# ASSESSMENT AND MONITORING IN CT WHEN SURVIVAL CURVES HAVE DISTINCT SHAPES: A BAYESIAN APPROACH WITH WEIBULL MODELING

BRUNO LECOUTRE, BIENVENU MABIKA

*Laboratoire de Mathmatiques Raphaël Salem, UMR C.N.R.S et Université de Rouen*

*Mathématiques Site Colbert, 76821 Mont-Saint-Aignan Cedex, France*

*E-mail: bruno.lecoutre@univ-rouen.fr*

AND

GERARD DERZKO

*SANOFI Recherche*

*374 rue du Professeur Joseph Blayac, 34184 Montpellier Cedex, France.*

*E-mail: gerard.derzko@sanofi-synthelabo.com*

Revised version, November 2000

## SUMMARY

The comparison of Weibull distributions with unequal shape parameters, in the case of right censored survival data obtained for independent samples, is considered within the Bayesian statistical methodology. The procedures are illustrated with the example of a mortality study where a new treatment is compared to a placebo. The posterior distributions about relevant parameters allowing to search for a conclusion of clinical superiority of the treatment, and the predictive distributions used to obtain an early stopping rule at an interim analysis, are considered for a class of appropriate priors.

# INTRODUCTION

An extensive use of time-to-event data is made in the latest stage of clinical research and development of new drugs. In the most common setting, two or more new treatments are to be compared with a standard in patients at risk of developing serious acute diseases, or of dying. A new treatment is deemed beneficial if it delays the occurrence of the first event of a pre-specified category, and the study objective is either to demonstrate the superiority of a new therapy over the standard, or to show that the two are equivalent.

These studies are designed on the basis of estimates drawn from previous studies in similar conditions, and of assumptions, simple enough and supposed to be robust; usual assumptions are that the event rates would be constant on the diverse studied therapies, and/or that the hazard ratio of events in the two therapies would be constant over time. Interim analyses of the data, while patients are still being accrued, sometimes induce significant changes in the study design (like its size), and may involve a re-evaluation of the validity of the initial assumptions made. The final analysis of the collected data is usually performed with non parametric Kaplan-Meier estimation, log-rank testing, or semi-parametric Cox's model analysis. Again the validity of the initial assumptions is to be considered at this stage, especially if it is apparent from the survival curves that they might be questionable. Further explorations of the data are required when treatment effects cannot be simply stated as uniform hazard ratio less than 1. They allow to design future studies of similar conditions more appropriately.

A common deviation from the usual assumptions is that the hazard rates are not constant. Weibull's model then represents an alternative to the more simple exponential model. This model may account for possible "shape" effects of treatments as well as "scale" effects. A Bayesian approach was proposed with this model to demonstrate the superiority of a treatment over a standard [1].

In this paper, we assume that the exponential model cannot appropriately fit the observed survival curves in all of the treatment groups, but a Weibull model does. Moreover, the treatment effect, if any, applies on shape, rather than scale parameter. The Bayesian statistical methodology is proposed to compare the groups, as well as to stop the study early if necessary. The Bayesian predictive approach [2, 3, 4, 5, 6] is a very appealing method [7] for stopping a study early. Similarly to stochastic curtailment [8, 9, 10] it simulates the probability of achieving the study target, conditionally on available data and simple conjectures about the future observations; but the simulations are explicitly based on either the hypotheses used to design the study, expressed in terms of the prior distribution, or on available data, or on both (see especially [11]). It is usual in clinical research to assume noninformative priors, as a study is expected to bring evidence by itself. Indeed, in this case, the posterior distribution at interim analysis and the predictive distribution for future observations are based solely on the data. But we will show how alternative choices of priors may be used to refining inference.

## EMIAT CLINICAL TRIAL

EMIAT (European Myocardial Infarction Amiodarone Trial) was a double blind, randomized, placebo-controlled mortality trial [12]. Patients included in the study were post-MI (myocardial infarction) patients, with a damaged left ventricular function (i.e. impaired ejection fraction). The observation period for each patient was up to two years. Furthermore the patients were stratified according to their ejection fraction, 31%-40% (stratum 1, moderately damaged) and 30% or less (stratum 2, severely damaged). Of particular interest in this population was the cardiac mortality. All studies in the same indication performed at the time EMIAT started had six month follow-up or less, on the well known ground that the death hazard is higher shortly before an acute myocardial infarction. Three month has thus been considered here a relevant duration for the assessment of early drug effect, in addition to the two-year cardiac mortality rate.

Interim analyses were planned approximately every third of the total expected number of deaths (namely 225), with the purpose of stopping the study early in case of manifest efficacy of amiodarone. Because of external circumstances, the first interim analysis was delayed until respectively 239 and 226 patients in stratum 1 and 200 and 214 patients in stratum 2 had been randomized to placebo and amiodarone. All these patients were considered in the analysis. Figure 1 displays the survival curves at this time.

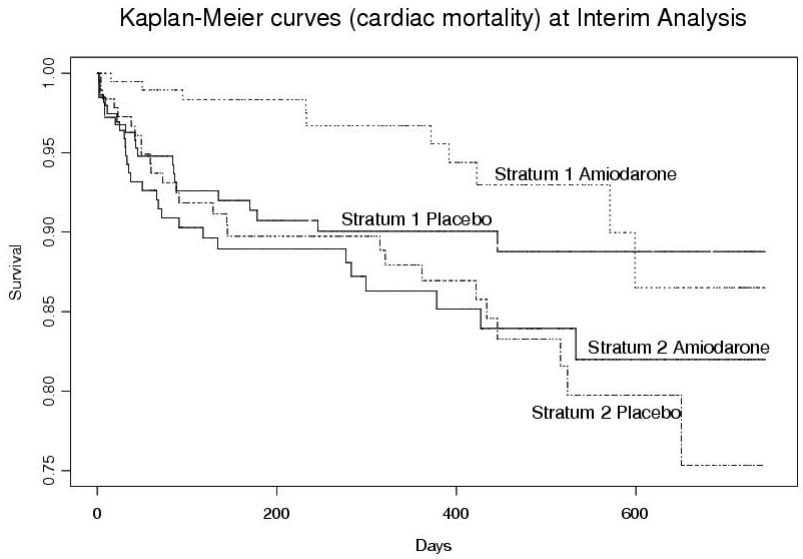


Figure 1: Kaplan-Meier estimates of cardiac mortality by group at interim analysis.

In stratum 1, there had been $\frac{20}{239}$ events in the placebo group and $\frac{10}{226}$ events in the amiodarone group. In this stratum, deaths tended to occur earlier among placebo patients than among treated patients [12]. In particular, 15 out of the 20 deaths occurred before three months in the placebo group versus only 2 out of the 10 in the amiodarone group. In stratum 2, there was no such differences. Respectively $\frac{26}{200}$ events and $\frac{26}{214}$ events had been observed in the two groups. A majority of deaths occurred before three months, $\frac{14}{26}$ among placebo patients and $\frac{18}{26}$ among treated patients. Thus it appeared that a constant mortality rate model was questionable to describe the observed evolution. Therefore we assumed that survival times within each of the

four groups followed a Weibull distribution, with unequal shape parameters.

# BAYESIAN METHODS

Conditionally upon the shape parameters, Weibull model comes down to the exponential model. For a class of appropriate priors, the corresponding conditional posterior distributions for the scale parameters and the marginal posterior distributions for the shape parameters can be made explicit. This extends previous results obtained for a single sample and the uncensored case [13]. Inferences for the parameters of interest are obtained by numerical methods. In particular straightforward and efficient simulation techniques can be easily implemented.

Concerning the predictive distributions at the time of interim analysis, the situation is complex. Indeed, in the case of censored survival data, the patients to be considered at this time are divided into four categories: (1) included patients for whom the event of interest has been observed, (2) Included patients definitely censored, (3) included patients under current observation for whom the maximum observation period is not ended, (4) new patients planned to be included after the time of the interim analysis. Thus we have to simultaneously consider two types of predictions, respectively relating to the third and fourth categories, taking into account the maximum observation period of these patients. Nigm [14] called these two type of predictions respectively "one sample" and "two sample" predictions. Here we extend the procedures previously developed in the case of exponential distributions [15] and in the case a single Weibull sample for uncensored data [13].

## Basic Results for One Sample

Let $X = (X_1, X_2, \cdots, X_n)$ a random sample from a Weibull distribution with pdf:

$$f(x|\alpha,\beta) = \frac{\beta}{\alpha} x^{\beta-1} \exp\Big(-\frac{x^\beta}{\alpha}\Big) \qquad x > 0$$

where $\beta > 0$ and $\alpha > 0$ (or more exactly $\alpha^{\frac{1}{\beta}}$) are usually referred to as shape and scale parameters. The corresponding likelihood function has two components, one for the $r$ patients

for whom the event of interest has occurred and corresponding to the survival (or failure) times $(x_1, x_2, \cdots, x_r)$, and the other for the remaining $(n-r)$ patients who were censored. Thus it is:

$$L(\alpha,\beta) = \beta^r \alpha^{-r} U^{\beta-1} exp\Big(-\tfrac{T_\beta}{\alpha}\Big) \quad \text{where } U = \textstyle\prod_{i=1}^r x_i \text{ and } T_\beta = \sum_{i=1}^n x_i^\beta$$

When $\beta$ is known the statistic $T_\beta$ is sufficient for $\alpha$. In this case $y = x^\beta$ has an exponential distribution with parameter $\alpha$, that is a gamma distribution with parameter 1 and scale factor $\alpha$. Consequently, a convenient family of prior densities is $\pi(\beta,\alpha) = \pi(\alpha|\beta)\pi(\beta)$ where, conditionally on $\beta$, $\alpha$ has an inverse gamma distribution $\alpha|\beta \sim r_o g_o(\beta) IG(r_o)$ and where $r_o \geq 0$ and $g_o(\beta)$ is an increasing function of $\beta$. Note that the constant $r_o$ could be replaced by an increasing function of $\beta$. For instance $\beta$ may have a gamma [14], uniform over $[b_1, b_2]$ [16], or beta over $[b_1, b_2]$ marginal distribution. Furthermore the parameters $\alpha$ and $\beta$ may be considered a priori independent if $g_o(\beta) = a_o$ $(a_o > 0)$. Lastly the usual noninformative prior $\pi(\alpha,\beta) \propto \frac{1}{\alpha\beta}$ [17] is included as a particular case for $r_o = 0$ and ?$\pi(\beta) \propto \frac{1}{\beta}$.

The joint posterior pdf of $\alpha$ and $\beta$ is proportional to $L(\alpha,\beta)\ \pi(\alpha|\beta)\pi(\beta)$. It is given by:

$$\pi(\alpha,\beta|data) \propto \beta^r U^{\beta-1} \Big(r_o g_o(\beta)\Big)^{r_o} \alpha^{-(r_o+r+1)}\ exp\Big(-\tfrac{1}{\alpha}\ \Big(T_\beta + r_o g_o(\beta)\Big)\Big)\pi(\beta)$$

Hence the marginal posterior pdf of $\beta$ is obtained from the inverted gamma integral (Box and Tiao [18], page 144):

$$\pi(\beta|data) = \textstyle\int_0^{+\infty} \pi(\alpha,\beta|data)d\alpha = K\beta^r U^{\beta-1}\Big(r_o g_o(\beta)\Big)^{r_o}\Big(T_\beta + r_o g_o(\beta)\Big)^{-(r_o+r)}\ \pi(\beta)$$

where $K$ is a normalizing constant. The conditional on $\beta$ posterior distribution of $\alpha$ can be identified as an inverse gamma distribution:

$$\alpha|\beta, data \sim T_{1\beta}\, IG(r_1) \quad \text{where } T_{1\beta} = T_\beta + r_o g_o(\beta) \text{ and } r_1 = r_o + r$$

Then the distributions of various derived parameters of interest can be easily deduced, in particular the mean of the Weibull distribution, $\alpha^{\frac{1}{\beta}}\Gamma(1+\frac{1}{\beta})$, and the mortality rate $\varphi_{[t]}$ at time $t$, $\varphi_{[t]} = 1 - \exp(-\frac{t^\beta}{\alpha})$.

## Predictive Distributions

The density of the posterior predictive distribution of a future observation $x'$ given available data is:

$$p(x'|data) = \int_0^\infty \int_0^\infty f(x'|\alpha,\beta,data)\pi(\alpha,\beta|data)d\alpha d\beta$$

Two cases are to be considered. For a new subject $f(x'|\alpha,\beta,data)$ does not depend on the previous observations and is the sampling density $f(x'|\alpha,\beta)$. Conditionally on $\beta$, $x'^{\beta}|\alpha,\beta,data \sim \mathcal{E}xp(\alpha)$, hence $\frac{x'^{\beta}}{\alpha}|\beta \sim G(1)$. From $\frac{T_{1\beta}}{\alpha}|\beta,data \sim G(r_1)$, since the ratio of two independent gamma distributions is an $F$ distribution, it is deduced that:

$$x'|\beta,data \sim \left(\frac{T_{1\beta}}{r_1}F_{2,2r_1}\right)^{\frac{1}{\beta}}$$

where $F_{2,2r_1}$ is the usual $F$ distribution with 2 and $2r_1$ degrees of freedom.

For a subject already observed until time $c$ (without event), the predictive distribution must be conditioned by the constraint $C : x' > c$. Consequently the sampling density of $y' = x'^{\beta}$ is replaced by:

$$f(y'|\alpha,\beta,data) = f^*(y'|\alpha,\beta,data)\ \frac{Pr^*(C|\alpha,\beta,data,y')}{Pr^*(C|\alpha,\beta,data)}$$

where an asterisk indicates the unconstrained distribution. Since the factor $Pr^*(C|\alpha,\beta,data,y')$ is equal to one when $y' > c^{\beta}$ and zero otherwise, we get (for $y' > c^{\beta}$):

$$f(y'|\alpha,\beta,data) = \frac{f(y'|\alpha,\beta)}{Pr(y'>c^{\beta}|\alpha,\beta)} = \frac{1}{\alpha}\exp\left(-\frac{c^{\beta}+y'}{\alpha}\right)$$

and conditionally on $\beta$:

$$p(y'|\beta,data) = \int_0^{+\infty} f(y'|\alpha,\beta,data)\,\pi(\alpha|\beta,data)d\alpha \propto \int_0^{+\infty} \alpha^{-(r_1+2)}\ \exp\left(-\frac{c^{\beta}+T_{1\beta}+y'}{\alpha}\right)d\alpha$$

Applying again the inverted gamma integral, it is found that:

$$p(y'|\beta,data) \propto (c^{\beta}+T_{1\beta}+y')^{r_1+1} \qquad y' > c^{\beta}$$

which can be identified as a truncated $F$ distribution with 2 and $2r_1$ degrees of freedom and scale parameter $\frac{c^{\beta}+T_{1\beta}}{r_1}$.

### Numerical Computations

The density of the posterior distribution of $\beta$ can be obtained by standard numerical integration. In the cases where this distribution is slightly dispersed, the posterior distribution of $\alpha$ and derived parameters can be approximated by the distribution conditional upon the mean of the posterior distribution of $\beta$. In any event the required inferences about parameters just as prediction of future observations can be easily obtained by simulation. Furthermore it can be expected that the posterior distribution of $\beta$ could be well approximated by a Gamma distribution with position and scale parameters having the same three first moments. Once this approximation has been obtained, all the computations only involve simulating gamma distributions.

All the procedures above are readily extended to the comparison of several survival curves, if the observations correspond to independent samples, each from a Weibull distribution with parameters ($\beta_g$,$\alpha_g$), and if the corresponding priors are assumed to be independent. In this case, the posterior distribution of any derived parameter of interest, such as a linear combination of the mortality rates in each group, can be straightforwardly obtained from the joint simulation of the corresponding independent marginals.

## RESULTS

### Interim Analysis

A noninformative prior distribution was first used for the interim analysis. The posterior distributions for the shape parameters $\beta_g$ are shown in Figure 2.

Their moments and approximations by Gamma distributions are given in Table 1. These approximations are excellent and virtually no difference was found in subsequent analyses between the simulations based on the exact densities and the one based on their approximations. It

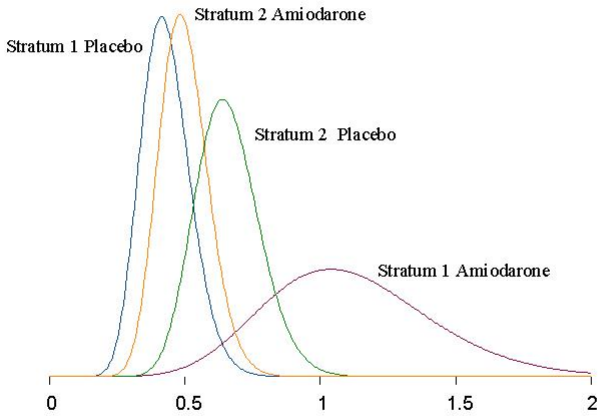


FIGURE 2

Figure 2: Posterior distributions for the shape parameters $\beta_g$ at interim analysis (noninformative prior).

appears in Table 1 that the assumption of a constant hazard rate was doubtful in three of the four subgroups. Moreover, in stratum 1 one may infer that the drug effect was large on the shape parameter.

**Table 1. Interim analysis with a noninformative prior: moments and approximations of the posterior distributions for the shape parameters $\beta_g$**

| | | $\mathbb{E}(\beta_g)$ | $\mathbb{E}(\beta_g^2)$ | $\mathbb{E}(\beta_g^3)$ | Approximations |
|---|---|---|---|---|---|
| Stratum 1 | Placebo | 0.432 | 0.195 | 0.092 | -0.032 + $0.018G(26.10)$ |
| Stratum 1 | Amiodarone | 1.116 | 1.341 | 1.724 | -0.126 + $0.077G(16.23)$ |
| Stratum 2 | Placebo | 0.659 | 0.448 | 0.314 | -0.057 + $0.019G(37.20)$ |
| Stratum 2 | Amiodarone | 0.498 | 0.256 | 0.136 | -0.047 + $0.015G(36.65)$ |

Let $\varphi_{p,j}$ and $\varphi_{t,j}$ denote the respective cardiac mortality rates in the placebo and treated groups, within each stratum $j$ ($j = 1, 2$). In order to evaluate the treatment and stratum effects we considered the differences between the mortality rates:

Placebo vs Amiodarone: $\delta = \frac{\varphi_{p,1}+\varphi_{p,2}}{2} - \frac{\varphi_{t,1}+\varphi_{t,2}}{2}$

Stratum 2 vs Stratum 1: $\delta = \frac{\varphi_{p,2}+\varphi_{t,2}}{2} - \frac{\varphi_{p,1}+\varphi_{t,1}}{2}$

Furthermore the interaction between treatments and strata was defined as the difference of differences :

Interaction: $\delta = (\varphi_{p,2} - \varphi_{t,2}) - (\varphi_{p,1} - \varphi_{t,1})$

Results are summarized in Table 2, both for the three-month and the two-year cardiac mortality. They showed a clear-cut effect of the strata, which confirmed the appropriateness of the protocol design, and no evidence of overall treatment effect. The most outstanding finding was the interaction at three months, where the treatment effect could consist of an early mortality reduction in stratum 1. These results suggested to consider further separate analyses within each stratum.

**Table 2. Interim analysis with a noninformative prior: characteristics of the posterior distributions for $\delta$**

| | | $\mathbb{E}(\delta)$ | $Pr(\delta > 0)$ |
|---|---|---|---|
| Placebo vs Amiodarone | 3 months | +0.014 | 0.871 |
| Stratum 2 vs Stratum 1 | 3 months | +0.032 | 0.996 |
| Interaction | 3 months | +0.051 | 0.979 |
| Placebo vs Amiodarone | 2 years | +0.024 | 0.784 |
| Stratum 2 vs Stratum 1 | 2 years | +0.088 | 0.997 |
| Interaction | 2 years | +0.008 | 0.545 |

## Separate Analysis within Each Stratum

### Stratum 1

For illustrative purposes, we present in detail the analysis within stratum 1 at the time of the interim analysis. Assuming again a noninformative prior, the posterior probability that the ratio $\frac{\beta_{t,1}}{\beta_{p,1}}$ was more than 1.37 was equal to 0.95. This indicated a substantive difference in the curve shapes. The 0.95 credibility intervals were respectively:

$$\varphi_{p,1[3 \text{ months}]} : [0.020 \ , \ 0.077] \qquad \varphi_{p,1[2 \text{ years}]} : [0.072 \ , \ 0.171]$$

$$\varphi_{t,1[3 \text{ months}]} : [0.002 \ , \ 0.025] \qquad \varphi_{t,1[2 \text{ years}]} : [0.041 \ , \ 0.154]$$

More specifically the analysis stated the superiority of amiodarone on the three-month cardiac mortality:

$$Pr(\varphi_{p,1[3 \text{ months}]} > \varphi_{t,1[3 \text{ months}]}) = 0.999$$

$$Pr(\varphi_{p,1[3 \text{ months}]} - \varphi_{t,1[3 \text{ months}]} > 0.018) = 0.95$$

$$Pr(\frac{\varphi_{p,1[3 \text{ months}]}}{\varphi_{t,1[3 \text{ months}]}} > 1.99) = 0.95$$

but not on the two year mortality:

$$Pr(\varphi_{p,1[2 \text{ years}]} > \varphi_{t,1[2 \text{ years}]}) = 0.78$$

Up to this point a noninformative prior was assumed. The sensitivity of conclusions to the choice of the prior should nevertheless be investigated. As an illustration, *skeptical* prior distributions [19] were here of special interest for the shape parameter in the placebo group. Thus we considered for this parameter gamma prior distributions with different means and standard deviations (a gamma distribution with mean $m$ and standard deviation $s$ is $(\frac{s^2}{m})G(\frac{m^2}{s^2})$). For the placebo group, a prior distribution that favored an exponential model (especially a distribution centered around 1) could be considered as skeptical with regard to the conclusion of early mortality. We assumed the means 0.75 and 1, and the standard deviations 0, 0.09 and 0.18. Note that 0.09 was close to the standard deviation of the posterior distribution of the shape parameter $\beta_{p,1}$ for the noninformative prior. Using these priors revealed some interesting features. In particular the choice of the prior clearly influenced the estimation of the three month cardiac mortality rate $\varphi_{p,1[3\text{ months}]}$: the mean of the posterior distribution decreased from 0.050 for the noninformative prior to 0.021 for the exponential model ($m = 1$ and $s = 0$) that "ignored" the early mortality. In the same way the mean of the posterior distribution of the ratio of the shape parameters $\frac{\beta_{t,1}}{\beta_{p,1}}$ decreased from 2.66 to 1.10. Table 3 gives the posterior 95% credibility intervals for these two quantities. These intervals showed that the precision of the estimation also heavily depended on the *a priori* information about the shape parameter: assuming a known value (s=0) substantially improved the inference for a value $m$ close to 1 (exponential model), but had less and less impact when $m$ decreased. To directly investigate the conclusion of clinical superiority of amiodarone at three months, Table 3 also gives the 95% lower credibility limit for the ratio of the mortality rates $\frac{\varphi_{p,1[3\text{ months}]}}{\varphi_{t,1[3\text{ months}]}}$. The greater this limit, the more reliably the clinical superiority of amiodarone could be asserted.

**Table 3. Interim analysis in stratum 1: impact of the prior distribution relative to the shape parameter in the placebo group.**

| | 95% credibility limits for: | | |
|---|---|---|---|
| Prior for the placebo | $\varphi_{p,1[3\text{ months}]}$ | $\frac{\beta_{t,1}}{\beta_{p,1}}$ | $\frac{\varphi_{p,1[3\text{ months}]}}{\varphi_{t,1[3\text{ months}]}}$ (lower) |
| *noninformative* | [0.029 , 0.077] | [1.22 , 4.93] | 1.99 |
| $m = 0.75$ $s = 0.18$ | [0.026 , 0.069] | [1.08 , 3.88] | 1.79 |
| $m = 0.75$ $s = 0.09$ | [0.022 , 0.058] | [0.89 , 2.95] | 1.52 |
| $m = 0.75$ $s = 0$ | [0.019 , 0.046] | [0.77 , 2.35] | 1.29 |
| $m = 1$ $s = 0.18$ | [0.022 , 0.059] | [0.88 , 3.06] | 1.51 |
| $m = 1$ $s = 0.09$ | [0.017 , 0.041] | [0.70 , 2.22] | 1.15 |
| $m = 1$ $s = 0$ | [0.013 , 0.031] | [0.57 , 1.76] | 0.88 |

Such a sensitivity analysis would allow a Monitoring Committee to assess the strength of conclusions induced by the data, especially if the skeptical prior was specified beforehand. In this way, the trial would only stop if the partial data give sufficient evidence to counterbalance it.

From the posterior distribution, we can simulate a predictive distribution and assess the consequence of continuing the trial. We assume solely here the reference noninformative prior, but other priors could be considered as well. Given the intermediate data, 300000 "final" samples of $2 \times 400$ patients were generated by simulating the missing future data. Each sample included the data of the 465 patients available at the time of interim analysis. This concerned the 30 patients for whom the death had been observed and 44 patients censored at the time of interim analysis (20 placebo patients and 24 amiodarone patients). For the remaining 391 patients again under observation, the missing data were simulated to get the planned observation period of two years. In addition, the data of 174 new placebo patients and 161 new amiodarone patients were simulated.

At three months, the numbers of events in each of these samples had means 4.0 (amiodarone) and 23.6 (placebo) and respective standard deviations 3.5 and 1.8. All the samples except two had a mortality rate lower for the amiodarone group, and 99.6% allowed to conclude to the superiority of amiodarone (i.e. the posterior probability that $\varphi_{t,1[3 \text{ months}]}$ was smaller than $\varphi_{p,1[3 \text{ months}]}$ was superior to 0.95). At two years, the numbers of events had respective means 35.4 and 46.6 and standard deviations 10.1 and 7.93. 81.7% samples had a mortality rate lower for the amiodarone group, but only 3.4% allowed to conclude to the superiority of amiodarone. Thus it appeared very likely that the conclusion of early superiority of amiodarone in stratum 1 should be confirmed by the additional data, but unlikely that this superiority should be stated for the two-year mortality.

**Stratum 2**

For stratum 2 the superiority of amiodarone cannot be demonstrated (with a noninformative prior):

$$Pr(\varphi_{p,2[3 \text{ months}]} > \varphi_{t,2[3 \text{ months}]}) = 0.285$$

$$Pr(\varphi_{p,2[2 \text{ years}]} > \varphi_{t,2[2 \text{ years}]}) = 0.663$$

The 95% credibility intervals were $[0.42, 1.60]$ for the ratio $\varphi_{p,2[3\text{ months}]}/\varphi_{t,2[3\text{ months}]}$ and $[0.66, 1.90]$ for $\varphi_{p,2[2\text{ years}]}/\varphi_{t,2[2\text{ years}]}$.

Moreover, given the interim data, there was an about null predictive probability that the final analysis could state the superiority of amiodarone within stratum 2.

## Final Analysis

The study was ended as initially planned. At the final analysis, in stratum 1, 30 events out of 407 patients were observed in the placebo group and 30 events out of 390 patients in the amiodarone group, with respective three-month mortality rates equal to $\frac{16}{407}$ and $\frac{4}{390}$. In stratum 2, 59 events out of 336 patients were observed in the placebo group and 55 events out of 353 patients in the amiodarone group, with respective three-month mortality rates equal to $\frac{27}{336}$ and $\frac{22}{353}$. The survival curves are shown in Figure 3. The posterior distributions for the shape parameters essentially revealed the same features as the interim analysis (see Table 4).

**Table 4. Final analysis with a noninformative prior: moments and approximations of the posterior distributions for the shape parameters $\beta_g$**

| | | $\mathbb{E}(\beta_g)$ | $\mathbb{E}(\beta_g^2)$ | $\mathbb{E}(\beta_g^3)$ | Approximations |
|---|---|---|---|---|---|
| Stratum 1 | Placebo | 0.457 | 0.216 | 0.105 | -0.005 + $0.015G(31.54)$ |
| Stratum 1 | Amiodarone | 1.005 | 1.042 | 1.114 | -0.023 + $0.031G(32.82)$ |
| Stratum 2 | Placebo | 0.639 | 0.414 | 0.273 | -0.013 + $0.010G(65.02)$ |
| Stratum 2 | Amiodarone | 0.484 | 0.238 | 0.119 | -0.008 + $0.008G(59.83)$ |

The interim results were confirmed. Table 5 shows a clear-cut effect of the strata, no evidence of overall treatment effect, and an interaction at three months.

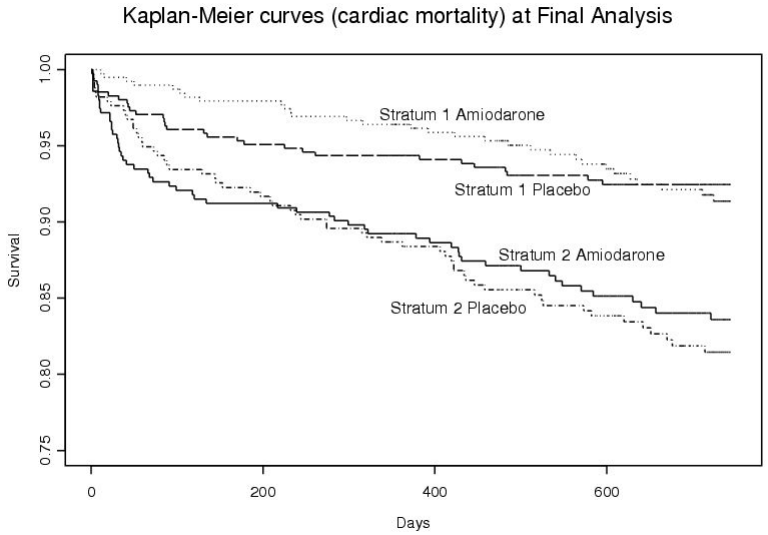


Figure 3: Kaplan-Meier estimates of cardiac mortality by group at final analysis.

**Table 5. Final analysis with a noninformative prior: characteristics of the posterior distributions for $\delta$**

| | | $\mathbb{E}(\delta)$ | $Pr(\delta > 0)$ |
|---|---|---|---|
| Placebo vs Amiodarone | 3 months | +0.006 | 0.726 |
| Stratum 2 vs Stratum 1 | 3 months | +0.035 | 0.9997 |
| Interaction | 3 months | +0.032 | 0.934 |
| Placebo vs Amiodarone | 2 years | +0.007 | 0.663 |
| Stratum 2 vs Stratum 1 | 2 years | +0.093 | $> 0.9999$ |
| Interaction | 2 years | -0.029 | 0.209 |

The separate analyses for each stratum revealed that the results were fully compatible with the predictions, except for the placebo group in stratum 1. In this group, after the time of the interim analysis, only one cardiac death out of 168 new patients had been observed at three months. Nevertheless the posterior probabilities relative to the mortality rates in stratum 1

were fairly compatible with the predictions. They are the following

$$Pr(\varphi_{p,1[3\text{ months}]} > \varphi_{t,1[3\text{ months}]}) = 0.990$$

$$Pr(\varphi_{p,1[3\text{ months}]} - \varphi_{t,1[3\text{ months}]} > 0.006) = 0.95$$

$$Pr(\frac{\varphi_{p,1[3\text{ months}]}}{\varphi_{t,1[3\text{ months}]}} > 1.34) = 0.95$$

$$Pr(\varphi_{p,1[2\text{ years}]} > \varphi_{t,1[2\text{ years}]}) = 0.38$$

If the superiority of amiodarone at three months in stratum 1 could be asserted again, its clinical superiority, that is the existence of a *substantial* decrease of the early mortality was more questionable.

## CONCLUSION

The example above illustrates some interesting features of the Bayesian approach. In this framework, general and flexible procedures for the analysis of complex models are available. As an example, in the case of the comparison of Weibull survival curves, we can easily overcome the assumption that shape parameters have equal values, when it appears grossly unrealistic. Furthermore the sensitivity of the conclusions *vis-à-vis* the value of the shape parameters can be investigated by the means of prior distributions. In an interim analysis, the inference conditional on the available data can be complemented with a predictive inference about the complete planned data. It is of special interest for providing the Safety and Efficacy Monitoring Committees with arguments for or against stopping a study, or even reshaping it, at an early stage.

## References

[1] Abrams, K., Ashby, D., and Errington, D. 'A Bayesian approach to Weibull survival models: application to a cancer clinical trial', *Life Time Data Analysis*, **2**, 159–174 (1996).

[2] Berliner, L.M. and Hill, B.M. 'Bayesian nonparametric survival analysis', *Journal of the American Statistical Association*, **83**, 772–779 (1988).

[3] Spiegelhalter, D.J., Freedman, L.S., and Blackburn, P.R. 'Monitoring clinical trials: Conditional or predictive power?', *Controlled Clinical Trials*, **7**, 8–17 (1986).

[4] Choi, S.C. and Pepple, P.A. 'Monitoring clinical trials based on predictive probability of significance', *Biometrics*, **45**, 317–323 (1989).

[5] Chang, M.N. and Shuster, J.J. 'Interim analysis for randomized clinical trials: simulating the predictive distribution of the log-rank test statistic', *Biometrics*, **50**, 827–833 (1994).

[6] Lecoutre, B., Derzko, G., and Grouin, J.M. 'Bayesian predictive approach for inference about proportions', *Statistics in Medicine*, **14**, 1057–1063 (1995).

[7] Baum, M., Houghton, J., and Abrams, K.R. 'Early stopping rules: clinical perspectives and ethical considerations', *Statistics in Medicine*, **13**, 1459–1469 (1989).

[8] Halperin, M., Lan, K.K.G., Ware, J.H., Johnson, N.J., and Demets, D.L. 'An aid to data monitoring in long-term clinical trials' *Controlled Clin Trials*, **3**, 311–323 (1982).

[9] Jennisson, C. and Turnbull, B.W. 'Statistical approaches to interim monitoring of medical trials: a review and commentary', *Statistical Science*, **5**, 299–317 (1990).

[10] Lan, K.K.G., Simon, R., and Halperin, M. 'Stochastically curtailed tests in long-term clinical trials', *Communications in Statistics - Sequential Analysis*, **1**, 207–219 (1982).

[11] Dignam, J.J., Bryant, J., Wieand, H.S., Fisher, B., and Wolmark, N. 'Early stopping of a clinical trial when there is evidence of no treatment benefit: protocol B-14 of the National Surgical Adjuvant Breast and Bowel Project', *Controlled Clinical Trials*, **19**, 575–588 (1998).

[12] Julian, D.G., Camm, A.J., Frangin, G., et al. 'Randomised trial of effect of amiodarone on mortality in patients with left-ventricular dysfunction after recent myocardial infarction: EMIAT', *Lancet*, **349**, 667–674 (1997).

[13] Sinha, S.K. and Guttman, I. 'Bayesian analysis of life-testing problems involving the Weibull distribution', *Communications in Statistics - Theory and Methods*, **17**, 343–356 (1988).

[14] Nigm, A.M. 'An informative Bayesian prediction for the Weibull lifetime distribution', *Communications in Statistics - Theory and Methods*, **18**, 897–911 (1989).

[15] Geisser, S. Comment about the article of Berliner L.M., Bruce M.H., 'Bayesian nonparametric survival analysis', *Journal of the American Statistical Association*, **83**, 772–881 (1993).

[16] Tziafetas, G.N. 'On the construction of Bayesian predictive limits for the Weibull distribution', *Statistics*, **18**, 623–628 (1987).

[17] Evans, I.G. and Nigm, A.M. 'Bayesian prediction for two-parameter Weibull lifetime models', *Communications in Statistics - Theory and Methods*, **9**, 649–658 (1980).

[18] Box, G.E.P. and Tiao, G.C. *Bayesian Inference in Statistical Analysis*, Addison Wesley, Reading, Massachusetts, 1973.

[19] Spiegelhalter, D.J., Freedman, L.S., and Parmar, M.K.B. 'Bayesian approaches to randomized trials', *Journal of the Royal Statistical Society, Series A*, **157**, 357–416 (1994).