future observations unless one has established some statistical relationship between $\xi$ and $\xi_f$. If we assume homogeneity of past studies and future observations (Figure 1), then we have $\xi_f = \xi$, and the posterior on $\xi$ applies to $\xi_f$ as well. However, if future observations do not sample the same population as past studies, then the random effects model we discuss below may be more realistic.

### Sample Size and Model Choice

The models in Table 1.2 and below require large samples, but how large is large? These models are built on 2 layers of approximations: First, the fact that large-sample binomial and Poisson distributions are approximately normal, and second, the fact that normality is maintained for large samples by the logit and log transformations (a result known as the Delta method—see Appendix D).

Consider first the binomial case. Widespread advice[21-23] indicates that a binomial distribution with parameters $n$ and $p$ may be accurately approximated by a normal distribution with the same mean and variance as long as that variance $\sigma^2 = np(1 - p)$ is at least 10, or alternately, as long as $np$ and $n(1 - p)$ are both at least 10. We are aware of one respected source[24] that claims that $np(1 - p)$ need only exceed 3. Thus, values of $np(1 - p)$ between 3 and 10 constitute a "gray area" where normal approximation begins to break down but may still be adequate.

For a Poisson distribution with parameter $\nu$ (equal to $\lambda \Delta t$ in Table 1.2), the normal approximation becomes adequate when the standard deviation $\sqrt{\nu}$ is small compared to the mean $\nu$.[25] As in the binomial case, $\nu \geq 10$ gives good approximations, and $3 \leq \nu \leq 10$ is a gray area.

The fact that logit and log transformations maintain normality for large samples depends on the fact that linear transformations of normally distributed variables are normally distributed and that any differentiable transformation is approximately linear over a small range. With $\hat{p} = k/n$ in Table 1.2, a rule

of thumb is that approximate normality is maintained for the logit transform as long as $n\hat{p}(1-\hat{p}) \geq 10$, if by approximate we mean cumulative probability errors of at most 0.05. This bound is a relaxation of the tighter approximate bound

$$n\hat{p}(1-\hat{p}) \geq 2\frac{\left(\hat{p}-\frac{1}{2}\right)^2}{\pi e^2 \delta^2}. \quad (1)$$

As long as (1) holds, the dominant term in the expression for error in cumulative probabilities will be at most $\delta$. Here $e = 2.718$ is the base of the natural logarithm. Inequality (1) is only an approximate guarantee based on second-order Taylor series expansions (see Appendix D). Nevertheless, this approximation seems to work well numerically.

For the log transform in either the binomial or Poison case, a rule of thumb is that one needs $k \geq 10$ to achieve error of at most 0.05 in the cumulative probabilities. In the binomial case, this is a relaxation of the tighter bound

$$k \geq \frac{1-\hat{p}}{2\pi e^2 \delta^2} \quad (2)$$

guaranteeing the dominant error term for cumulative probabilities is at most $\delta$. For the Poisson case, the tighter bound is

$$k \geq \frac{1}{2\pi e^2 \delta^2}.$$

Again, these are approximate bounds based on second-order Taylor series expansions. When these bounds fail for the logit or log transform, it may make sense to use the *untransformed* normal approximation in Table 1.2—see the examples immediately following. But of course one should check whether the resulting posterior on $p$ lies substantially within the interval [0,1], or whether the posterior on $\lambda$ is substantially in the nonnegative region—if not, then these posteriors will not be useful for Monte Carlo simulation.

### Example 1: Acceptance Rate for Rapid HIV Testing and Treatment

Mrus and Tsevat cite Rajegowda and colleagues,[26] who observed that in $n = 539$ patients, $k = 462$ (85.7%) were willing to accept rapid HIV test and treatment. To obtain a Bayesian posterior on the acceptance rate $p$, we may use the conjugate beta posterior in Table 1.1 or any of the large-sample approximate posteriors for binomial data in Table 1.2. The former yields a beta($k$ + 1, $n-k+1$) = beta(463, 78) posterior. For the latter, we note that for the estimate $\hat{p} = k/n = 0.857$, we have $n\hat{p}(1-\hat{p}) = 66$, substantially exceeding 10. An untransformed approximate normal posterior has mean $y = k/n = 0.857$ and standard deviation $\sigma = \sqrt{(k/n)(1-k/n)/n} = 0.015$.

If we consider using the logit transform, our rule of thumb $n\hat{p}(1-\hat{p}) \geq 10$ tells us immediately that we can easily maintain errors of 0.05 or less. The tighter bound (1) is satisfied by a maximum error of $\delta = 0.013$. So the logit maintains normality very accurately. Similarly, for the log transform, our rule of thumb $k \geq 10$ is easily satisfied, so accuracy is 0.05 or less. The tighter bound (2) is satisfied by maximum error of $\delta = 0.0026$. The beta posterior, the untransformed normal posterior, and the implied normal logit and lognormal posteriors on $p$ are graphed in Figure 3a and are virtually identical, reflecting the fidelity of the large-sample normal approximation expected with $n\hat{p}(1-\hat{p})$ substantially exceeding 10.

### Example 2: Specificity of Rapid HIV Testing

One of the 3 studies Mrus and Tsevat cite (see Example 5 below) on the specificity of rapid HIV testing revealed $k = 446$ negative test results among $n = 451$ subjects not HIV infected, a specificity of $\hat{p} = 446/451 = 0.989$. We have $n\hat{p}(1-\hat{p}) = 4.945$, so we are in the gray area with respect to the normal approximation to the binomial.

Consider using the logit transformation. Our rule of thumb $n\hat{p}(1-\hat{p}) \geq 10$ for 0.05 accuracy fails, and the tighter bound (1) is satisfied by maximum error $\delta = 0.065$. So the logit transformation may be poor. For the log transformation, on the other hand, our rule of thumb $k \geq 10$ easily holds, and the bound (2) is satisfied by maximum error $\delta = 0.00073$, so the log transform is very accurate.

These assertions are confirmed by the graphs in Figure 3b. The true maximum error in cumulative probabilities between the logit and the untransformed normal is 0.071 (versus the $\delta = 0.065$ estimate from (1)). We can also see that the normal approximation is only roughly adequate, as expected in that we are in the gray area. It has a roughly 1% chance of generating infeasible values $p > 1$ in a Monte Carlo simulation, a not unacceptable figure.

### Example 3: Mother-to-Infant HIV Transmission Risk

One of the 7 studies Mrus and Tsevat cite (see Example 4 below) regarding mother-to-infant HIV

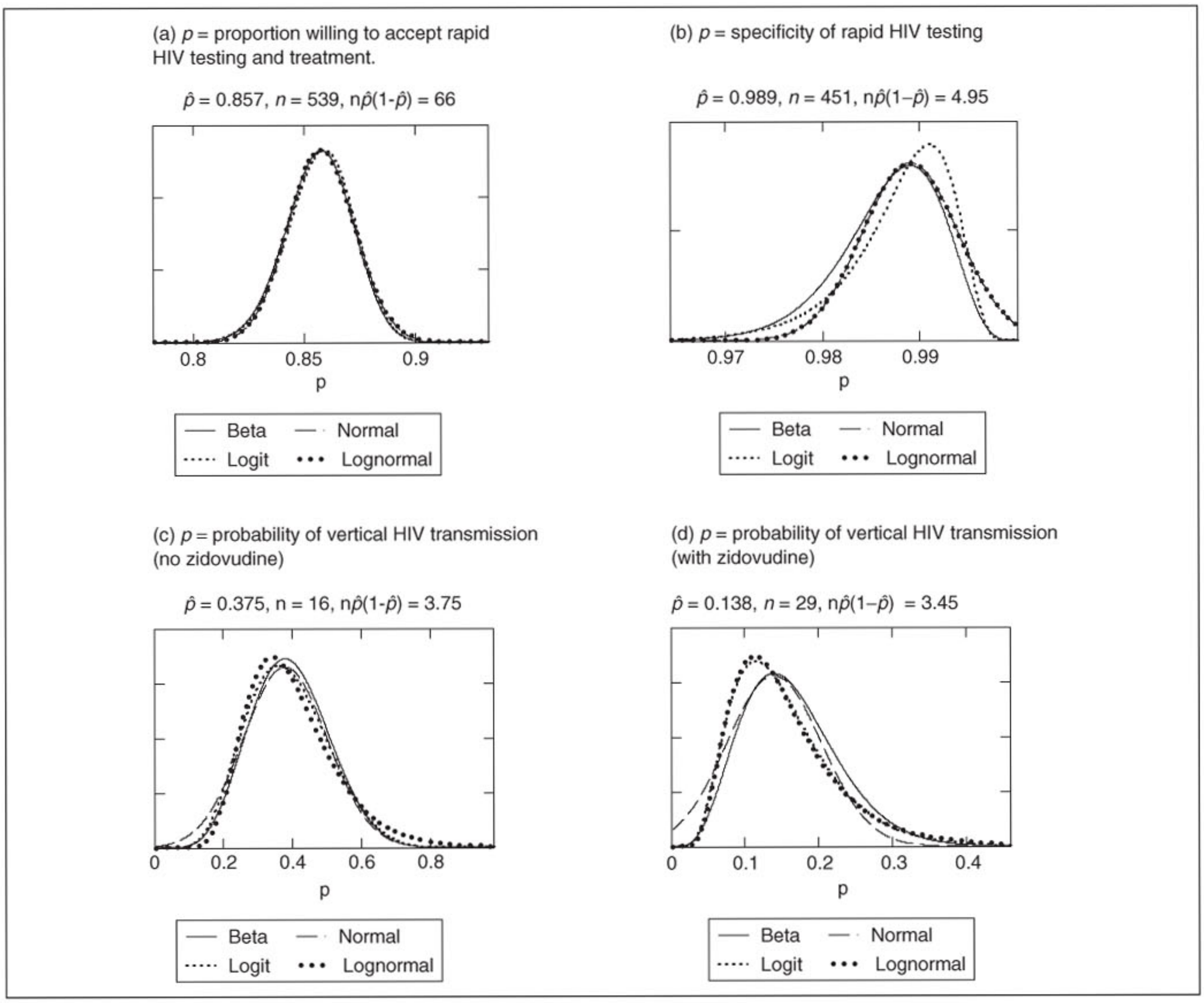


*Figure 3 Beta posterior distribution for a probability p of a critical event, along with normal, normal log-odds, and lognormal approximate posterior distributions. In (a), $n\hat{p}(1-\hat{p})$ is much greater than 10 and the approximate normal, normal log-odds and lognormal posteriors are so close to the true beta posterior that it is difficult to distinguish them. In (b), (c), and (d), $n\hat{p}(1-\hat{p})$ lies in a "gray area" below 10 where approximate normal posteriors begin to lose their accuracy.*

transmission risk gives 6 instances in 16 of mother-to-infant HIV transmission without zidovudine and 4 instances in 29 with zidovudine. We have $\hat{p}_0 = 6/16 = 0.375$, $n\hat{p}_0(1-\hat{p}_0) = 3.75$, and also $\hat{p}_1 = 4/29 = 0.138$, $n\hat{p}_1(1-\hat{p}_1) = 3.45$, so with respect to the normal approximations to the binomial from Table 1.2, we are in the gray area for both samples.

Consider using the logit transform. Our rule of thumb $n\hat{p}(1-\hat{p}) \geq 10$ for 0.05 accuracy fails for both samples, but the bound (1) yields achievable accuracy $\delta = 0.019$ for the no zidovudine sample and $\delta = 0.057$ for the zidovudine sample. So the logit transform is accurate for the first sample and marginally accurate for the second. Regarding the log transform, our rule of thumb $k \geq 10$ for 0.05 accuracy fails for both samples, but the tighter bound (2) is satisfied by $\delta = 0.047$ for the no zidovudine sample and $\delta = 0.068$ for the zidovudine sample. So the log transform is not as accurate as the logit for the first sample and is comparable to the logit for the second sample. These assertions are confirmed by Figure 3c and d, where the true maximum errors in cumulative probabilities between

**Table 2.1** Normally Distributed Observations from a Controlled Study

| Parameters | Observations | Prior Distributions | Posterior Distribution | Generating New Parameter Values $\xi_0$, $\xi_1$ for Monte Carlo Simulation |
|---|---|---|---|---|
| $\xi_0$, $\xi_1$<br>$\varepsilon = \xi_1 - \xi_0$ | $y_0 \sim \text{normal}(\xi_0, \sigma_0^2)$<br>$y_1 \sim \text{normal}(\xi_1, \sigma_1^2)$<br>$\sigma_0^2$, $\sigma_1^2$ known or estimated from data | $\xi_0$, $\varepsilon$ are independent, each with a noninformative prior | $\xi_0 \sim \text{normal}(y_0, \sigma_0^2)$<br>$\xi_1 \sim \text{normal}(y_1, \sigma_1^2)$<br>$\xi_0$, $\xi_1$ independent | Generate<br>$\xi_0 \sim \text{normal}(y_0, \sigma_0^2)$<br>$\xi_1 \sim \text{normal}(y_1, \sigma_1^2)$ |

the logit and the untransformed normal are 0.029 and 0.068 (v. the estimates $\delta = 0.019$ and 0.057 from (1)), and between the log and the untransformed normal are 0.051 and 0.075 (v. the estimates $\delta = 0.047$ and 0.068). The untransformed normal approximation has for the no zidvudine sample about a 1/1000 chance of generating an infeasible negative random variate $p$ in a Monte Carlo simulation, and for the zidovudine sample, the chance is 1.6%.

## OBSERVATIONS FROM HOMOGENEOUS CONTROLLED STUDIES

We consider now the case of observations from homogeneous controlled studies that have been pooled. Table 2.1 shows the situation in which the observations $y_0, y_1$ from the control and treatment groups are normally distributed and the prior distribution is noninformative. Here the parameter vector $(\xi_0, \varepsilon)$ consists of the unknown mean $\xi_0$ in the control group and the unknown efficacy $\varepsilon$. The mean in the treatment group is then $\xi_1 = \xi_0 + \varepsilon$. Table 2.1 gives the posterior distribution of $(\xi_0, \xi_1)$ under a noninformative prior when all likelihoods are normal. This result is a consequence of Theorem 2 in Appendix B. Generating parameter values $\xi_0$ and $\xi_1$ for Monte Carlo simulation is done in the obvious way, as summarized in the table.

It may seem strange that $\xi_0, \xi_1$ are a posteriori independent even though they are dependent a priori (because $\xi_1 = \xi_0 + \varepsilon$). The reason for this is that the noninformative priors on $\xi_0$ and $\varepsilon$ are too weak to maintain this dependence in the face of independent data $y_0, y_1$.

For the situation in which control group and treatment effect are independent parameters (as assumed in Table 2.1), studies involving binomial or Poisson observations do not have tractable conjugate Bayesian updates. However, when large-sample normal approximations are valid, then results from Table 2.1 may be applied. Table 2.2 gives the resulting approximate normal posterior distributions for the binomial case, and Table 2.3 for the Poisson case. To generate random variates for a probabilistic sensitivity analysis, one should use the procedure from Table 2.1 and then transform back as indicated in Table 2.2 or Table 2.3 to obtain the desired probability or rate random variates.

### Example 4: The Effect of Zidovudine Prophylaxis on HIV Transmission

Mrus and Tsevat cite 7 controlled studies[27-33] giving data for estimating the effect of zidovudine prophylaxis on mother-to-infant HIV transmission risk. Observed data are listed as follows,

| | Without Zidovudine | | | With Zidovudine | | |
|---|---|---|---|---|---|---|
| Group i | $k_i^0$ | $n_i^0$ | $k_i^0/n_i^0$ | $k_i^1$ | $n_i^1$ | $k_i^1/n_i^1$ |
| 1 | 47 | 152 | 0.309 | 19 | 416 | 0.0457 |
| 2 | 51 | 216 | 0.236 | 16 | 204 | 0.0784 |
| 3 | 30 | 95 | 0.316 | 21 | 336 | 0.0625 |
| 4 | 6 | 16 | 0.375 | 4 | 29 | 0.1379 |
| 5 | 30 | 115 | 0.261 | 19 | 115 | 0.1652 |
| 6 | 37 | 198 | 0.187 | 18 | 194 | 0.0928 |
| 7 | 1019 | 5571 | 0.18291 | 223 | 2269 | 0.09828 |

where $k_i^0$ of $n_i^0$ infants in the control group $i$ are HIV infected, and $k_i^1$ of $n_i^1$ infants in the prophylaxis group $i$ are HIV infected, $i = 1,2, \ldots 7$.

Assuming homogeneity tests are passed (we will discuss this further below), we can pool the 7 study groups into one as follows.

| Without Zidovudine | | With Zidovudine | |
|---|---|---|---|
| $k_0$ | 1220 | $k_1$ | 320 |
| $n_0$ | 6363 | $n_1$ | 3563 |
| $k_0/n_0$ | 0.191733 | $k_1/n_1$ | 0.089812 |
| $y_0 = \ln(k_0/n_0)$ | −1.65165 | $y_1 = \ln(k_1/n_1)$ | −2.41004 |
| $\sigma_0^2 = 1/k_0 - 1/n_0$ | 0.000663 | $\sigma_1^2 = 1/k_1 - 1/n_1$ | 0.002844 |

**Table 2.2** Controlled Studies Involving Binomial Observations and Their Large-Sample Normal Approximation

| Parameters | Observations | Transformation to an Approximate Normal Model (Table 2.1) | Generating a New Parameter Value for MC Simulation: Generate $\xi_0$, $\xi_1$ as in Table 2.1. Then |
|---|---|---|---|
| Probabilities $p_0$, $p_1$ | $k_0$ events in $n_0$ independent trials without intervention<br>$k_1$ events in $n_1$ independent trials with intervention<br>$k_0 \sim$ binomial$(n_0, p_0)$<br>$k_1 \sim$ binomial$(n_1, p_1)$ | $\xi_0 = p_0$<br>$\xi_1 = p_1$<br>(no transformation)<br>$y_0 = k_0/n_0$<br>$y_1 = k_1/n_1$<br>$\sigma_0^2 = (k_0/n_0)(1-k_0/n_0)/n_0$<br>$\sigma_1^2 = (k_1/n_1)(1-k_1/n_1)/n_1$ | Set $p_0 = \xi_0$, $p_1 = \xi_1$ |
| | | $\xi_0 = \text{logit}(p_0)$<br>$\xi_1 = \text{logit}(p_1)$<br>$y_0 = \text{logit}(k_0/n_0)$<br>$y_1 = \text{logit}(k_1/n_1)$<br>$\sigma_0^2 = 1/k_0 + 1/(n_0 - k_0)$<br>$\sigma_1^2 = 1/k_1 + 1/(n_1 - k_1)$ | Calculate<br>$p_0 = \dfrac{e^{\xi_0}}{1+e^{\xi_0}}$<br>$p_1 = \dfrac{e^{\xi_1}}{1+e^{\xi_1}}$. |
| | | $\xi_0 = \ln(p_0)$<br>$\xi_1 = \ln(p_1)$<br>$y_0 = \ln(k_0/n_0)$<br>$y_1 = \ln(k_1/n_1)$<br>$\sigma_0^2 = 1/k_0 - 1/n_0$<br>$\sigma_1^2 = 1/k_1 - 1/n_1$ | Calculate<br>$p_0 = e^{\xi_0}$<br>$p_1 = e^{\xi_1}$<br>$RR = p_1/p_0 = e^{\xi}$ |

MC = Monte Carlo.

According to Table 2.1, the transformed log-risk parameter $\xi_0 = \ln p_0$ has a posterior approximate normal distribution with mean –1.65 and standard deviation $0.000663^{1/2} = 0.0257$, and $\xi_1$ is approximately normal with mean –2.41 and standard deviation $\sigma_1 = 0.002844^{1/2} = 0.0533$.

## OBSERVATIONS FROM HETEROGENEOUS STUDIES

Consider now the random effects situation depicted in Figure 2, in which observations relevant to a parameter $\xi$ arise from $n$ heterogeneous studies randomly drawn from an overall population with mean $\mu$. Let $\sigma_\xi^2$ be the variance of this overall population. If $\sigma_\xi^2$ is positive, then the studies are heterogeneous, whereas if $\sigma_\xi^2 = 0$, then all $\xi_i$ are equal and the homogenous case of Figure 1 obtains. A popular test[34] of the null hypothesis $\sigma_\xi^2 = 0$ uses the statistic $Q$ given by

$$Q = \sum_i \omega_i^{null} (y_i - \bar{y}^{null})^2 \qquad (3)$$

$$\omega_i^{null} = \frac{1}{\sigma_i^2}$$

$$\bar{y}^{null} = \frac{\sum_i \omega_i^{null} y_i}{\sum_i \omega_i^{null}}.$$

where $\sigma_i$ is the known or estimated standard deviation from the $i^{th}$ study. $Q$ has a large-sample approximate $\chi^2$ distribution with $n - 1$ degrees of freedom under the null hypothesis, so one would reject this hypothesis if $Q > \chi_\alpha^2$ (n – 1) for some appropriate test level $\alpha$. However, it is known that this test has low power, so failure to reject may not be adequate reason for assuming $\sigma_\xi^2 = 0$ (see the example below), and anyway the large sample requirement is usually not met in this context, as often only a few studies are available. In fact, one can argue, with Dumouchel,[13]p. 511 that this

**Table 2.3** Controlled Studies Involving Poisson Observations and Their Large-Sample Normal Approximation

| Parameters | Observations | Transformation to an Approximate Normal Model (Table 2.1) | Generating a New Parameter Value for MC Simulation: Generate $\xi_0$, $\xi_1$ as in Table 2.1. Then |
|---|---|---|---|
| Rates $\lambda_0$, $\lambda_1$ | $k_0$ events in duration $\Delta t_0$ without intervention<br>$k_1$ events in duration $\Delta t_1$ with intervention<br>$k_0 \sim \text{Poisson}(\lambda_0 \Delta t_0)$<br>$k_1 \sim \text{Poisson}(\lambda_1 \Delta t_1)$ | $\xi_0 = \lambda_0$<br>$\xi_1 = \lambda_1$<br>(no transformation)<br>$y^0 = k_0/\Delta t_0$<br>$y^1 = k_1/\Delta t_1$<br>$\sigma_0^2 = k_0/\Delta t_0^2$<br>$\sigma_1^2 = k_1/\Delta t_1^2$ | Set $\lambda_0 = \xi_0$, $\lambda_1 = \xi_1$ |
| | | $\xi_0 = \ln(\lambda_0)$<br>$\xi_1 = \ln(\lambda_1)$<br>$y_0 = \ln(k_0/\Delta t_0)$<br>$y_1 = \ln(k_1/\Delta t_1)$<br>$\sigma_0^2 = 1/k_0$<br>$\sigma_1^2 = 1/k_1$ | Calculate<br>$\lambda_0 = e^{\xi_0}$<br>$\lambda_1 = e^{\xi_1}$<br><br>$RR = \lambda_1/\lambda_0 = e^{\xi}$ |

MC = Monte Carlo.

method "ignores the distinction between failing to reject interstudy homogeneity and proving it is present." Some study heterogeneity is inevitable, even if undetected by a hypothesis test, so why not simply estimate it? It can be shown that any estimate of the form

$$\hat{\sigma}_\xi^2 = \frac{\sum_i \alpha_i (y_i - \bar{y})^2 - \sum_i \alpha_i (1-\alpha_i)\sigma_i^2}{\sum_i \alpha_i (1-\alpha_i)} \qquad \bar{y} = \sum_i \alpha_i y_i$$

is unbiased as long as $\sum_i \alpha_i = 1$. Choosing $\alpha_i = \sigma_i^{-2} / \sum_i \sigma_i^{-2}$ yields the moment estimator[34]

$$\hat{\sigma}_\xi^2 = \frac{Q - (n-1)}{\sum_i \omega_i^{null} - \left(\sum_i (\omega_i^{null})^2 / \sum_i \omega_i^{null}\right)}. \qquad (4)$$

Unfortunately, for a small number of studies, these estimates are highly variable and can even be negative. A sampling distribution would therefore be desirable, but no simple one is available except in the equal-variance case $\sigma_1^2 = \ldots = \sigma_n^2 = \sigma^2$, when $\alpha_i = 1/n$ and

$$\hat{\sigma}_\xi^2 = \frac{1}{n-1}\sum_i (y_i - \bar{y})^2 - \sigma^2. \qquad \bar{y} = \sum_i y_i / n$$

in which case $\hat{\sigma}_\xi^2$ is equal in distribution to $\sigma^2 + (\sigma^2 + \sigma_\xi^2)\chi_{n-1}^2/(n-1)$, which has upper-$\alpha$ percentile

$$\hat{\sigma}_\xi^2(\alpha) = (\sigma^2 + \sigma_\xi^2)\frac{\chi_{n-1}^2(\alpha)}{n-1} - \sigma^2 \qquad (5)$$

Following Larsen and colleagues,[35] Spiegelhalter and colleagues[14] point out that because $\xi$ is normal($\mu$, $\sigma_\xi^2$), the 95% range of sampled $\xi$ values is $3.92\sigma_\xi$, and median absolute difference between 2 randomly sampled values of $\xi$ is $1.09\sigma_\xi$ ; these facts can be used to aid in a rough subjective estimate of $\sigma_\xi$ when (4) is inadequate due to a small number of studies (see Example 5 below). This includes the case of only a single study, where (4) is not even defined.

When there is only one past study, one must be particularly careful not to unthinkingly fall back on the homogeneous model (Figure 1 and Table 1.1), which assumes $\sigma_\xi^2 = 0$, unless future observations of interest arise only from the same population sampled by the study. The proper procedure is to subjectively estimate $\sigma_\xi^2$, admittedly prone to error but certainly more accurate than taking $\sigma_\xi^2 = 0$. Concern about the reliability of subjective estimates should be addressed by a sensitivity analysis.

Table 3.1 gives the posterior distribution for $\mu$ for heterogeneous studies when all likelihoods are normal (see Theorem 2 in Appendix B). The normal $(\bar{y}, (\sum_i \omega_i)^{-1})$ posterior on $\mu$ combines with the normal($\mu$, $\sigma_\xi^2$) prior on $\xi$ given $\mu$ to produce a normal $(\bar{y}, \sigma_\xi^2 + (\sum_i \omega_i)^{-1})$ posterior on $\xi$, and this may be used to generate values for a Monte Carlo probabilistic sensitivity analysis, as indicated. Table 3.2 gives large-sample approximate posterior distributions for transformed probabilities and rates under binomial

**Table 3.1** Normally Distributed Observations from Heterogeneous Groups (Random Effects Model)

| Parameter for Study *i* | Observations $y_i$ for Study *i* | Prior Distributions | Posterior Distribution | Generating a New Parameter Value $\xi$ for Monte Carlo Simulation |
|---|---|---|---|---|
| $\xi_i$ | $y_i \sim \text{normal}(\xi_i, \sigma_i^2)$ | $\mu$ has a noninformative prior | $\mu \sim \text{normal}(\bar{y}, (\sum_i \omega_i)^{-1})$ | Generate $\xi \sim$ normal $(\bar{y}, \sigma_\xi^2 + (\sum_i \omega_i)^{-1})$ |
| | $\sigma_i^2$ known or estimated | $\xi_i \mid \mu \sim \text{normal}(\mu, \sigma_\xi^2)$ | where $\omega_i = (\sigma_i^2 + \sigma_\xi^2)^{-1}$ | |
| | | $\sigma_\xi^2$ known or estimated | $\bar{y} = (\sum_i \omega_i)^{-1} \sum \omega_i y_i$ | |

**Table 3.2** Binomial and Poisson Observations from Heterogeneous Groups and Their Large-Sample Normal Approximation

| Parameter for Study *i* | Observations for Study *i* | Transformation to an Approximate Normal Random Effects Model (Table 3.1) | Generating New Parameter Value for MC Simulation: Generate $\xi$ as in Table 3.1. Then |
|---|---|---|---|
| Probability $p_i$ | $k_i$ events in $n_i$ independent trials.<br>$k_i \mid p_i \sim \text{binomial}(n_i, p_i)$ | $\xi_i = p_i$<br>(no transformation)<br>$y_i = k_i/n_i$<br>$\sigma_i^2 = (k_i/n_i)(1-k_i/n_i)/n_i$ | Set $p = \xi$ |
| | | $\xi_i = \text{logit}(p_i)$<br>$y_i = \text{logit}(k_i/n_i)$<br>$\sigma_i^2 = 1/k_i + 1/(n_i - k_i)$ | Calculate $p = \dfrac{e^\xi}{1+e^\xi}$. |
| | | $\xi_i = \ln(p_i)$<br>$y_i = \ln(k_i/n_i)$<br>$\sigma_i^2 = 1/k_i - 1/n_i$ | Calculate $p = e^\xi$ |
| Rate $\lambda_i$ | $k_i$ events in duration $\Delta t_i$<br>$k_i \mid \lambda_i \sim \text{Poisson}(\lambda_i \Delta t_i)$ | $\xi_i = \lambda_i$<br>(no transformation)<br>$y_i = k_i/\Delta t_i$<br>$\sigma_i^2 = k_i/\Delta t_i^2$ | Set $\lambda = \xi$ |
| | | $\xi_i = \ln(\lambda_i)$,<br>$y_i = \ln(k_i/\Delta t_i)$<br>$\sigma_i^2 = 1/k_i$ | Calculate $\lambda = e^\xi$ |

MC = Monte Carlo.

and Poisson sampling, drawing on the results in Table 3.1.

### Example 5: Specificity of Rapid HIV Testing

Mrus and Tsevat cite 3 studies[36-38] providing data on the specificity of rapid HIV testing, observing that in $n_i$ people who were not HIV-infected, there were $k_i$ having negative test results, $i = 1,2,3$. The data are as follows.

| Group *i* | $k_i$ | $n_i$ | $\hat{p}_i = k_i/n_i$ | $\sigma_i = \sqrt{\hat{p}_i(1-\hat{p}_i)/n_i}$ |
|---|---|---|---|---|
| 1 | 446 | 451 | 0.988914 | 0.00493 |
| 2 | 783 | 790 | 0.991139 | 0.00333 |
| 3 | 486 | 496 | 0.979839 | 0.00631 |

We use Table 3.1 and Table 3.2 without transformation, as the variances $n\hat{p}_i(1-\hat{p}_i)$ do not exceed 10 (see Example 2 above), and the between-study variance $\sigma_\xi^2$ is not large. In fact, the test statistic (3) for the null hypothesis $\sigma_\xi^2 = 0$ of homogeneity yields $Q = \sum_{i=1}^{3} \omega_i^{\text{null}} (y_i - \bar{y}^{\text{null}})^2 = 2.508$, with $P$-value = 0.285, insufficient to reject $\sigma_\xi^2 = 0$. However, this test has low power, and anyway, we consider it a better strategy to estimate heterogeneity rather than test for it. The moment estimator (4) yields $\hat{\sigma}_\xi = 0.00238$. The upper-0.10 percentile (5) for the equal-variance estimate $\hat{\sigma}_\xi = 4.078 \times 10^{-3}$ is $6.172 \times 10^{-3}$, taking $\sigma^2$ in (5) to be the $\alpha_i$-weighted average over $i$ of the $\sigma_i^2$ values given here. So it appears that $\sigma_\xi$ could be $6.172/4.078 \approx 1.5$ times as large as its estimate $\hat{\sigma}_\xi = 0.00238$, that is, $\sigma_\xi$ could

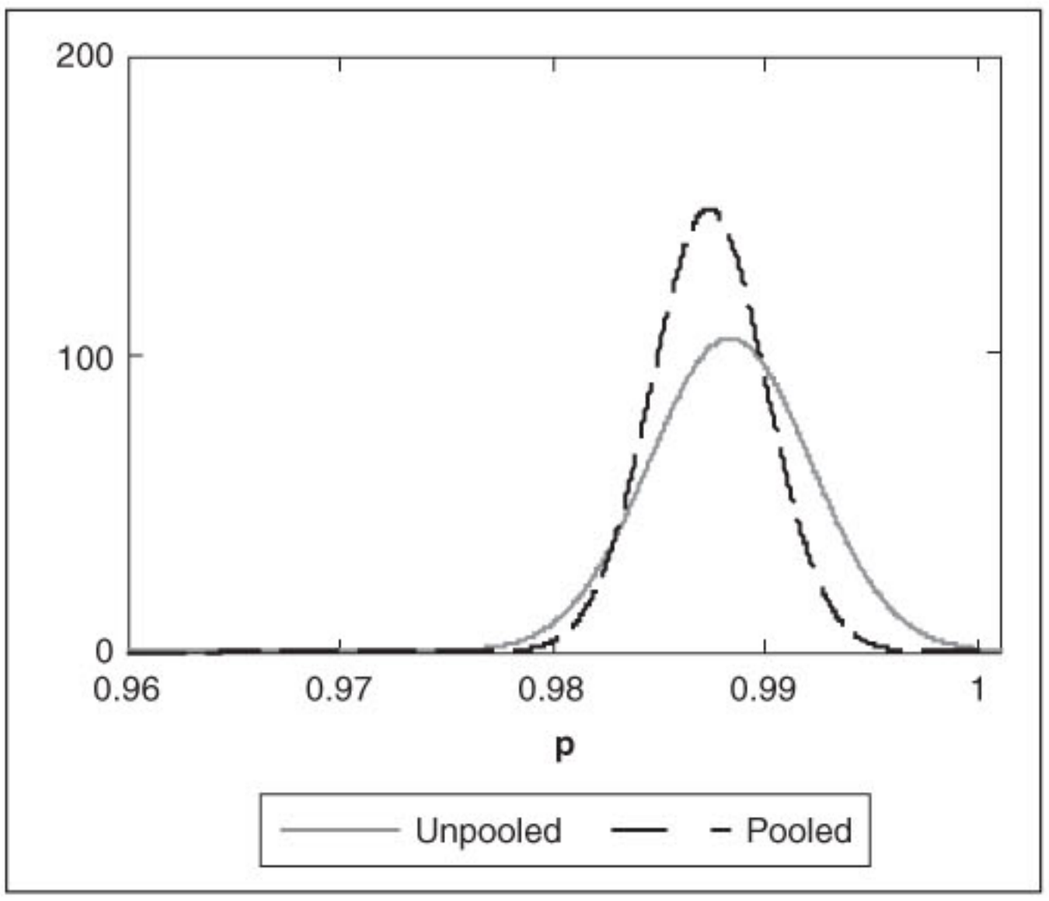


*Figure 4 Normal posterior distributions on the specificity p of rapid HIV test, one based on pooling 3 studies under the assumption they are homogeneous, and the other unpooled result based on combining studies using the random effects model. The random effects model gives a wider posterior distribution, reflecting residual uncertainty regarding whether the studies are really homogeneous.*

be as much as $1.5 \times 0.00238 \approx 0.00361$. In this light, $\hat{\sigma}_\xi = 0.00238$ seems not an unreasonable estimate.

Using this estimate, we obtain $(\omega_1,\omega_2,\omega_3) = (3.335, 5.954, 2.197) \times 10^4$, $(\sum_i \omega_i)^{-1} = 8.706 \times 10^{-6}$, $\bar{y} = 0.988$. Hence the approximate posterior distribution of parameter $\mu$ is normal with mean 0.988 and standard deviation $(8.706 \times 10^{-6})^{1/2} = 0.00295$.

To generate a random $\xi = p$, we sample from a normal distribution with mean 0.988 and standard deviation $(8.706 \times 10^{-6} + \sigma_\xi^2)^{1/2} = 0.00379$. This has a 95% range 0.981–0.996 and a 0.10% chance of generating an infeasible $p > 1$, certainly acceptable.

Had we accepted homogeneity as indicated by the $Q$ statistic, we would have pooled the 3 studies to obtain $k = 1715$ negative test results in $n = 1737$ HIV-negative individuals. This yields an approximate predictive distribution on $p$ that is normal with mean 0.987 and standard deviation 0.00268. The mean is virtually the same as the unpooled case, but the standard deviation is only 2/3 as large.

The pooled and unpooled posterior distributions for the specificity $p$ are graphed in Figure 4. Even though homogeneity is not rejected and pooling might therefore be defended by some, the pooled posterior is narrower than the unpooled posterior. Were the homogeneity hypothesis correct and $\sigma_\xi^2$ really equal to zero, the 2 posteriors would be identical[a]—here the wider posterior for the unpooled case reflects residual uncertainty about homogeneity. So in this light (and in general, we believe), the decision to retain heterogeneity as a possibility seems prudent.

### Example 6: Sensitivity of Rapid HIV Testing

Mrus and Tsevat cite 3 studies[36-38] providing data on rapid HIV test sensitivity. These studies observe $k = 262$ positive test results out of $n = 262$ HIV-infected individuals (100% sample sensitivity). The data, which we pool, provides no information on study heterogeneity, which must be subjectively estimated. Ideally the estimate $\sigma_\xi^2$ should be informed by some knowledge of how or why sensitivity might vary across studies. In the absence of such information here, we take an estimated range 0.95–0.999 on study mean sensitivity. This translates into a range of logit(0.999)–logit(0.95) = 3.962 on logit sensitivity, yielding a standard deviation $\sigma_\xi = 3.962/3.92 = 1.011$ in the logit model. Applying Table 3.1 to the logit portion of Table 3.2 with half-integer corrections $k = 262.5$, $n = 263$ yields a normal posterior on $\mu$ with mean 6.26 and standard deviation 1.74. The predictive distribution on $\xi = \text{logit}(p)$ is normal with the same mean and standard deviation $(\sigma_\xi^2 + 1.74^2)^{1/2} = 2.012$. This translates into a median sensitivity $p = 0.998$ and 95% range 0.911–0.99996.

### Example 7 (continuation of Example 1): Acceptance Rates for Rapid HIV Testing and Treatment

We have mentioned the single Rajegowda study,[39] observing that $k = 462$ of $n = 539$ patients (85.7%) were willing to accept rapid HIV test and treatment. To generalize beyond the population from which this study sampled, one must estimate the cross-study variance $\sigma_\xi^2$. As there is only one study, there is no data to inform this estimate. In the absence of any information on acceptance rates, we estimate a range of 0.70–0.95 on study mean sensitivity. This translates into a range of logit(0.95)–logit(0.70) = 2.097 on logit acceptance rate, yielding a standard deviation $\sigma_\xi = 2.097/3.92 = 0.535$ in the logit model. Applying Table 3.1 to the logit portion of Table 3.2 yields a normal posterior on $\mu$ with mean 1.792 and standard deviation 0.549. The predictive distribution

[a]The statement remains true when a transformation has been applied if the transformed data (not the original data) are pooled.

**Table 4.1** Normally Distributed Observations from Heterogeneous Controlled Studies (Random Effects Model)

| Parameters for Study $i$ | Observations $y_i$ for Study $i$ | Prior Distributions | Posterior Distribution | Generating New Parameter Values $\xi^0, \xi^1$ for Monte Carlo Simulation |
|---|---|---|---|---|
| $\xi_i^0$, $\xi_i^1$<br>$\varepsilon_i = \xi_i^1 - \xi_i^0$<br>$\xi_i = (\xi_i^0, \varepsilon_i)$ | $y_i^0 \sim \text{normal}(\xi_i^0, \sigma_{i0}^2)$<br>$y_i^1 \sim \text{normal}(\xi_i^1, \sigma_{i1}^2)$<br>$\sigma_{i0}^2$, $\sigma_{i1}^2$ known or estimated | $\mu = (\mu^0, \mu^\varepsilon)$ has a noninformative prior.<br>$(\xi_i^0, \varepsilon_i)$ bivariate normal with mean $\mu = (\mu^0, \mu^\varepsilon)$, and covariance matrix<br>$\Sigma_{\xi 0} = \begin{pmatrix} \sigma_{\xi^0}^2 & \sigma_{\xi^0,\varepsilon} \\ \sigma_{\xi^0,\varepsilon} & \sigma_\varepsilon^2 \end{pmatrix}$.<br>$\sigma_{\xi 0}^2$, $\sigma_\varepsilon^2$, $\sigma_{\xi 0,\varepsilon}$ known or estimated. | $\mu = (\mu^0, \mu^\varepsilon) \sim \text{normal}(\bar{y}, (\Sigma_i \Omega_i)^{-1})$<br>where $\Omega_i = (\Sigma_i + \Sigma_0^\varepsilon)^{-1}$<br>$\Sigma_i = \begin{pmatrix} \sigma_{i0}^2 & -\sigma_{i0}^2 \\ -\sigma_{i0}^2 & \sigma_{i0}^2 + \sigma_{i1}^2 \end{pmatrix}$<br>$\bar{y} = \left(\sum_i \Omega_i\right)^{-1} \left(\sum_i \Omega_i y_i\right)$,<br>$y_i = \begin{pmatrix} y_i^0 \\ y_i^1 - y_i^0 \end{pmatrix}$ | Step 1. Generate $(\xi^0, \varepsilon) \sim \text{normal}(\bar{y}, \Sigma_{\xi 0} + (\Sigma_i \Omega_i)^{-1})$.<br>Step 2. Calculate $\xi^1 = \xi^0 + \varepsilon$. |

on $\xi = \text{logit}(p)$ is normal with the same mean, and standard deviation $(\sigma_\xi^2 + 0.549^2)^{1/2} = 0.767$. This translates into a 95% range 0.572–0.964 on future observed acceptance rates.

## OBSERVATIONS FROM HETEROGENEOUS CONTROLLED STUDIES

For heterogeneous controlled studies, each element in Figure 2 is a 2-dimensional vector consisting of baseline and efficacy components: $\mu = (\mu^0, \mu^\varepsilon)$, $\xi_i = (\xi_i^0, \varepsilon_i)$. If $y_i^0, y_i^1$ are the (possibly transformed) control and treatment observations in study $i$, we adopt, for consistency, the notation $y_i = (y_i^0, y_i^1 - y_i^0)$, which records the control observation and the treatment effect $y_i^1 - y_i^0$. The parameter $\xi_i$ is assumed to have a *bivariate* normal distribution with mean $\mu$ and $2 \times 2$ covariance matrix $\Sigma_{\xi 0} = \begin{pmatrix} \sigma_{\xi^0}^2 & \sigma_{\varepsilon,\xi^0} \\ \sigma_{\varepsilon,\xi^0} & \sigma_\varepsilon^2 \end{pmatrix}$. Note that this allows the baseline and efficacy components $\xi_i^0$, $\varepsilon_i$ of study $i$ to be correlated.

Table 4.1 gives the posterior distribution for the population mean $\mu$ (see Theorem 2 in Appendix B) and a Monte Carlo procedure for probabilistic sensitivity analysis. This posterior distribution is bivariate normal with covariance matrix equal to the matrix inverse of the sum $\sum_i \Omega_i$, where each $\Omega_i$ is itself the matrix inverse of the sum $\Sigma_{\xi 0} + \Sigma_i$, and $\Sigma_i = \begin{pmatrix} \sigma_{i0}^2 & -\sigma_{i0}^2 \\ -\sigma_{i0}^2 & \sigma_{i0}^2 + \sigma_{i1}^2 \end{pmatrix}$ is the covariance matrix for $y_i = (y_i^0, y_i^1 - y_i^0)$. The unknown population mean $\mu$ has posterior mean vector $\bar{y}$ given by the product of the covariance matrix $(\sum_i \Omega_i)^{-1}$ with the sum of matrix products $\Omega_i y_i$.

The Monte Carlo procedure described in the table requires the generation of a bivariate normal vector $\xi = (\xi^0, \varepsilon)$. We discuss how to accomplish this in Appendix A. The posterior distribution on $\xi = (\xi^0, \varepsilon)$ is bivariate normal with mean $\bar{y}$ and covariance matrix $(\sum_i \Omega_i)^{-1} + \sum_{\xi 0}$.

These methods require estimates of the prior covariance matrix $\Sigma_{\xi 0}$. It can be shown that any estimate of the form

$$\hat{\Sigma}_{\xi 0} = \frac{\hat{\Sigma} - \sum_i \alpha_i (1 - \alpha_i) \Sigma_i}{\sum_i \alpha_i (1 - \alpha_i)} \tag{6}$$

$$\bar{y} = \sum_i \alpha_i y_i \qquad \hat{\Sigma} = \sum_i \alpha_i (y_i - \bar{y})(y_i - \bar{y})^T$$

is unbiased as long as $\sum_i \alpha_i = 1$. A natural choice is

$$\alpha_i = \frac{\sigma_{i0}^{-2} \sigma_{i1}^{-2}}{\sum_i \sigma_{i0}^{-2} \sigma_{i1}^{-2}}.$$

Unfortunately, for a small number of studies this estimate is highly variable and may give an estimate $\hat{\Sigma}_{\xi 0}$ that is not a legitimate covariance matrix. The covariance matrix $\Sigma_{\xi 0} = \begin{pmatrix} \sigma_{\xi^0}^2 & \sigma_{\varepsilon,\xi^0} \\ \sigma_{\varepsilon,\xi^0} & \sigma_\varepsilon^2 \end{pmatrix}$ is legitimate provided that both $\sigma_{\xi 0}^2$ and the conditional variance $\sigma_{\varepsilon|\xi^0}^2 = \sigma_\varepsilon^2 - \frac{\sigma_{\varepsilon,\xi^0}^2}{\sigma_{\xi^0}^2}$ of effect $\varepsilon$ given baseline $\xi^0$ are

**Table 4.2** Binomial Observations from Heterogeneous Controlled Studies, and Their Large-Sample Normal Approximation

| Parameters for Study $i$ | Observations for Study $i$ | Transformation to an Approximate Normal Random Effects Model (Table 4.1) | Generating New Parameter Values: Do Steps 1 and 2 in Table 4.1. Then |
|---|---|---|---|
| Probabilities $p_i^0$, $p_i^1$ | $k_i^0$ events in $n_i^0$ independent trials without intervention<br>$k_i^1$ events in $n_i^1$ independent trials with intervention<br>$k_i^0 \sim \text{binomial}(n_i^0, p_i^0)$<br>$k_i^1 \sim \text{binomial}(n_i^1, p_i^1)$ | $\xi_i^0 = p_i^0$<br>$\xi_i^1 = p_i^1$<br>(no transformation)<br>$y_i^0 = k_i^0/n_i^0$<br>$y_i^1 = k_i^1/n_i^1$<br>$\sigma_{i0}^2 = (k_i^0/n_i^0)(1-k_i^0/n_{i0})/n_i^0$<br>$\sigma_{i1}^2 = (k_i^1/n_i^1)(1-k_i^1/n_i^1)/n_i^1$ | Set $p^0 = \xi^0$, $p^1 = \xi^1$ |
| | | $\xi_i^0 = \text{logit}(p_i^0)$<br>$\xi_i^1 = \text{logit}(p_i^1)$<br>$y_i^0 = \text{logit}(k_i^0/n_i^0)$<br>$y_i^1 = \text{logit}(k_i^1/n_i^1)$<br>$\sigma_{i0}^2 = 1/k_i^0 + 1/(n_i^0 - k_i^0)$<br>$\sigma_{i1}^2 = 1/k_i^1 + 1/(n_i^1 - k_i^1)$ | Calculate<br>$p_0 = \dfrac{e^{\xi^0}}{1+e^{\xi^0}}$,<br>$p_1 = \dfrac{e^{\xi^1}}{1+e^{\xi^1}}$. |
| | | $\xi_i^0 = \ln(p_i^0)$<br>$\xi_i^1 = \ln(p_i^1)$<br>$y_i^0 = \ln(k_i^0/n_i^0)$<br>$y_i^1 = \ln(k_i^1/n_i^1)$<br>$\sigma_{i0}^2 = 1/k_i^0 - 1/n_i^0$<br>$\sigma_{i1}^2 = 1/k_i^1 - 1/n_i^1$ | Calculate<br>$p^0 = e^{\xi^0}$,<br>$p^1 = e^{\xi^1}$, or $RR = e^{\varepsilon}$ |

**Table 4.3** Poisson Observations from Heterogeneous Controlled Studies, and Their Large-Sample Normal Approximation

| Parameters for Study $i$ | Observations for Study $i$ | Transformation to an Approximate Normal Random Effects Model (Table 4.1) | Generating New Parameter Values: Do Steps 1 and 2 in Table 4.1. Then |
|---|---|---|---|
| Rates $\lambda_i^0$, $\lambda_i^1$ | $k_i^0$ events in duration $\Delta t_i^0$ without intervention<br>$k_i^1$ events in duration $\Delta t_i^1$ with intervention<br>$k_i^0 \sim \text{Poisson}(\lambda_i^0 \Delta t_i^0)$<br>$k_i^1 \sim \text{Poisson}(\lambda_i^1 \Delta t_i^1)$ | $\xi_i^0 = \lambda_i^0$<br>$\xi_i^1 = \lambda_i^1$<br>(no transformation)<br>$y_i^0 = k_i^0/\Delta t_i^0$<br>$y_i^1 = k_i^1/\Delta t_i^1$<br>$\sigma_{i0}^2 = k_i^0/\Delta t_i^{0\,2}$<br>$\sigma_{i1}^2 = k_i^1/\Delta t_i^{1\,2}$ | Set $\lambda^0 = \xi^0$, $\lambda^1 = \xi^1$ |
| | | $\xi_i^0 = \ln(\lambda_i^0)$<br>$\xi_i^1 = \ln(\lambda_i^1)$<br>$y_i^0 = \ln(k_i^0/\Delta t_i^0)$<br>$y_i^1 = \ln(k_i^1/\Delta t_i^1)$<br>$\sigma_{i0} = 1/k_i^0$<br>$\sigma_{i1}^2 = 1/k_i^1$ | Calculate<br>$\lambda^0 = e^{\xi^0}$,<br>$\lambda^1 = e^{\xi^1}$, or $\text{RR} = e^{\varepsilon}$ |

non-negative. (Note it is not enough that $\sigma_{\xi 0}^2$ and $\sigma_\varepsilon^2$ both be nonnegative.) In the case where one or both of $\hat{\sigma}_{\xi 0}^2$, $\hat{\sigma}_{\varepsilon|\xi 0}^2$ are negative, one must revise these estimates subjectively.

Table 4.2 and Table 4.3 summarize the corresponding large-sample normal approximations for binomial and Poisson observations in heterogeneous controlled studies.

**Example 8: The Effect of Zidovudine Prophylaxis on HIV Transmission (continuation of Example 4)**

In Example 4, we pooled the 7 studies estimating HIV transmission risk with and without zidovudine prophylaxis. In fact, if we calculate the test statistic (3) with the control observations $y_i^0$, we obtain $Q_0 = 39.83$ with $P$-value $4.9 \times 10^{-7}$. The corresponding statistic $Q_\varepsilon$ using the treatment effects $y_i^1 - y_i^0$ gives $Q_\varepsilon = 38.6$ with $P$ value $8.7 \times 10^{-7}$. The null hypothesis of homogeneous studies is therefore untenable (and is better not entertained anyway, as we have argued above).

A random effects analysis using the log transformation from Table 4.2 may be performed as follows.

| Group $i$ | $y_i^0 = \log(k_i^0/n_i^0)$ | $\sigma_{i0}^2 = 1/k_i^0 - 1/n_i^0$ | $y_i^1 = \log(k_i^1/n_i^1)$ | $\sigma_{i1}^2 = 1/k_i^1 - 1/n_i^1$ | $y_i^1 - y_i^0$ |
|---|---|---|---|---|---|
| 1 | −1.174 | 0.0147 | −3.086 | 0.05 | −1.913 |
| 2 | −1.443 | 0.015 | −2.546 | 0.058 | −1.102 |
| 3 | −1.153 | 0.0228 | −2.773 | 0.045 | −1.62 |
| 4 | −0.9808 | 0.10417 | −1.981 | 0.216 | −0.94 |
| 5 | −1.344 | 0.0246 | −1.8 | 0.044 | −0.457 |
| 6 | −1.677 | 0.022 | −2.377 | 0.05 | −0.7 |
| 7 | −1.6988 | 0.0008 | −2.3199 | 0.004 | −0.621 |

The matrix estimate (6) of $\Sigma_{\xi 0}$ yields $\hat{\sigma}_{\xi 0} = 0.267$ and $\hat{\sigma}_\varepsilon = 0.541$, $\hat{\sigma}_{\varepsilon,\xi 0} = -0.134$. The corresponding estimated conditional variance is $\hat{\sigma}^2_{\varepsilon|\xi 0} = 0.040 \geq 0$, so $\hat{\Sigma}_{\xi 0}$ is a legitimate covariance matrix. Results from Table 4.1 are

$$\bar{y} = \begin{pmatrix} -1.41 \\ -1.06 \end{pmatrix} \qquad \left(\sum_i \Omega_i\right)^{-1} = \begin{pmatrix} 0.013 & -0.023 \\ -0.022 & 0.053 \end{pmatrix},$$

and we conclude that $\mu = (\mu^0, \mu^\varepsilon)$ has approximate bivariate normal posterior distribution with this mean and this covariance matrix. The predictive distribution of $\xi = (\xi_0, \varepsilon)$ is normal with the same mean and covariance matrix

$$\Sigma_{\xi 0} + \left(\sum_i \Omega_i\right)^{-1} = \begin{pmatrix} 0.0847 & -0.1571 \\ -0.1571 & 0.3451 \end{pmatrix}.$$

We may compare the implied joint posterior distribution on $p_0$, $p_1$ with the implied posterior obtained by (incorrectly) pooling the 7 studies as we did in Example 4, and the result is in Figure 5. The posterior distributions are quite different, which is consistent with the strong rejection of the hypothesis of homogeneity that we noted above.

The reader may recall (Example 3 and Figure 3) that the lognormal and logit-normal approximations for study 6 were less than exact, suggesting the use of the untransformed option in Table 4.2. Unfortunately, the resulting bivariate posterior substantially overlaps $p_1 < 0$ and is therefore not logically feasible. Because the remaining 6 studies all have values of $n\hat{p}(1-\hat{p})$ safely exceeding 10, we felt comfortable in using the log transform for this analysis.

**Example 9: Relative Risk of No Prenatal Care on HIV Prevalence**

Mrus and Tsevat cite studies by Lindsay and colleagues[40] and Donegan and colleagues[41] on the relative risk of no prenatal care on HIV infection. We analyze these studies using Table 4.1 and the log transformation in Table 4.2. The relevant observations are as follows.

| | With Prenatal Care | | | Without Prenatal Care | | |
|---|---|---|---|---|---|---|
| Study i | $k_i^0$ | $n_i^0$ | $k_i^0/n_i^0$ | $k_i^1$ | $n_i^1$ | $k_i^1 - 1/n_i^1$ |
| 1 | 26 | 7356 | 0.0035 | 12 | 834 | 0.0144 |
| 2 | 82 | 3891 | 0.0211 | 11 | 254 | 0.0433 |

With only 2 studies, there is very little information on study heterogeneity and we may expect difficulty in estimating $\hat{\Sigma}_{\xi 0}$. Indeed, the estimate (6) is

$$\hat{\Sigma}_{\xi 0} = \begin{pmatrix} \hat{\sigma}^2_{\xi^0} & \hat{\sigma}_{\varepsilon,\xi^0} \\ \hat{\sigma}_{\varepsilon,\xi^0} & \hat{\sigma}^2_\varepsilon \end{pmatrix} = \begin{pmatrix} 1.569 & -0.585 \\ -0.585 & 0.124 \end{pmatrix},$$

not a legitimate covariance matrix, because

$$\sigma^2_{\varepsilon|\xi^0} = \sigma^2_\varepsilon - \frac{\sigma^2_{\varepsilon,\xi^0}}{\sigma^2_{\xi^0}} = -0.148 < 0.$$

We therefore subjectively revise $\sigma^2_{\varepsilon|\xi 0}$. Based on the data above, it seems reasonable to follow Mrus and Tsevat and assign a mean relative risk ($RR$) in

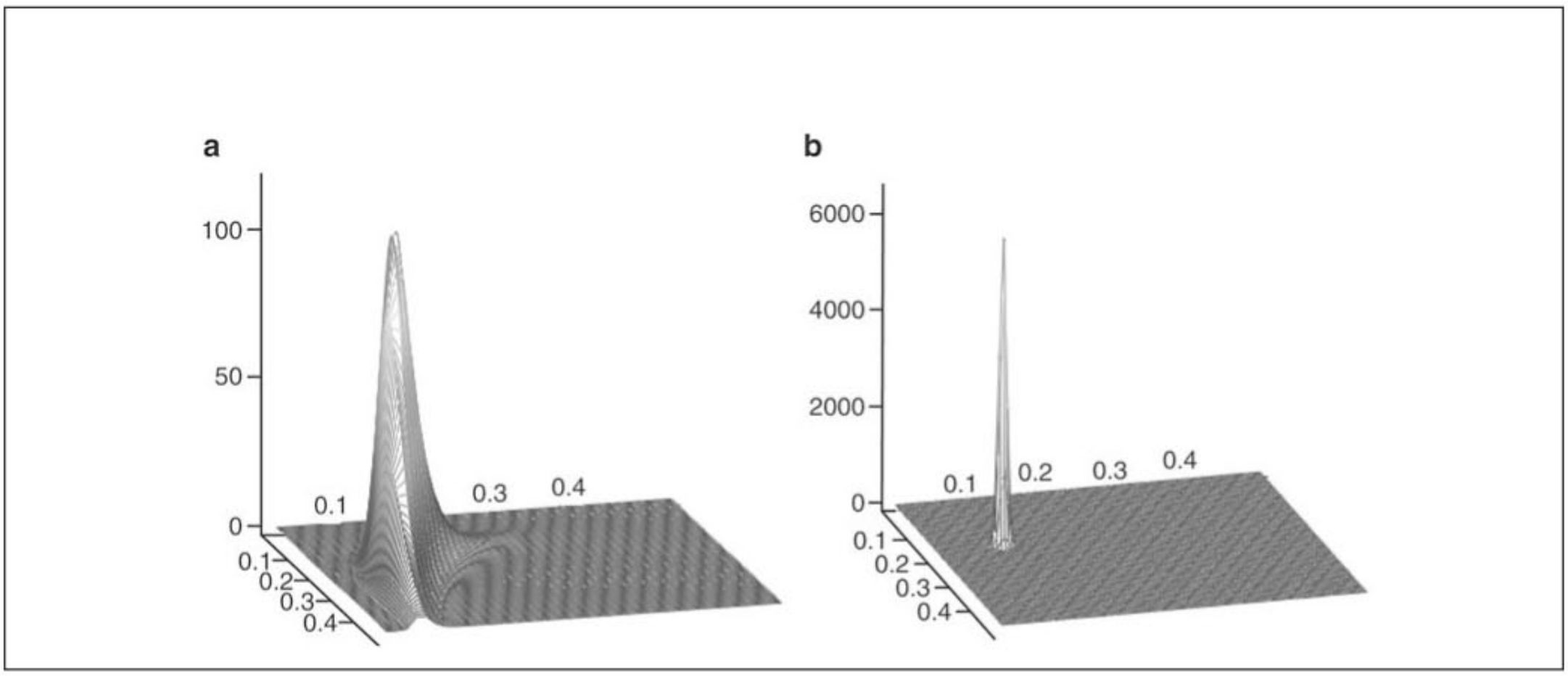


*Figure 5 Joint log-normal approximate posterior distributions of HIV vertical transmission risk ($p_0$) without intervention and vertical transmission risk ($p_1$) with zidovudine prophylaxis based on data from 7 studies. In (a), the studies are treated as heterogeneous and not pooled. In (b) the studies are (incorrectly) pooled. The posterior in (b) is a tight spike centered near (0.2, 0.1) that does not adequately account for study heterogeneity.*

the range 2–4 to the lack of prenatal care. With $\varepsilon = \ln(RR)$, this corresponds to a range of $\ln(4)$–$\ln(2) = 0.69$ on $\varepsilon$ and a standard deviation $0.69/3.92 = 0.177$. The conditional standard deviation $\sigma_{\varepsilon|\xi^0}$ is likely to be smaller. Nevertheless, we err on the conservative side and take $\hat{\sigma}_{\varepsilon|\xi^0} = 0.177$. This gives

$$\hat{\sigma}_{\varepsilon}^2 = \hat{\sigma}_{\varepsilon|\xi^0}^2 + \frac{\hat{\sigma}_{\varepsilon,\xi^0}^2}{\hat{\sigma}_{\xi^0}^2} = 0.250$$

and a revised, legitimate estimate

$$\hat{\Sigma}_{\xi 0} = \begin{pmatrix} 1.569 & -0.585 \\ -0.585 & 0.250 \end{pmatrix}.$$

Given the data, the approximate posterior distribution of parameter $\mu = (\mu^0, \mu^\varepsilon)$ is then multivariate normal with mean vector $\begin{pmatrix} -4.745 \\ 1.055 \end{pmatrix}$ and covariance matrix $\Sigma_\mu = \begin{pmatrix} 0.797 & -0.305 \\ -0.305 & 0.179 \end{pmatrix}$. The predictive distribution of $\xi = (\xi^0, \varepsilon)$ has the same mean and covariance matrix $\Sigma_\mu + \hat{\Sigma}_{\xi 0} = \begin{pmatrix} 2.365 & -0.890 \\ -0.890 & 0.429 \end{pmatrix}$. The resulting 95% prediction interval on $\varepsilon$ is $1.055 \pm (1.96)(0.429)^{1/2} = 1.055 \pm 1.284$. This translates to a 95% interval (0.795, 10.364) on $RR = e^\varepsilon$, much wider than the Mrus and Tsevat range 2–4, which does not account for study heterogeneity.

## Combining Heterogeneous Controlled and Uncontrolled Studies

It may happen that in addition to controlled studies, there may be studies that have no treatment arm. The methods from Table 4.1 and Table 4.2 still apply in this case if they are modified as follows. For studies $i$ without treatment arm, in Table 4.2 set the missing observation $y_i^1$ to an arbitrary value and omit the calculation of $\sigma_{i1}^2$. Then in Table 4.1 set

$$\Omega_i = \begin{pmatrix} \dfrac{1}{\sigma_{i0}^2 + \sigma_{\xi^0}^2} & 0 \\ 0 & 0 \end{pmatrix}$$

(which is the form taken by $\Omega_i$ when $\sigma_{i1}^2$ is infinitely large). Then apply the procedures in Table 4.1 and Table 4.2 as before. The zero entries in $\Omega_i$ ensure that the arbitrary value of $y_i^1$ does not affect the computation of $\bar{y}$ given in Table 4.1.

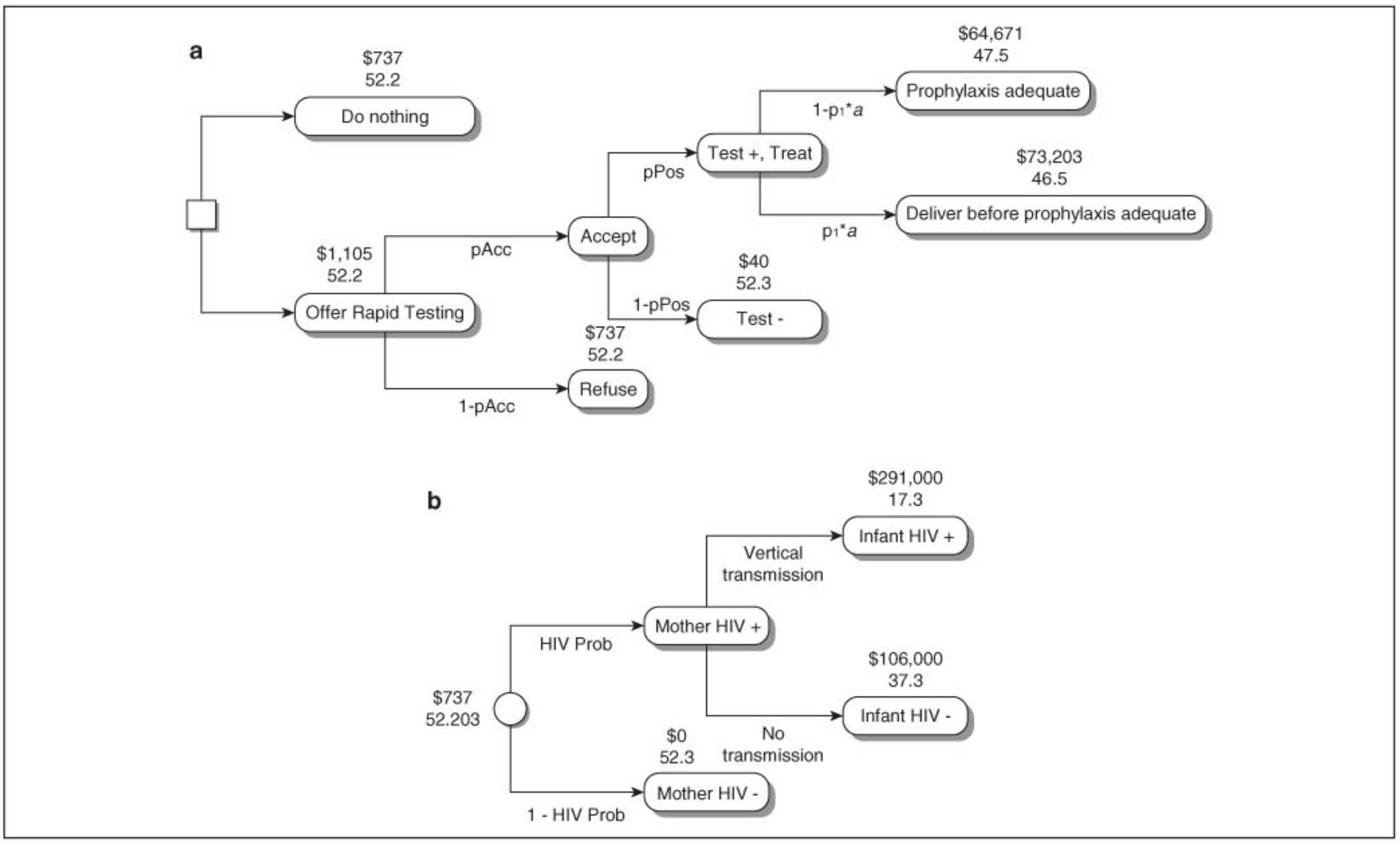


*Figure 6 The decision to offer rapid testing in labor followed by zidovudine prophylaxis, based on Mrus and Tsevat.[16] Figures adjacent to nodes are expected cost and quality-adjusted life years (QALYs)—for instance, Offer Rapid Testing has an expected cost of $1105 and expected QALY of 52.2 years. In (a), a mother who accepts rapid HIV testing receives zidovudine upon a positive test, unless delivery occurs before treatment can take effect. The probability pPos of a positive test depends on sensitivity, specificity, and HIV prevalence. The subtree (b), in which mother and infant HIV status is revealed, follows each of the terminal nodes in (a). Branch probabilities in (b) depend on the path taken in (a).*

## AN ILLUSTRATIVE PROBABILISTIC SENSITIVITY ANALYSIS

We now apply these methods to conduct a probabilistic sensitivity analysis for all probability and efficacy parameters in Mrus and Tsevat's recently published analysis of zidovudine prophylaxis following rapid HIV testing in labor to prevent vertical HIV transmission in pregnant women.[16] Figure 6 shows a decision tree depicting this problem, and in Table 5 we list all 8 probability and efficacy parameters. In the examples above, we have already calculated posterior distributions for 6 of these parameters, as the table indicates. The baseline values and ranges listed there are the medians and 95% credible intervals based on these posterior distributions.

The following discussion provides further detail on our use of these parameters.

*Prevalence of HIV infection in pregnant women without prenatal care:* Mrus and Tsevat estimate the prevalence of HIV infection in pregnant women in the United States to be $p_0 = 0.0017$ according to a completed survey for the year 1991. They assign a relative risk (*rrHIV*) in the range 2–4 to women without prenatal care. The resulting prevalence of HIV infection without prenatal care is $p = p_0 \cdot rrHIV$.

Because the estimate $p_0$ is based on a very large sample, we take the value $p_0 = 0.0017$ as fixed and include only the relative risk parameter *rrHIV* in our probabilistic sensitivity analysis. As we have mentioned, Mrus and Tsevat cite studies by Lindsay and colleagues[40] and Donegan and colleagues[41] on the relative risk of HIV infection without prenatal care. In Example 9, we used these studies to obtain a posterior on *rrHIV* using Table 4.1 and the log transformation in Table 4.2.

*Percentage of women delivering before preventative therapy can take effect:* Mrus and Tsevat estimate the percentage $p$ of women delivering before preventative therapy can take effect as the product $p = p_1 \cdot a$ of the percentage $p_1$ of women without prenatal care

**Table 5** Probability and Relative Risk Parameters in the Mrus and Tsevat Model. The Baseline Values Listed Are Posterior Medians Obtained Using Techniques from Table 1 through Table 4

| Parameter | Description | Baseline Value (range) |
|---|---|---|
| $p_0$ | HIV prevalence | 0.0017 |
| *rrHIV* | Relative risk of no prenatal care on HIV prevalence (Example 9) | 2.872 (0.796–10.368) |
| *pTransNo* | Vertical transmission risk without intervention (Example 8) | 0.244 (0.138–0.431) |
| *rrTransZid* | Relative reduction in vertical transmission risk of zidovudine prophylaxis (Example 8) | 0.347 (0.109–1.096) |
| *pAcc* | Test and treatment acceptance rate (Example 7) | 0.852 (0.572–0.964) |
| $p_1$ | Probability of delivery within 4 hours | 0.549 (0.467–0.631) |
| *a* | Percentage of women delivering before preventative therapy can take effect | 0.5 (0.25–0.75) |
| *Sens* | Rapid HIV antibody test Sensitivity (Example 6) | 0.998 (0.910–0.99996) |
| *Spec* | Rapid HIV antibody test Specificity (Example 5) | 0.988 (0.981–0.995) |

who deliver within 4 h of presentation and the percentage $a$ of such women who would deliver before preventative therapy can take effect.

Mrus and Tsevat take $a = ½$ based on an estimated 2 h needed to test, treat, and derive benefit from the drug. Our re-analysis uses $a = ½$ as baseline, but for probabilistic sensitivity analysis, we take $a$ to have a beta distribution with mean ½ and 95% credible interval 0.25–0.75, that is, a beta(7,7) distribution.

Regarding $p_1$, Mrus and Tsevat cite Grobman and Garcia[42] as observing $k = 306$ of $n = 557$ women ($\hat{p}_1 = 54.9\%$, $\sigma_1 = 2.1\%$) without prenatal care delivering within 4 h of presentation. We take this as a single study from a heterogeneous population (Table 3.1) and use the "no transformation" results in Table 3.2. With only 1 study, we must subjectively estimate $\sigma_\xi^2$ where $\xi$ is in this case the (unobserved) proportion $p_1$ estimated by a randomly chosen study. Ideally the estimate $\sigma_\xi^2 = \sigma_{p_1}^2$ should be informed by some knowledge of how or why $p_1$ might vary across studies. In the absence of such information here, we take an approximate 2-standard-deviation range for $p_1$ to be 0.10, yielding $\sigma_{p_1} = 0.10/3.92 = 0.0255$. Invoking Table 3.1, we obtain an approximate normal posterior on $\mu$ with mean $\hat{p} = 0.549$ and standard deviation $(0.021^2 + \sigma_{p_1}^2)^{1/2} = 0.0331$. A future observation $p_1$ has approximate normal distribution with mean 0.549 and standard deviation $(0.0331^2 + \sigma_{p_1}^2)^{1/2} = 0.0418$. This is twice the standard deviation that would have been obtained assuming homogeneity.

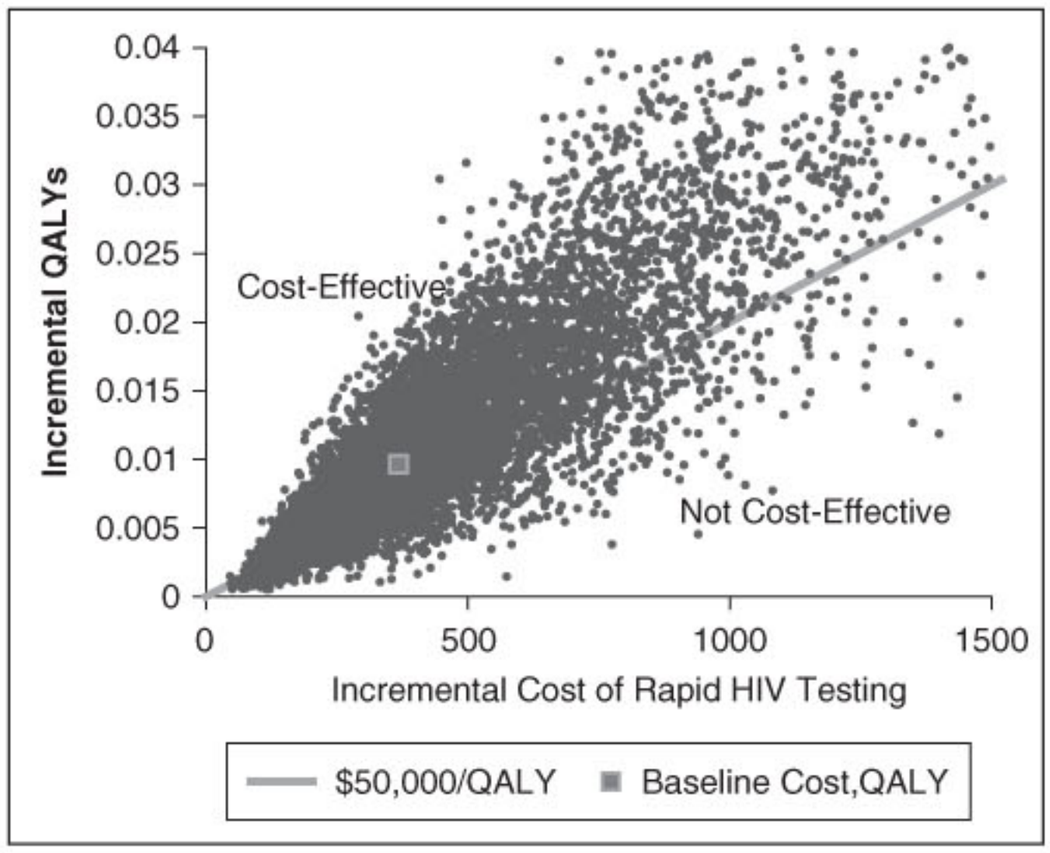


*Figure 7 Incremental cost and quality-adjusted life year (QALY) results for 10,000 iterations of a probabilistic sensitivity analysis on all probability and rate parameters in Mrus and Tsevat's analysis of zidovudine prophylaxis for HIV transmission in pregnancy. Baseline expected cost and QALYs are also shown.*

## Results of Probabilistic Sensitivity Analysis

For illustrative purposes, we re-analyze the version of Mrus and Tsevat's model that includes $100,000 additional lifetime costs associated with early HIV detection and treatment. (Mrus and Tsevat estimated

a range of $0–$200,000 on this cost and used $0 in their primary analysis.) In our base case analysis, we used the medians of the posterior distributions for the 8 probability and efficacy parameters discussed in this article, along with Mrus and Tsevat's estimates for the remaining cost and quality-of-life parameters. This analysis and the subsequent probabilistic sensitivity analysis was conducted in Microsoft Excel using the *Stotree* software written by one of the authors.[43-46] Using the standard value of $50,000/QALY, we found a net benefit of $114 per pregnancy for rapid HIV testing followed by zidovudine prophylaxis (C/E ratio of $38,142).

We performed a probabilistic sensitivity analysis jointly on the 8 probability and efficacy parameters discussed in this article, leaving the remaining cost and quality parameters at the levels specified by Mrus and Tsevat. We note, however, that the uncertainty surrounding the cost and quality parameters in this model is substantial, and these should also be included in a probabilistic sensitivity analysis. For the purposes of this article, however, we confine our analysis to the probability and efficacy parameters.

Our Monte Carlo simulation takes random draws for parameters from their posterior distributions and estimates the probability of decision change and the expected improvement[b] in net benefit were the values of all 8 parameters known prior to the decision to offer rapid HIV testing.

Based on 10,000 Monte Carlo iterations (see Figure 7), zidovudine prophylaxis has a 77.3% (s.e. 0.4%) probability of optimality and the expected improvement in net benefit on all 8 relative effects and probabilities equal to $18.20 ($\bar{X}s_x$ $0.64) per pregnancy. This figure is not large compared to the baseline expected net benefit ($114), and we conclude that the optimality of zidovudine prophylaxis is only somewhat sensitive to simultaneous variation in the 8 probability and efficacy parameters in the model.

It is of interest to determine which parameters contribute most strongly to the expected improvement value. As Table 6 shows, the vertical transmission parameters *pTransNo* and *rrTransZid* together make the largest contribution, with an estimated expected improvement of $13.02, and no other parameter contributes significantly on an individual basis. (Note that it is not meaningful to isolate individual expected improvement values for *pTransNo* and *rrTransZid*, as they are not probabilistically independent variables.)

[b]As noted by Felli and Hazen,[3] expected improvement is equal to expected value of perfect information in many decision problems. However, here net benefit is a nonlinear function of the probability and efficacy parameters, so equality does not hold.

**Table 6** Estimated Expected Improvement Values in Our Mrus and Tsevat Reanalysis for All Parameters and Parameter Subsets of Size 1 or 2, Based on 10,000 Monte Carlo Iterations Each

| Parameters | Expected Improvement Given Foreknowledge (with Estimated Standard Error) |
|---|---|
| All parameters | $18.20 (0.064) |
| *pTransNo, rrTransZid* | $13.02 (0.39) |
| $p_1$, *a* | $0.560 (0.048) |
| *rrHIV* | $0.123 (0.012) |
| *Sens, Spec* | $0.004 (0.003) |
| *pAcc* | $0 (0) |

Expected improvement given foreknowledge is equal to the expected increase in net benefit if the value of parameter in question were revealed prior to the decision to offer rapid HIV testing, and is approximately equal to the expected value of perfect information. Only the vertical transmission parameter combination *pTransNo, rrTransZid,* with expected improvement of $13.02, contributes significantly to the expected improvement for all 8 parameters jointly, which was equal to $18.20.

## CONCLUSION

We have presented large-sample Bayesian methods for obtaining approximately normal posterior distributions for transformations of probability, rate, and relative-effect parameters given data from either controlled or uncontrolled studies. These methods can account for issues that are rarely addressed in standard probabilistic sensitivity analyses, such as heterogeneity across studies and correlations between control and treatment parameters. Analysts can then generate random variates from these distributions and reverse transform to obtain probability, rate, or relative-effect random variates for the purposes of probabilistic sensitivity analysis. This variate generation and transformation procedure is simple enough to be conducted on a spreadsheet, as we have done. Because these are large-sample approximations, caution must be exercised in applying these techniques to studies with small sample sizes, as we have discussed. The results we present may also be used for the calculation of expected value of *sample* information using Monte Carlo simulation, much in the manner of Ades and colleagues.[19]

Our results assume that population variances within each study—the quantities $\sigma^2$ in Table 1.1; $\sigma_0^2$, $\sigma_1^2$ in Table 2.1; $\sigma_i^2$ in Table 3.1; $\sigma_{i0}^2$, $\sigma_{i1}^2$ Table 4.1—are well estimated by sample variances, and this is very likely to be so for large samples. However, as we have pointed out, when the number of studies is small, there is likely to

be little or no information on cross-study heterogeneity, and the estimates of $\sigma_\xi^2$ in Table 3.1 and the covariance matrix $\Sigma_{\xi 0}$ in Table 4.1 must rely on subjective input. A fully Bayesian approach would calculate posterior distributions of these parameters given study data, but such posteriors do not take on a simple form and may even be improper when a noninformative prior distribution is employed.[13] Because we use point estimates for these cross-study variances instead of calculating posterior distributions, the predictive distributions from which we sample during probabilistic sensitivity analysis are too tight to some unknown degree, much in the same way that a normal distribution based on known variance is tighter than the corresponding *t*-distribution that includes uncertainty about variance. The resulting probabilistic sensitivity analysis may therefore underestimate problem sensitivity. A fully Bayesian approach to combining heterogeneous studies would require the specification of a proper prior distribution for cross-study variance and the use of Markov chain Monte Carlo techniques to estimate posterior distributions.

## APPENDIX A
## BIVARIATE NORMAL RANDOM VARIATE GENERATION

See Tong,[47] Anderson,[48] or similar references for multivariate normal characteristics. When $x = (x_1, x_2)$ is bivariate normal with mean $(m_1, m_2)$ and covariance matrix $\begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}$, generate a bivariate normal vector $x$ using the following steps:

Step 1: Generate $x_1$ from a normal distribution with mean $m_1$ and variance $\sigma_1^2$.

Step 2: Generate $x_2$ from a normal distribution with mean $m_2 + (x_1 - m_1)\sigma_{12}/\sigma_1^2$ and variance $\sigma_2^2 - \sigma_{12}^2/\sigma_1^2$, which is the conditional distribution of $x_2$ given $x_1$.

## APPENDIX B
## NORMAL CONJUGATE POSTERIOR DISTRIBUTIONS

**Lemma 1**[14,17]**:** If a random vector $y$ has a multivariate normal distribution $y \mid \mu \sim \text{normal}(\mu, \Sigma)$, and $\mu$ is parameterized as normal$(\mu_0, \Sigma_0)$, then

**i)** The posterior on $\mu$ given $y$ is also normally distributed: $\mu \mid y_1, \ldots, y_n \sim \text{normal}(\mu_n, \Sigma_n)$ where

$$\mu_n = (\Sigma^{-1} + \Sigma_0^{-1})^{-1} (\Sigma^{-1} y + \Sigma_0^{-1} \mu_0), \quad \Sigma_n = (\Sigma^{-1} + \Sigma_0^{-1})^{-1}$$

**ii)** The marginal distribution of $y$ is normal$(\mu_0, \Sigma_0 + \Sigma)$.

**Theorem 1**[14]**:** Suppose the independence structure of Figure 2 holds, and moreover that $y_1, \ldots, y_n, \xi_1, \ldots, \xi_n$, and $\mu$ are univariate parameters with $y_i \mid \xi_i \sim \text{normal}(\xi_i, \sigma_i^2)$, and $\xi_i \mid \mu \sim \text{normal}(\mu, \sigma_\xi^2)$.

**i)** If $\mu$ has prior distribution normal$(\mu_0, \sigma_0^2)$, then the posterior on $\mu$ given $y_1, \ldots, y_n$ is also normally distributed: $\mu \mid y_1, \ldots, y_n \sim \text{normal}(\mu_n, \sigma_n^2)$.

**ii)** If $\mu$ has noninformative uniform prior distribution that is, $f(\mu) \propto_\mu 1$, then $\mu \mid y_1, \ldots, y_n \sim \text{normal}\left(\bar{y}, \frac{1}{\sum_i \omega_i}\right)$.

Here

$$\mu_n = \frac{\sum_i \omega_i \bar{y} + \tau_0 \mu_0}{\sum_i \omega_i + \tau_0}, \quad \sigma_n^2 = \frac{1}{\tau_n}, \quad \tau_n = \sum_i \omega_i + \tau_0,$$

$$\omega_i = \frac{1}{\sigma_i^2 + \sigma_\xi^2}, \quad \bar{y} = \frac{\sum_i \omega_i y_i}{\sum_i \omega_i}, \quad \tau_0 = \frac{1}{\sigma_0^2}.$$

We are unable to find the following result in the literature and have derived it ourselves. A proof is available from the authors upon request.

**Theorem 2:** Suppose the independence structure of Figure 2 holds, and moreover that $y_1, \ldots, y_n, \xi_1, \ldots, \xi_n$, and $\mu$ are $d$-dimensional parameters with $y_i \mid \xi_i \sim \text{normal}(\xi_1, \Sigma_i)$, and $\xi_i \mid \mu \sim \text{normal}(\mu, \Sigma_{\xi 0})$.

**i)** If $\mu$ has prior distribution normal$(\mu_0, \Sigma_0)$, then the posterior on $\mu$ given $y_1, \ldots, y_n$ is also normally distributed: $\mu \mid y_1, \ldots, y_n \sim \text{normal}(\mu_n, \Sigma_n)$.

**ii)** If $\mu$ has noninformative uniform prior distribution, that is, $f(\mu) \propto_\mu 1$, then $\mu \mid y_1, \ldots, y_n \sim \text{normal}\ (\overline{y}, (\sum_i \Omega_i)^{-1})$.

Here

$$\mu_n = \left(\sum_i \Omega_i + \Sigma_0^{-1}\right)^{-1} \left(\sum_i \Omega_i y_i + \Sigma_0^{-1} \mu_0\right),$$

$$\Sigma_n = \left(\sum_i \Omega_i + \Sigma_0^{-1}\right)^{-1},$$

$$\Omega_i = (\Sigma_i + \Sigma_{\xi 0})^{-1}, \qquad \bar{y} = \left(\sum_i \Omega_i\right)^{-1} \left(\sum_i \Omega_i y_i\right).$$

## APPENDIX C:
## MODELS WITH UNBALANCED DATA

Suppose there are $m$ controlled studies and $n$-$m$ studies assessing only baseline risk (no treatment

groups). The $n$-$m$ baseline study groups give data $y_i^0$, which estimate parameters $\xi_i^0$, $i = 1,2, \ldots n$-$m$. For these studies, $y_i^0 \mid \xi_i^0$ are independent and have normal distributions with mean $\xi_i^0$ and variance $\sigma_{i0}^2$. And $\xi_1^0, \ldots, \xi_{n-m}^0$ given $\mu_0$ are i.i.d. random variables, each is normal$(\mu_0, \sigma_{\xi^0}^2)$. In the other $m$ study groups with treatment arms, we have $y_i = (y_i^0, x_i)^T$, $\xi_i = (\xi_i^0, \varepsilon_i)^T$, where $x_i = y_i^1 - y_i^0$, $\varepsilon_i = \xi_i^1 - \xi_i^0$. Here $y_i \mid \xi_i$ is normal $(\xi_i, \varepsilon_i)$, where $\Sigma_i$ is estimated by $\Sigma_i = \begin{pmatrix} \sigma_{i0}^2 & -\sigma_{i0}^2 \\ -\sigma_{i0}^2 & \sigma_{i0}^2 + \sigma_{i1}^2 \end{pmatrix}$. And $\xi_{n-m+1}, \ldots \xi_n$ given $\mu = (\mu^0, \mu^\varepsilon)^T$ are i.i.d. (2-dimensional) random vectors, each is normal$(\mu, \Sigma_{\xi 0})$, with $\Sigma_{\xi 0} = \begin{pmatrix} \sigma_{\xi^0}^2 & \sigma_{\varepsilon,\xi^0} \\ \sigma_{\varepsilon,\xi^0} & \sigma_\varepsilon^2 \end{pmatrix}$.

A Bayesian approach to obtain the posterior for $\mu = (\mu^0, \mu^\varepsilon)^T$ would be as follows: Under the assumption of noninformative uniform prior distribution for $\mu = (\mu^0, \mu^\varepsilon)^T$, update the distribution for $\mu_0$ (using the methods developed in Table 3.1 and Table 3.2) according to the data from the ($n$-$m$) groups, which do not contain studies for treatment. Then use this distribution for $\mu^0$ and noninformative distribution for $\mu^\varepsilon$ as the prior to obtain the posterior distribution for $\mu = (\mu^0, \mu^\varepsilon)^T$ given the data from the remaining $m$ control groups and $m$ treatment groups (using the conclusion in Theorem 2). This approach leads to the following result. A proof is available from the authors upon request.

**Theorem 3:** For a random effects model with unbalanced data, described above, suppose $\mu = (\mu^0, \mu^\varepsilon)^T$ has noninformative uniform prior distribution. Then $\mu \mid y_1, \ldots, y_n \sim$ normal$(\bar{y}, (\sum_{i=1}^n \Omega_i)^{-1})$, where

$$\bar{y} = \left(\sum_{i=1}^n \Omega_i\right)^{-1} \left(\sum_{i=1}^n \Omega_i y_i\right),$$

$$\Omega_i = \begin{pmatrix} \omega_i & 0 \\ 0 & 0 \end{pmatrix},$$

$$\omega_i = \frac{1}{\sigma_{i0}^2 + \sigma_{\xi^0}^2} \quad i = 1, 2, \ldots n\text{-}m,$$

and

$$\Omega_i = \left(\Sigma_i + \Sigma_{\xi 0}\right)^{-1}$$

$$= \begin{pmatrix} \sigma_{i0}^2 + \sigma_{\xi^0}^2 & -\sigma_{i0}^2 + \sigma_{\varepsilon,\xi^0} \\ -\sigma_{i0}^2 + \sigma_{\varepsilon,\xi^0} & \sigma_{i0}^2 + \sigma_{i1}^2 + \sigma_\varepsilon^2 \end{pmatrix}^{-1} \quad i = n\text{-}m+1, \ldots n.$$

## APPENDIX D
## ACCURACY OF LOG AND LOGIT TRANSFORMS

The results for transformed parameters in Tables 1-4 draw on the following well-known result.

**Lemma 2**[18] **(Delta-Method):** If $X_n$ is approximately normal$(\mu, \sigma_n^2)$, where $\sigma_n$ is a sequence of constants tending to zero, and $g$ is a differentiable function, then for large $n$, $g(X_n)$ is approximately normal with mean $g(\mu)$ and standard deviation $\sigma_n \mid g'(\mu) \mid$.

If $X$ is normal$(\mu, \sigma^2)$ and $Y = g(X)$, then what is the error in approximating $Y$ by a normal$(g(\mu), \sigma^2 g'(\mu)^2)$ random variable $Y'$? Assuming $g$ is increasing, a good measure of error is the difference in cumulative probabilities

$$\mathrm{Err}(t) = \mathrm{P}(Y' \le g(\mu + t\sigma)) - \mathrm{P}(Y \le g(\mu + t\sigma))$$

$$= \Phi\left(\frac{g(\mu + t\sigma) - g(\mu)}{g'(\mu)\sigma}\right) - \Phi(t),$$

where $\Phi(\cdot)$ is the standard normal distribution function. The following result is based on a Taylor series approximation to $\mathrm{Err}(t)$ around $t = 0$. A proof is available from the authors upon request.

**Lemma 3:** If $g$ is an increasing differentiable function, and $\varphi$ is the standard normal density, then

$$\mathrm{Err}(t) = \tfrac{1}{2}\varphi(t)\frac{g''(\mu)}{g'(\mu)}t^2\sigma + o(t^2).$$

and

$$|\mathrm{Err}(t)| \le \frac{1}{2}(2\pi)^{-\frac{1}{2}}e^{-1}\left|\frac{g''(\mu)}{g'(\mu)}\right|\sigma + o(t^2).$$

The bounds (1) and (2) in the text follow from the lemma by taking

$$g(t) = \mathrm{logit}(t) \qquad \frac{g''(t)}{g'(t)} = \frac{2(t-1)}{t(1-t)}$$

$$g(t) = \ln(t) \qquad \frac{g''(t)}{g'(t)} = \frac{1}{t}.$$

$$\mu = \hat{p} \qquad \sigma = \sqrt{\tfrac{1}{n}\hat{p}(1-\hat{p})}$$

For the Poisson case, take

$$\mu = k/\Delta t \qquad \sigma = \sqrt{k}/\Delta t.$$

## REFERENCES


1. Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo simulation. Med Decis Making. 1985;5(2):157–77.

2. Critchfield GG, Willard KE. Probabilistic analysis of decision trees using Monte Carlo simulation. Med Decis Making. 1986; 6(2):85–92.

3. Felli JC, Hazen GB. Sensitivity analysis and the expected value of perfect information. Med Decis Making. 1998;18(1):95–109.

4. Felli JC, Hazen GB. Erratum: sensitivity analysis and the expected value of perfect information. Med Decis Making. 2003; 23(1):97.

5. Critchfield GG, Willard KE, Connelly DP. Probabilistic analysis methods for general decision models. Comput Biomed Res. 1986;19:254–65.

6. Williams P, Dowson AJ, Rapoport AM, Sawyer J. The cost effectiveness of stratified care in the management of migraine. Pharmacoeconomics. 2001;19(8):819–29.

7. Murakami Y, Ohashi Y. Projected number of diabetic renal disease patients among insulin-dependent diabetes mellitus children in Japan using a Markov model with probabilistic sensitivity analysis. Int J Epidemiol. 2001;30(5):1078–83.

8. Ng AK, Weeks JC, Mauch PM, Kuntz KM. Decision analysis on alternative treatment strategies for favorable-prognosis, early-stage Hodgkin's disease. J Clin Oncol. 1999;17(11):3577–85.

9. Pasta DJ, Taylor JL, Henning JM. Probabilistic sensitivity analysis incorporating the bootstrap: an example comparing treatments for the eradication of Helicobacter pylori. Med Decis Making. 1999;19(3):353–63.

10. Goodman CA, Coleman PG, Mills AJ. Cost-effectiveness of malaria control in subSaharan Africa. Lancet. 1999;354(9176): 378–85.

11. Lord J, Asante MA. Estimating uncertainty ranges for costs by the bootstrap procedure combined with probabilistic sensitivity analysis. Health Econ. 1999;8(4):323–33.

12. Spiegelhalter DJ, Best NG. Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling. Stat Med. 2003;22:3687–709.

13. DuMouchel WH. Bayesian metaanalysis. In: Berry DA, ed. Statistical Methodology in the Pharmaceutical Sciences. New York: Marcel Dekker; 1990.

14. Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian Approaches to Clinical Trials and Health-Care Evaluation. Chichester (UK): Wiley; 2004.

15. Ades AE, Cliff S. Markov chain Monte Carlo estimation of a multiparameter decision model: consistency of evidence and the accurate assessment of uncertainty. Med Decis Making. 2002;22:359–71.

16. Mrus JM, Tsevat J. Cost-effectiveness of interventions to reduce vertical HIV transmission from pregnant women who have not received prenatal care. Med Decis Making. 2004;24(1):30–9.

17. Bernardo JM, Smith AFM. Bayesian Theory. Chichester (UK): Wiley; 2000.

18. Sen PK, Singer JM. Large Sample Methods in Statistics: An Introduction With Applications. Boca Raton (FL): Chapman & Hall/CRC; 1994.

19. Ades AE, Lu G, Claxton K. Expected value of sample information calculations in medical decision modelling. Med Decis Making. 2004;24(4):207–27.

20. DeGroot MH. Optimal Statistical Decisions. New York: McGraw-Hill; 1970.

21. Devore JL. Probability and Statistics for Engineering and the Sciences. Pacific Grove (CA): Duxbury; 2000.

22. Johnson NL, Kotz S, Kemp AW. Univariate Discrete Distributions. 2nd ed. New York: John Wiley; 1993.

23. Ross S. A First Course in Probability. 7th ed. Englewood Cliffs (NJ): Prentice Hall; 2006.

24. Olkin I, Gleser LJ, Derman C. Probability Models and Applications. 2nd ed. New York: Macmillan College Publishing; 1994.

25. Pittman J. Probability. New York: Springer-Verlag; 1993.

26. Rajegowda BK, Das BB, Lala R, Rao S, McNeeley OF. Expedited human immunodeficiency virus testing of mothers and newborns with unknown HIV status at time of labor and delivery. J Perinat Med. 2000;28:458–63.

27. Connor EM, Sperling RS, Gelber R, et al. Reduction of maternal-infant transmission of human immunodeficiency virus type 1 with zidovudine treatment. Pediatric AIDS Clinical Trials Group Protocol 076 Study Group. N Engl J Med. 1994;331:1173–80.

28. Fiscus SA, Schoenbach VJ, Wilfert C. Short courses of zidovudine and perinatal transmission of HIV. N Engl J Med. 1999; 340:1040–1.

29. Wade NA, Birkhead GS, Warren BL, et al. Abbreviated regimens of zidovudine prophylaxis and perinatal transmission of the human immunodeficiency virus. N Engl J Med. 1998;339:1409–14.

30. Guay LA, Musoke P, Fleming T, et al. Intrapartum and neonatal single-dose nevirapine compared with zidovudine for prevention of mother-to-child transmission of HIV-1 in Kampala, Uganda: HIVNET 012 randomised trial. Lancet. 1999;354:795–802.

31. The International Perinatal HIV Group. The mode of delivery and the risk of vertical transmission of human immunodeficiency virus type 1—a meta-analysis of 15 prospective cohort studies. N Engl J Med. 1999;340:977–87.

32. Wiktor SZ, Ekpini E, Karon JM, et al. Short-course oral zidovudine for prevention of mother-to-child transmission of HIV-1 in Abidjan, Cote d'Ivoire: a randomised trial. Lancet. 1999;353:781–5.

33. Shaffer N, Chuachoowong R, Mock PA, et al. Short-course zidovudine for perinatal HIV-1 transmission in Bangkok, Thailand: a randomised controlled trial. Bangkok Collaborative Perinatal HIV Transmission Study Group. Lancet. 1999;353:773–808.

34. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7:177–88.

35. Larsen K, Petersen JH, Budtz-Jorgensen E, Endahl L. Interpreting parameters in the logistic regression model with random effects. Biometrics. 2000;56:909–14.

36. Kelen GD, Shahan JB, Quinn TC. Emergency department-based HIV screening and counseling: experience with rapid and standard serologic testing. Ann Emerg Med. 1999;33:147–55.

37. Irwin K, Olivo N, Schable CA, Weber JT, Janssen R, Ernst J. Performance characteristics of a rapid HIV antibody assay in a hospital with a high prevalence of HIV infection. CDC-Bronx-Lebanon HIV Serosurvey Team. Ann Intern Med. 1996;125:471–5.

38. Malone JD, Smith ES, Sheffield J, et al. Comparative evaluation of six rapid serological tests for HIV-1 antibody. J Acquir Immune Defic Syndr. 1993;6:115–9.

39. Rajegowda BK, Das BB, Lala R, Rao S, McNeeley DF. Expedited human immunodeficiency virus testing of mothers and newborns with unknown HIV status at time of labor and delivery. J Perinat Med. 2000;28:458–63.

40. Lindsay MK, Feng TI, Peterson HB, et al. Routine human immunodeficiency virus infection screening in unregistered and registered inner-city parturients. Obstet Gynecol. 1991;77: 599–603.

41. Donegan SP, Steger KA, Recla L, Slade BA, Willis S, Klein L. Seroprevalence of human immunodeficiency virus in parturients at Boston City Hospital: implications for public health and obstetric practice. Am J Obstet Gynecol. 1992;167:622–9.

42. Grobman WA, Garcia PM. The cost-effectiveness of voluntary intrapartum rapid human immunodeficiency virus testing for women without adequate prenatal care. Am J Obstet Gynecol. 1999;181:1062–71.

43. Hazen GB. Stochastic trees: a new technique for temporal medical decision modeling. Med Decis Making. 1992;12:163–78.

44. Hazen GB. Factored stochastic trees: a tool for solving complex temporal medical decision models. Med Decis Making. 1993;13:227–36.

45. Hazen GB. Stochastic trees and the StoTree modeling environment: models and software for medical decision analysis. J Med Syst. 2002;26:399–413.

46. Hazen GB. Stochastic trees and the StoTree modeling software. Available from http://users.iems.nwu.edu/~hazen/.

47. Tong YL. The Multivariate Normal Distribution. New York: Springer-Verlag; 1990.

48. Anderson TW. An Introduction to Multivariate Statistical Analysis. 3rd ed. New York: John Wiley; 2003.