Use of a Bayesian Approach to Decide When to Stop a Therapeutic Trial: The Case of a Chemoprophylaxis Trial in Human Immunodeficiency Virus Infection

Alphonse Kpozéhouen¹, Ahmadou Alioum², Xavier Anglaret¹, Philippe Van de Perre³, Geneviève Chêne¹, and Roger Salamon¹

¹ INSERM U 593, Université Victor Segalen Bordeaux 2, France.
² INSERM E 0338, Université Victor Segalen Bordeaux 2, France.
³ Regional Hospital Center, Laboratory of Bacteriology-Virology, Montpellier, France.

Received for publication March 5, 2003; accepted for publication July 29, 2004.

From 1996 to 1998, a phase III, placebo-controlled, therapeutic trial was conducted in Abidjan, Ivory Coast, to assess the efficacy of cotrimoxazole prophylaxis in reducing severe morbidity in adults at early stages of human immunodeficiency virus infection. The authors used the real data from this trial to simulate three Bayesian interim analyses. Three prior distributions were considered: a noninformative one, a skeptical one, and one based on external information. The posterior distribution was calculated by using directed acyclic graphs and Gibbs sampling. This Bayesian approach showed different results according to the prior distribution chosen. Although use of the noninformative prior would have led to stopping the trial at the same time that the frequentist approach would have, the skeptical prior would have led to continuing it, and the prior based on external data would have led to stopping it 1 year earlier.

Bayesian approach; cotrimoxazole; counting process; Cox model; directed acyclic graph; Gibbs sampling

Abbreviations: ANRS, Agence Nationale de Recherche sur le Sida; CI, Côte d’Ivoire; HIV, human immunodeficiency virus.
A randomized trial funded by ANRS (France) was begun in Abidjan, Ivory Coast, by a Franco-Ivorian team at the beginning of April 1996. This double-blind, placebo-controlled trial aimed to assess the tolerance and efficacy of cotrimoxazole (800 mg of sulfamethoxazole + 160 mg of trimethoprim in one daily dose) in preventing serious complications of HIV infection in adult patients not having reached the acquired immunodeficiency syndrome stage and followed up as outpatients at a dispensary in Yopougon, a suburb of Abidjan. “Serious complications,” the primary outcome in this trial, were defined as all morbid events leading to hospitalization and/or death. The patients included in the trial received all medical care free of charge (consultations, drugs, hospitalizations, transport) (11). The trial was planned to last 3 years (1 year of recruitment and 2 years of follow-up after the final recruitment). The number of subjects required, 730 in total, had been calculated to detect a 30 percent reduction in the incidence of serious complications with \( \alpha = 5 \) percent and a power of 90 percent, according to the hypothesis that the incidence of serious events would be 40 percent in the placebo arm. Analysis was conducted on an intent-to-treat basis. An independent data safety monitoring board was responsible for controlling the running of the trial and performing interim analyses. The first interim analysis was set to be performed at 18 months or, in the case of more numerous events than planned, at the moment when half of the expected events had occurred. The stopping rule used in this trial was the O’Brien-Fleming rule (12). Two interim analyses were performed by the members of the data safety monitoring board of this trial: one in October 1997 using the data obtained until July 31, 1997, and concerning 394 patients included until that date; and another in March 1998 using data obtained until December 1, 1997, and concerning 509 patients included (figure 1). On March 17, 1998, the data safety monitoring board recommended terminating the trial because the results of the second interim analysis showed a 49 percent reduction (11) in serious events in the cotrimoxazole group in relation to the placebo group. The final analysis was performed on data obtained until March 17, 1998 (stopping day of trial) and concerned 541 patients included until that date (figure 1). It showed a 43 percent reduction in serious events in the cotrimoxazole group compared with the placebo group.

In the present study, we simulated three interim analyses by using a Bayesian approach. The first one was performed with data obtained until December 31, 1996 (7 months before the first frequentist interim analysis), and the two others were conducted with data obtained at the same dates as those for the two interim frequentist analyses.
The Bayesian approach

Model. The model considered here is the Bayesian proportional hazards model proposed by Spiegelhalter et al. (13). Let \( T \) be the delay of occurrence of the events of interest, \( \mathbf{z} \) the vector of covariates, and \( n \) the total number of individuals included in the study. Let \( N_i(t) \) be the number of serious events that have occurred up to time \( t \) for individual \( i \) (\( i = 1, \ldots, n \)), considered as a process in time with intensity \( I_i(t) \). In this case, \( I_i(t)dt \) is the conditional expected value of \( dN_i(t) \) given \( F_i \); that is, \( I_i(t) = E(dN_i(t) | F_i) \), where \( dN_i(t) \) represents the increment of \( N_i(t) \) over the short-time interval \([t, t + dt] \) and \( F_i \) is a filtration (all data collected just before date \( t \)). Therefore, \( dN_i(t) \) takes value 1 if individual \( i \) has experienced a serious event during interval \([t, t + dt] \) and 0 otherwise. We can model \( I_i(t) \) by \( I_i(t) = Y_i(t)\lambda_i(t)\exp(\mathbf{z}_i')\beta \), where \( Y_i(t) \) takes values 0 or 1 according to whether the individual was observed at time \( t \), and \( Y_i(t)\lambda_i(t)\exp(\mathbf{z}_i')\beta \) is the hazard function under the proportional hazards model; \( \lambda_i(t) \) represents the baseline hazard function, \( \beta \) is the vector of the regression coefficients, and \( \mathbf{z}_i \) is the vector of the covariates for individual \( i \). Concretely, the data may be represented by \( D = \{N_i(t), Y_i(t), \mathbf{z}_i \mid i = 1, \ldots, n \} \). The unknown parameters are \( \beta \) and the baseline cumulative incidence \( \Lambda_0(t) = \int_0^t \lambda_0(u)du \). These parameters are assumed to be independent. The posterior distribution is given by \( \Pr(\beta, \Lambda_0(t) | \text{Data}) \), which is proportional to \( \Pr(\beta) \times \Pr(\Lambda_0(t)) \times \Pr(\text{Data} | \beta, \Lambda_0(t)) \), where \( \Pr(\beta) \) and \( \Pr(\Lambda_0(t)) \) are prior distributions of parameters \( \beta, \Lambda_0(t) \) and \( \Pr(\text{Data} | \beta, \Lambda_0(t)) \) is the likelihood (13). Under noninformative censoring, the likelihood is proportional to

\[
\prod_{i=1}^n \left[ \prod_{t \geq 0} I_i(t)^{dN_i(t)} \right] \exp(- \int_0^t I_i(u)du).
\]

The counting-process increments \( dN_i(t) \) in the time interval \([t, t + dt] \) are considered independent Poisson random variables with mean \( I_i(t)dt \).

Prior distributions. We used three types of prior distributions on \( \beta \): a noninformative prior, a skeptical prior (9), and a prior based on previously known results about the natural history of HIV infection in Africa (14). Following the methods of Kalbfleisch (15), we took the prior on \( \Lambda_0(t) \) to follow a gamma distribution with parameters \( c \times d\Lambda_0^c(t) \) and \( c \). We chose a gamma distribution to facilitate determination of the conjugated distribution, which is also a gamma because \( dN_i(t) \) follows a Poisson distribution. Here, we set \( d\Lambda(t) = r \times dt \), which can be thought of as a prior guess at the unknown hazard function, the parameter \( c \) represents the degree of confidence in this guess, and \( r \) corresponds to the number of serious events per unit of time. A low value of \( c \) corresponds to a low value of the prior distribution, whereas a high value of \( c \) corresponds to a high prior value. We took \( r = 500 \) and \( c = 0.001 \). As usual, parameter \( \beta \) represents the logarithm of the ratio of instantaneous risk of experiencing a serious event (placebo group vs. cotrimoxazole group) and makes it possible to assess the effect of treatment in the placebo group compared with the cotrimoxazole group. During the simulations, we kept the prior distribution on \( \lambda_0(t) \) proposed by Kalbfleisch (15) but modified the prior distribution on \( \beta \) and took it to be a normal distribution with mean \( \mu \) and variance \( \sigma^2 \).

Noninformative prior (vague prior). Because we had no information about \( \beta \), we assigned a noninformative prior to it, that is, \( \beta \sim N(0, 10^3) \). This prior distribution represented very weak prior information, so it has a little influence on the posterior distribution. The noninformative prior has the advantage of corresponding roughly to the classical frequentist approach.

Skeptical prior. The skeptical prior distribution attempts to formalize the belief that treatments are unlikely to be different; this prior represented a position that a null hypothesis is likely to be true. Let \( \beta \) be the logarithm of the hazard ratio between the placebo and cotrimoxazole groups and \( p_1 \) and \( p_2 \) the incidence of serious events in the placebo and cotrimoxazole groups, respectively. Suppose that the trial was designed with type I error \( \alpha \) and power \( 1 - \xi \). We have to test \( H_0: \beta = 0 \) versus \( H_1: \beta = \beta_s = \log(p_1/p_2) \). Consider that the aim of the trial was to assess the alternative hypothesis \( H_1 \). The skepticism concerning \( \beta_s \) is formulated by considering a prior normal distribution \( \beta \sim N(0, \sigma^2_{\text{scop}}) \) such that the probability \( \Pr(\beta > \beta_s) \) takes a small value \( \gamma \) (16, 17). Setting \( \gamma = 0.05 \), we can solve \( \sigma^2_{\text{scop}} \) by using the following equation: \( \beta_s/\sigma^2_{\text{scop}} = Z_1 - \gamma \), where \( Z_1 - \gamma = 1.64 \) is the corresponding standard normal deviate. In the Cotrimo-CI trial, the sample size was calculated to detect a reduction of 30 percent in the incidence of severe events, which was assumed to be 40 percent in the placebo group after a follow-up of 2 years. This translates to the alternative hypothesis with a log-hazard ratio of \( \beta_s = \log(p_1/p_2) = \log(0.40/0.28) = 0.36 \). Therefore, \( \sigma^2_{\text{scop}} = 0.05 \) and we have the skeptical prior \( -N(0, 0.05) \).

Prior based on external information. To determine the prior distribution based on previously known data about the history of HIV infection in Africa, we used the results of the study by Leroy et al. (14) regarding the natural history of HIV in Rwandan women. In that study, the relative risk of death was eightfold higher in the HIV-positive women than in the HIV-negative women. We supposed that the mortality of HIV-positive and HIV-negative women could reflect severe morbidity in women not taking cotrimoxazole and taking cotrimoxazole, respectively. In their study, the authors presented the number of deaths per group of women, the relative risk (RR), and the total number of women at risk. Let \( O_+ \) be the number of women having died when HIV-positive and \( O_- \) the number of HIV-negative women having died in the group of seronegative individuals. The prior on \( \beta \) was \( N(\log(4), 4/\sigma^2_{\text{scop}}) \) (18). According to the Leroy et al. (14) data, this gives \( \beta \sim N(2.09, 0.30) \).

Posterior distribution. The posterior distribution is given by

\[
\Pr(\beta, \Lambda_0(t) | \text{Data}) \propto \Pr(\beta) \times \Pr(\Lambda_0(t)) \times \prod_{i=1}^n \left[ \prod_{t \geq 0} I_i(t)^{dN_i(t)} \right] \exp(- \int_0^t I_i(u)du).
\]
To calculate the posterior distribution, we used a hierarchical model (19), shown in figure 2 by a directed acyclic graph (20). We took the hyperparameters to be independent. It is possible to calculate the distributions of the parameters or the conditional expectations of these parameters. Let $g(b, K_0(t))$ be a function of the parameters; then

$$E[g(b, K_0(t)|Data)] = \int_B \int_\Lambda g(\theta, A_0(t)|Data) Pr(\theta, A_0(t)|Data) d\theta dA_0(t).$$

$B$ and $\Lambda$ are the domains of variation of $b$ and $A_0(t)$. Because of the complexity of expressing a posterior distribution, we used Gibbs sampling (21, 22) and BUGS software (Bayesian Using Gibbs Sampling) (9) to estimate the parameters. To monitor convergence, we followed Gelman and Rubin’s recommendations (23). To diagnose approximate convergence, we estimated the potential reduction scale $R$ (23). Convergence of algorithm was established by using the program BOA (Bayesian Output Analysis) (24). We chose to form a sample of 6,000 values by taking every fifth iterate. We performed 1,200 iterations. The first 200 were used for burn-in and were discarded (25, 26). The lowest and upper limits of the credible interval were, respectively, 2.5 percent and 97.5 percent of the sample (27).

Comparison of the effect of treatments. There are several methods of comparing treatment effects in the Bayesian framework. It is possible to do the following:

1. Calculate $D(t) = \Pr(S(t, z = \text{placebo}) < S(t, z = \text{cotrimoxazole})|Data)$ (6), the predictive probability that the survival function in the placebo group is smaller than the survival function in the cotrimoxazole group at time $t$. The trial is stopped at time $t$ if $D(t)$ is at least equal to 95 percent.

2. Calculate $d(t) = E(S(t; z = \text{placebo}) - S(t; z = \text{cotrimoxazole})|Data)$ (6), the predictive mean difference between the survival functions in the control and treatment groups at time $t$. The trial is stopped at time $t$ if $d(t)$ is at least equal to 95 percent.

3. Use a rule based on the hazard ratio (the placebo group being in the numerator of this ratio). According to the Cotrimo-CI trial investigators, the incidence of severe events was expected to be 40 percent in the placebo group and 28 percent in the cotrimoxazole group, that is, a 30 percent reduction in the incidence (11). That means a hazard ratio at least equal to 1.43 = 0.4/0.28; on a log-hazard ratio scale, $\beta > \log(1.43) = 0.36$. We therefore calculated $\Delta(t) = \Pr(\beta > 0.36|Data)$. The trial investigators decided that the trial would be stopped if $\Delta(t)$ was at least equal to 95 percent.

For this study, we decided to use the third method and calculated $\Delta(t)$ by using a Monte Carlo integration (28).
We simulated three intermediate analyses (29, 30) taking place 253, 465, and 618 days after the first patient was recruited (figure 1). The second and third simulations corresponded to the first and second interim analyses really performed in the Cotrimo-CI trial (figure 1). The date of the first interim simulation was set arbitrarily at 7 months before the first interim analysis was performed in the Cotrimo-CI trial.

**RESULTS**

After 1,200 iterations, we obtained a good convergence for the potential reduction scale $R$ (23). Table 1 shows, for each randomization group, the number of patients enrolled, the number of patients who experienced at least one severe event, and the result of the log-rank test on the date of the three simulated Bayesian analyses.

<table>
<thead>
<tr>
<th>Interim analysis</th>
<th>Enrolled</th>
<th>At least one severe event</th>
<th>Log-rank test result</th>
</tr>
</thead>
<tbody>
<tr>
<td>First: 253 days; December 31, 1996 ($n = 150$)</td>
<td>75 75</td>
<td>34 47</td>
<td>0.16</td>
</tr>
<tr>
<td>Second: 465 days; July 31, 1997 ($n = 394$)</td>
<td>198 196</td>
<td>70 99</td>
<td>0.06</td>
</tr>
<tr>
<td>Third: 618 days; December 31, 1997 ($n = 509$)</td>
<td>254 255</td>
<td>83 121</td>
<td>0.001</td>
</tr>
</tbody>
</table>

* Log-rank test result.
† Information in parentheses, number of patients enrolled.
‡ Cx, cotrimoxazole; Pb, placebo.

Figures 3, 4, and 5 show the prior and the posterior distributions of $\beta$ at the three simulated interim analyses for noninformative, skeptical, and external information priors, respectively.

Noninformative prior distribution. Figure 3 shows the estimated density of $\beta$ obtained by using kernel density estimation (31). The values of $\beta$ were almost always positive at the three interim analyses, therefore pointing to the superiority of the cotrimoxazole treatment. The largest mode of approximately 0.67 (95% credible interval: 0.25, 1.07) was obtained at the third analysis and corresponded to a relative risk of 1.95 of experiencing at least one serious event in the placebo group compared with the cotrimoxazole group. There was a right shift of the logarithmic curve of relative risk, indicating an increased effect of cotrimoxazole compared with placebo at each analysis. We found no significant difference in $\beta$ values.
FIGURE 4. Posterior distribution of $\beta$ when its prior is normal $(0, 0.05)$ (skeptical prior), Cotrimo-CI ANRS 059 trial, Abidjan, Ivory Coast, 1996–1998.

FIGURE 5. Posterior distribution of $\beta$ when its prior is normal $(2.09, 0.30)$ (prior distribution based on external information), Cotrimo-CI ANRS 059 trial, Abidjan, Ivory Coast, 1996–1998.
A Bayesian Approach to Decide When to Stop a Trial

The aim of this paper was to illustrate the advantage of a Bayesian approach in calculating a posterior distribution of the log-hazard ratio between the two treatments. This right shift was more pronounced for the second and third analyses. The largest mode of \( \beta \) was 0.43 (95 percent credible interval: 0.08, 0.78) and was obtained at the third analysis.

Prior distribution based on external information. By including external information from the Rwandan study (14), the value of \( \beta \) was positive for all analyses (figure 5). The third analysis gave the lowest mode: 0.88 (95 percent credible interval: 0.46, 1.29). The risk of experiencing a serious event was then 2.41 times greater in the placebo group than in the cotrimoxazole group. These risks were 2.79 and 2.72 at the first and second analyses, respectively.

Table 2 shows the probabilities that the log-hazard ratio was greater than 0.36 for each interim analysis. When the vague prior was used, these probabilities were 0.17 and 0.92 at day 253 and day 465, respectively. These values were lower than 0.95, the chosen cutoff for stopping the trial. At the third analysis, 618 days after the first inclusion, \( \Delta(618) \) was equal to 0.98. A Bayesian approach based on this noninformative prior would therefore have led to stopping the trial at this third interim analysis, which took place at the time of the second interim frequentist analysis actually performed during the trial and that led to trial termination. When the skeptical prior was used, none of the three simulated interim analyses would have led to recommending that the trial be stopped, because each of the \( \Delta \) values remained below the chosen cutoff. Finally, if the prior distribution based on external data had been used, the Bayesian analyses would have led to stopping the trial 12 months before the analyses based on the vague prior and on frequentist analyses.

### DISCUSSION

In this study, we applied three interim Bayesian analytical simulations to the real data from a trial assessing the efficacy of cotrimoxazole in reducing severe morbidity in HIV-infected adults in Abidjan. In this trial, a frequentist interim analysis led to recommending that the trial be stopped. We found that the Bayesian approach would also have led to stopping the trial but that the choice of the prior could have influenced the time at which the trial would have been stopped.

The use of prior information is both a key component of the Bayesian approach and one of the reasons that it is often perceived as being marred by subjectivity. Indeed, all types of priors are not adapted to all types of situations, as illustrated by our simulations.

When results of previous studies or data linked to the topic are available, a prior distribution based on this external information can be chosen. This feature of the Bayesian approach is indeed one of the most interesting compared with the frequentist approach. In the case of cotrimoxazole prophylaxis in sub-Saharan Africa, no previous data were available when the Cotrimo-CI trial was designed. We used data comparing survival in HIV-infected and -noninfected women in Rwanda to attempt to illustrate the use of this type of prior distribution, but it was clearly not well adapted to the context. Unfortunately, external information is likely to be rarer in developing countries than in industrialized ones given the relative paucity of studies on the same subject and the difficulty in standardizing the results of studies performed in settings and with means that are often very different.

When no previous information is available, a noninformative prior distribution can be used. In our study, this one would probably have been the best choice. Noninformative, prior-based analyses would have led to recommending the trial to be stopped at the same time as the frequentist analyses did. This illustrates the fact that a noninformative prior is likely to lead to the same conclusions as the frequentist approach, because \( 1 - \Delta(t) \) can frequently correspond to the conventional (one-sided) \( p \) value where a single test is conducted (16).

Finally, use of a skeptical distribution is another interesting possibility of the Bayesian approach (17). In our case, this prior distribution would probably not have been the best one to use. The hypotheses used by the Cotrimo-CI trial team to calculate the number of subjects to be enrolled were very questionable, because knowledge of the natural history of early stages of HIV disease in sub-Saharan Africa was very scarce at the beginning of that trial. Indeed, at the end of the trial, the incidence of serious adverse events in participating HIV-infected adults appeared to be quite different from the initial hypotheses. Priors other than those discussed in this paper could also be considered, such as the enthusiastic prior (16, 17), the meta-analysis prior (17, 32), or the prior proposed by Kadane and Wolfson (33). Given that stopping rules are so sensitive to the choice of the prior, it is of the utmost importance that this choice be made according to standardized procedures. This goal can be achieved by questioning experts about the knowledge and beliefs regarding the issue to be studied. After being recorded on standardized questionnaires, the responses of the experts are processed and analyzed to deduce the choice of the most appropriate prior (17, 32–39).

The aim of this paper was to illustrate the advantage of a Bayesian approach in calculating a posterior distribution.
based on accumulated data up to the time of the analyses. There are other advantages of a Bayesian approach that were not shown in our study, such as the ability to find predictive probabilities by considering results that have not yet been observed (40–42).

The Bayesian approach has now become easier to implement thanks to the development of new algorithms and information tools. It would be interesting to popularize it so that it could be used by a wider public. The joint use of both frequentist and Bayesian analyses in as many further trials as possible, including trials conducted in developing countries, would make it possible to illustrate the advantages and disadvantages of both approaches and to better appreciate their respective positions in the research world.

ACKNOWLEDGMENTS

The first author was supported by a grant from the Agence Nationale de Recherche sur le Sida (ANRS, France), to which he is extremely grateful.

REFERENCES

24. Smith BJ. Bayesian Output Analysis Program (BOA), version 1.0.0. Iowa City, IA: Department of Biostatistics, University of Iowa College of Public Health, 2000.