STATISTICS IN MEDICINE
Statist. Med. 2006; **25**:3534–3547
Published online 12 December 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/sim.2445

Modelling menstrual status during and after adjuvant treatment for breast cancer

Suzanne E. Szwarc[1,*,†] and Marco Bonetti[2,‡]

[1]*Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115, U.S.A.*
[2]*Istituto di Metodi Quantitativi, Universitá Bocconi, Milano, 20137, Italy*

SUMMARY

Failure time data may consist of the observation of an event whose cause is unknown due to the censoring or lack of a second event that could identify the cause of the first event. Standard competing-risks methodology does not apply to this setting because the cause of the event is not always identifiable. Moreover, one cannot assume that the entire population will eventually experience the event of interest, and the observation is potentially censored for all patients. The model that we describe in this article is motivated by a breast cancer clinical trial conducted by the International Breast Cancer Study Group (IBCSG). Because some breast cancer adjuvant treatments for premenopausal patients who have undergone surgery cause the interruption of menses, or amenorrhoea, it is of interest to describe the process by which menses discontinue and resume after treatment is completed. The process is complicated by the fact that natural menopause also occurs in the patient population, and that treatment-induced amenorrhoea is not distinguishable from menopause unless menses are observed to resume after treatment completion. We discuss a parametric model for the time to amenorrhoea and for the time to the recovery of menses, also accounting for the presence of censoring and for the possibility that treatment causes an anticipation of natural menopause. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: competing risks; breast cancer; cure-rate model; missing data

*Correspondence to: Suzanne Szwarc, Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115, U.S.A.
†E-mail: sszwarc@hsph.harvard.edu
‡E-mail: marco.bonetti@unibocconi.it

Contract/grant sponsor: Swiss Group for Clinical Cancer Research (SAKK)
Contract/grant sponsor: Frontier Science and Technology Research Foundation
Contract/grant sponsor: Cancer Council, Australia
Contract/grant sponsor: Australian New Zealand Breast Cancer Trials Group
Contract/grant sponsor: United States National Cancer Institute; contract/grant number: CA-75362
Contract/grant sponsor: Swedish Cancer Society

Received 15 April 2005
Accepted 25 September 2005

1. INTRODUCTION

The use of some breast cancer adjuvant treatments for premenopausal women who have undergone surgery may interrupt menses, but little is known about the process by which menses stop and resume post-treatment, or how the process of natural menopause may be affected by these treatments. The demand for this information is growing because nearly a quarter of all patients diagnosed with breast cancer are still in their child-bearing years, and the number of breast cancer diagnoses among young women continues to increase [1]. Recent indications that pregnancy after breast cancer may have a protective effect on the patient also emphasize the necessity for a better understanding of the time to resumption of menses [2]. Because breast cancer survivors have also been shown to experience an increased risk of early menopause after chemotherapy, many women are concerned about the risks of early onset of osteoporosis and heart disease, in addition to hot flashes and sexual dysfunction [3]. Regardless of what information breast cancer patients are seeking before selecting a treatment, data suggest that regardless of age or disease stage, patients worry about the impact that adjuvant treatments could have on their fertility [4], and that the lack of information about fertility loss imposes a negative psychosocial impact on the patient [5].

In this paper we develop a model for the analysis of the time to treatment-induced amenorrhoea (TIA) and of the time to resumption of menses after treatment for patients enrolled in a breast cancer clinical trial in which they receive treatment that may temporarily interrupt menses. The model incorporates parametric assumptions about the distribution of the age at entry and of the age of natural menopause in the potential patient population before the eligibility requirement that the patients be premenopausal is enforced.

For the patients who enter the study our approach is based on age-adjusted cure-rate models for both times to TIA and to recovery of menses [6]. Cure-rate models are a natural choice for describing these events since not all patients experience TIA when undergoing treatment and not all patients who experience TIA recover their menses. Cure-rate models have become popular for analysing data from clinical trials because sometimes not all patients are susceptible to the outcome of interest, even if complete follow-up is possible. Therefore, it is of interest to estimate the proportion of patients who do not experience an event after extended follow-up and the latency of those patients that are not cured.

Several methods have been developed for the cure-rate model in its original formulation. Testing for differences in cure rates has been addressed by Gray and Tsiatis [7], who developed a test that focuses power against late differences in the survival distributions. Laska and Meisner [8] developed a non-parametric test for the equality of the cure rates, and Chen *et al*. [9] developed a Bayesian cure model that incorporates a proportional hazards structure to the model. Peng *et al*. [6] proposed a generalized F distribution for the latency, and Farewell [10] proposed using the logit link to model the cure proportion.

In the setting that we discuss here, the quantities of interest are observed only partially, both because of censoring (due to administrative reasons or by other events) and because it is not possible to distinguish between a temporary cessation of menses and the onset of natural menopause, unless menses are observed to resume after the cessation. These complications require the careful specification of the likelihood contribution associated with each of the several possible observed data configurations. The likelihood function is then maximized to obtain parameter estimates which can then be used to derive model-based quantities of interest. It should be noted that this is the first attempt at modelling menopause and TIA as *separate*

and clearly defined age-dependent outcomes for cancer patients. The model also incorporates the possibility of an acceleration of time to natural menopause due to chemotherapy treatment. In Section 2 we describe the model in detail. In Section 3 we fit the model to data arising from a clinical trial conducted by the International Breast Cancer Study Group (IBCSG), which motivated this work. We close with some discussion in Section 4.

2. THE MODEL

Assume that Z is a random variable describing the age at study entry of breast cancer and that M denotes the potential age at natural menopause of patients in the study. Suppose C is the censoring time from randomization, and $Txend$ represents the time from randomization to treatment end. Also, assume that A_1 is the random variable for time from randomization to TIA, and A_2 is the random variable for time from treatment end to recovery of menses.

The patient population is premenopausal, and thus there is a dependence between age at entry (Z) and potential age of menopause (M) among patients in the study, since it is required that $Z<M$.

We assume that before selection into the study, Z follows a normal distribution with parameters (μ_Z, σ_Z^2) with density f_Z given by $\phi_Z(\cdot)$, the normal density. Further, we assume that the joint distribution of M and Z, $f_{M,Z}(m,z)$, is bivariate normal, with support $\{Z<M\}$ for the patients who are enrolled in the study. We assume that Z and M are independent before enrolment in the study. Defining $I(z<m)$ to be the indicator that patients are premenopausal (1) or not (0), for patients who enter the study we therefore have

$$f_{M,Z}(m,z \mid Z<M) = \frac{f_{M,Z}(m,z)I(z<m)}{P(Z<M)}$$

Integrating over the possible values for M, the distribution of age for patients eligible for the study is

$$f_Z(z \mid Z<M) = \frac{\phi_Z(z)[1-\Phi_M(z)]}{\Phi_{Z-M}(0)}$$

Note the random variable $Z-M$ is a normal random variable with mean equal to $\mu_Z-\mu_M$ and variance equal to $\sigma_Z^2+\sigma_M^2$. One can derive $f_M(m \mid Z<M)$ in a similar manner.

It can be shown that the CDF for M given that $Z=z$ and that a patient is enroled in the study is

$$F_M(m \mid Z=z, Z<M) = \frac{\Phi_M(m)-\Phi_M(z)}{1-\Phi_M(z)}, \quad z \leqslant m$$

We assume that the distribution of the time from randomization to TIA, A_1, follows a cure-rate model. That is, we assume that it is equal to $+\infty$ with probability $(1-\alpha)$ and that it has density $f_{A_1}(a_1)$ with probability α, where $f_{A_1}(\cdot)$ has support $[0, Txend]$. In particular, we assume that, conditionally on being finite, A_1 follows a truncated Weibull distribution with parameters γ_1 and c_1, with γ_1 and $c_1>0$:

$$P(A_1 \leqslant a_1 \mid Z<M) = \begin{cases} \dfrac{1-\mathrm{e}^{-(a_1/\gamma_1)^{c_1}}}{1-\mathrm{e}^{-(Txend/\gamma_1)^{c_1}}} & 0 \leqslant a_1 \leqslant Txend \\ 1 & a_1 > Txend \end{cases}$$

Lastly, assume that the distribution of the time from end of treatment to recovery of menses, A_2, depends on the value of A_1 only through the indicator function $I(A_1 < +\infty)$, and that it is equal to 0 if $A_1 = +\infty$ with probability 1, since resumption is not defined if TIA did not occur. If $A_1 < +\infty$ (and therefore $A_1 < Txend$), then A_2 is assumed to also follow a cure-rate model, with probability $(1-\beta)$ of being equal to $+\infty$, and (if A_2 is finite) with density f_{A_2} where f_{A_2} is also a Weibull distribution, with parameters γ_2 and c_2, also positive.

Note that the occurrence of TIA is indistinguishable from the occurrence of natural menopause, so the occurrence of the former can only be confirmed if resumption of menses is observed.

The observed data for a patient of age z can be written as $(\mathbf{X}, \delta, z)$, with $\mathbf{X} = (X_1, X_2, X_3)$, where X_1 represents the observed time from study entry to a cessation of menses before treatment end, X_2 represents the time from treatment end to a recovery of menses before C, X_3 represents the time from treatment end to a cessation of menses before C. If an event is not observed, the corresponding random variable is set equal to $+\infty$. The variable δ indicates one of the seven possible observed configurations of X_1, X_2, and X_3, shown in Figure 1, where the y-axis indicates whether a cessation of menses is occurring (1) or not (0).

The first type of menses pattern is indicated by $\delta = 1$, and the corresponding observed data can be written as $(X_1, +\infty, +\infty)$; this configuration only contains the information that the value at X_1 could either be TIA from which the patient does not recover prior to C, or permanent menopause masked by the treatment effect.

When we observe $(X_1, X_2, +\infty)$, the assigned indicator for δ is 2. Because a patient with this configuration has a recovery of menses after treatment end, one knows that the cessation of menses that was observed at X_1 is treatment induced, rather than natural menopause.

The observation of (X_1, X_2, X_3) is equivalent to having observed all three events prior to censoring. A patient enters the study, achieves TIA, recovers her menses, and then experiences a second cessation which must therefore be natural menopause. This pattern is represented by $\delta = 3$, and it conveys precise information about A_1, A_2, and M.

When a patient does not experience TIA, and menopause is not observed during the follow-up after treatment end until C, the observed information is written as $(+\infty, +\infty, +\infty)$. We denote this situation by $\delta = 4$.

If a patient's menses are observed to stop for the first time between the end of treatment and the censoring time, the cessation can no longer be treatment induced, and it is therefore attributed to natural menopause, M. The observed data for that situation, $(+\infty, +\infty, X_3)$, is represented by $\delta = 5$.

Note that it seems reasonable to censor a patient's information if they experienced an event that is likely to influence their menses. These 'intervening events' may include hysterectomy, oophorectomy, disease recurrence, death, pregnancy, and undergoing systemic treatment that could influence menses. In general, censoring due to one of these events (denoted by C) may occur before treatment end. Two additional observed configurations must therefore be added, as shown in Figure 1.

For those patients who are censored before treatment end and who never experience TIA during that time, the observed data configuration is indicated by $\delta = 6$, and it is represented by $(+\infty, +\infty, +\infty)$. This situation is very similar to the situation when $\delta = 4$.

Also, patients may experience a cessation of menses with censoring after X_1 but before treatment end. Their configuration is denoted by $\delta = 7$ and their observed information is written as $(X_1, +\infty, +\infty)$. This scenario is similar to the pattern labelled as $\delta = 1$.

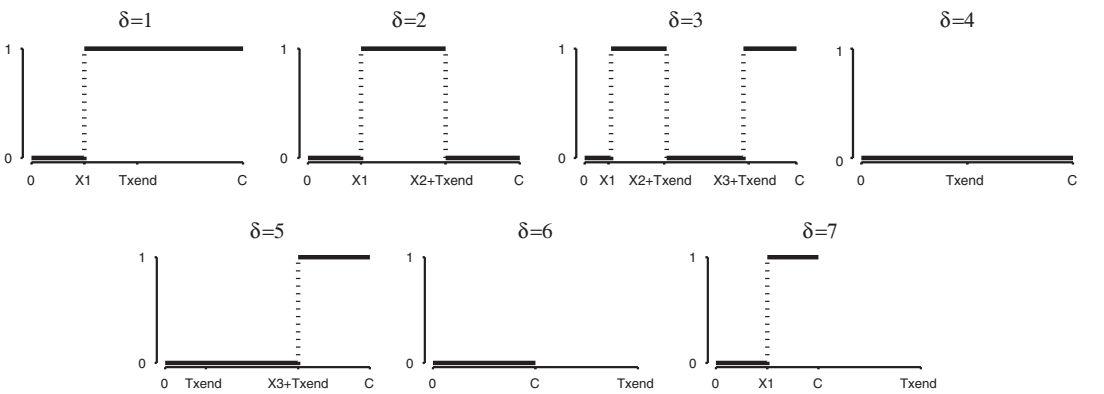

Figure 1. Possible observed data patterns, with associated value of the variable δ. The x-axis shows the time from randomization, while the y-axis represents the indicator of whether a cessation of menses is occurring (1) or not (0).

It should be noted that not all of the first five settings described above are affected by early censoring. When δ is equal to 3 or 5, the cessation the occurs at X_3 is defined to be natural menopause, an event after which recovery is not possible. Early censoring when $\delta = 2$ only censors information on the age at menopause.

The cure rate for the distributions of time to TIA (A_1) and time to recovery of menses (A_2) are expected to change greatly with the age at entry of the patient. These age effects can be added into the model, for example, via the logit link. In the application described in Section 3 we will treat age as a continuous random variable for the estimation of α, but we will categorize it for the estimation of β into the intervals $(0, 40), [40, 45), [45, +\infty)$. Note that the reference group here is $[45, +\infty)$. Thus,

$$\alpha(z) = \frac{e^{\alpha_1+\alpha_2 z}}{1+e^{\alpha_1+\alpha_2 z}} \quad \text{and} \quad \beta(z) = \frac{e^{\beta_1+\beta_2 I(z<40)+\beta_3 I(40\leqslant z<45)}}{1+e^{\beta_1+\beta_2 I(z<40)+\beta_3 I(40\leqslant z<45)}}$$

Evidence that some chemotherapeutic agents (such as the CMF regimen in the IBCSG trial) may cause early menopause in premenopausal breast cancer patients suggests, including in the model, the possibility of a relationship between age at study entry and natural menopause. So far, we have assumed that prior to entry in the study, $M \sim \mathrm{N}(\mu_M, \sigma_M^2)$ and that for a patient entering the study at age z, the age at natural menopause is unchanged by the treatment. We can enlarge the model by introducing a new random variable, M_{PR}, to denote post-randomization age at menopause.

We can define $M_{\mathrm{PR}}(Z, M) = Z + k(M - Z)$ for $k \in (0, 1]$. In general, we expect $z < m_{\mathrm{PR}} \leqslant m$, and if there is no effect of treatment on age at menopause, then conditionally on being premenopausal when entering the study at age $Z = z$, the two random variables (M and M_{PR}) are identical, or $k = 1$.

Because of the definition of M, the distribution of M_{PR} can be easily obtained as

$$f_{M_{\mathrm{PR}}}(m_{\mathrm{PR}} \mid Z = z, M, Z < M) = \frac{1}{k} f_M\left(\frac{m_{\mathrm{PR}} - (1-k)z}{k} \,\middle|\, Z = z, Z < M\right), \quad m_{\mathrm{PR}} \geqslant z$$

After careful consideration of all the events that produce each of the data configurations, each individual's contribution to the likelihood, $f(\mathbf{x};\delta)$, can be written as one of the following, depending on her observed data configuration:

$$
\begin{aligned}
f(\mathbf{x};1) &= f_{M_{\mathrm{PR}}}(x_1+z\,|\,Z=z,Z<M)\{\alpha(z)[F_{A_1}(Txend)-F_{A_1}(x_1)]+(1-\alpha(z))\}\\
&\quad+\alpha(z)f_{A_1}(x_1)\{[F_{M_{\mathrm{PR}}}(Txend+z\,|\,Z=z,Z<M)-F_{M_{\mathrm{PR}}}(x_1+z\,|\,Z=z,Z<M)]\\
&\quad+[1-F_{M_{\mathrm{PR}}}(Txend+z\,|\,Z=z,Z<M)][1-\beta(z)F_{A_2}(c-Txend)]+\beta(z)V\}\\
f(\mathbf{x};2) &= \alpha(z)f_{A_1}(x_1)\beta(z)f_{A_2}(x_2)[1-F_{M_{\mathrm{PR}}}(c+z\,|\,Z=z,Z<M)]\\
f(\mathbf{x};3) &= \alpha(z)f_{A_1}(x_1)\beta(z)f_{A_2}(x_2)f_{M_{\mathrm{PR}}}(x_3+z+Txend\,|\,Z=z,Z<M)\\
f(\mathbf{x};4) &= (1-\alpha(z))[1-F_{M_{\mathrm{PR}}}(c+z\,|\,Z=z,Z<M)]\\
f(\mathbf{x};5) &= (1-\alpha(z))f_{M_{\mathrm{PR}}}(x_3+z+Txend\,|\,Z=z,Z<M)\\
f(\mathbf{x};6) &= [\alpha(z)(1-F_{A_1}(c))+(1-\alpha(z))][1-F_{M_{\mathrm{PR}}}(c+z\,|\,Z=z,Z<M)]\\
f(\mathbf{x};7) &= \alpha(z)f_{A_1}(x_1)[1-F_{M_{\mathrm{PR}}}(x_1+z\,|\,Z=z,Z<M)]\\
&\quad+f_{M_{\mathrm{PR}}}(x_1+z\,|\,Z=z,Z<M)[\alpha(z)(1-F_{A_1}(x_1))+(1-\alpha(z))]
\end{aligned}
$$

where

$$
\begin{aligned}
V &= \int_{kTxend+z(2-k)}^{kc+z(2-k)} f_M(m\,|\,Z=z,Z<M)\mathrm{e}^{-(m-z(2-k)-kTxend/k\gamma_2)^{c_2}}\,\mathrm{d}m\\
&\quad-\mathrm{e}^{-(C-Txend/\gamma_2)^{c_2}}[F_M(kC+z(2-k)\,|\,Z=z,Z<M)-F_M(kTxend+z(2-k)\,|\,Z=z,Z<M)]
\end{aligned}
$$

For a given treatment arm the parameter $\theta=(\mu_M,\sigma_M^2,\mu_Z,\sigma_Z^2,\alpha_1,\alpha_2,\beta_1,\beta_2,\beta_3,\gamma_1,c_1,\gamma_2,c_2,k)$ can be estimated from the observed data on the n patients ($\mathscr{X}=\{(\mathbf{X}_i,\delta_i,z_i),\ i=1,\ldots,n\}$) by maximizing the likelihood:

$$
L(\theta\,|\,\mathscr{X})=\prod_{i=1}^{n}\prod_{j=1}^{7}f(\mathbf{x}_i;j)^{I(\delta_i=j)}f_Z(z_i\,|\,Z<M)
$$

where $I(\delta_i=j)$ is defined as the indicator that δ_i is equal to j, $j=1,\ldots,7$ for the ith patient. The integral in $f(\mathbf{x};1)$ can be calculated using a Monte Carlo approximation, or using numerical algorithms [11]. In our experience, the two approaches produced very similar results.

The variance–covariance matrix of the maximum likelihood estimates can then be obtained by inverting the estimated Hessian matrix, for example, by using the OPTIM function in *R* [12].

We now describe the implementation of the full model to the IBCSG Trial VIII, which motivated this work.

3. APPLICATION: THE INTERNATIONAL BREAST CANCER STUDY GROUP (IBCSG) TRIAL VIII

The International Breast Cancer Study Group (IBCSG) Trial VIII was designed to study how premenopausal node-negative breast cancer patients respond to combination therapy. Patients were randomized to four different treatment arms and stratified according to oestrogen-receptor (OR) status, whether radiotherapy was planned after surgery, and by the institution within which they had been randomized. Arm A patients received no adjuvant therapy; Arm B patients received goserelin injections for 24 months; Arm C patients received CMF (chemotherapy) for 6 months; Arm D patients received a combination of 6 months of CMF followed by 18 months of goserelin. There were 964 eligible patients available for our analyses, divided into 42 patients on the no-treatment arm, 304 patients on the goserelin arm, 316 on the CMF arm, and 302 patients on the CMF+goserelin arm. Median follow-up was 7 years. For a detailed description of the study and its clinical findings, we refer the reader to International Breast Cancer Study Group [13].

Figure 2 (left) displays a histogram of the ages at entry for Trial VIII. The mean age at entry was 44.3 years and the variance of age at entry was 31.4. The median age at entry was 45.2 years. The histogram shows clearly the effect of the eligibility requirement (that patients be premenopausal, or $Z < M$) on the distribution of the age at entry of the patients in the study.

As part of the study, monthly menses data were carefully recorded during the first 36 months, and every 6 months after that point. These data were used to determine which of the seven possible observable data configurations described in Section 2 occurred for each patient. A patient's monthly menses information was defined as normal, scanty, none, censored (by an intervening event or cessation of follow-up), or missing (no response). These data were collected very carefully, and missing menses information was minimal (2.4 per cent excluding missingness due to censoring).

Within the first 36 months of follow-up, we defined a cessation of menses as the observation of no menses, followed by a month of scanty/none/no response/censored menses, and then a month of none/no response/censored menses. We defined a recurrence of menses as the observation of normal menses, followed by any 2-month pattern that did not report 'none'.

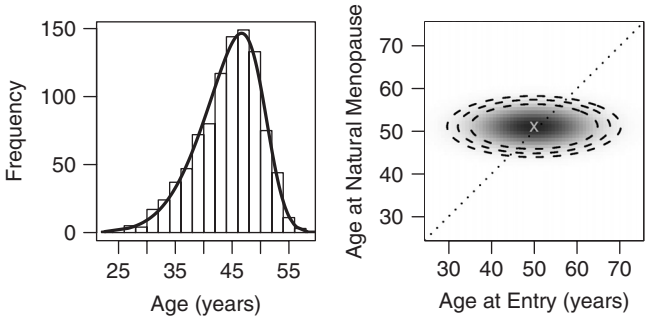

Figure 2. Histogram of ages at entry Z in the population eligible for study entry (left) and fitted bivariate normal for age at entry and potential age at menopause M (right). The solid line in the left panel represents the graph of the estimated density $f_Z(z \mid Z < M)$. The dotted line in the right panel represents the line $Z = M$. The dashed lines correspond to the 95, 90, and 80 per cent confidence ellipses for the joint distribution of Z and M.

Table I. Observed data configurations for Trial VIII.

	δ value							
Treatment arm	1	2	3	4	5	6	7	Total
No treatment	—	—	—	16	26	—	—	42
Goserelin	105	86	77	0	0	2	34	304
CMF	248	10	11	21	19	6	1	316
CMF+goserelin	227	26	20	2	2	5	20	302
Total	580	122	108	39	47	13	55	964

Beyond 36 months from randomization, we defined a cessation as the first time point at which there was an observation 'no menses' followed by no reports of normal menses. This definition does suffer from dependence on each patient's follow-up, but it seemed the most reasonable to clinical experts. Similarly, beyond 36 months of follow-up, a recovery was defined as the first time point at which normal and regular menses were reported. Because menses information collected after 36 months was only collected every 6 months, we decided to rely on only one time point to define a recovery to maintain a similar definition to that used for the events defined before 36 months. The forms asked the patient to summarize menses from 3 months prior to the date of the assessment, so the actual *date* we used for each event was 45 days prior to the date of the assessment.

Treatment end was defined to be equal to 0 for patients on the no-treatment arm, and equal to the end of the clinical administration time plus 1 month for those patients receiving treatment. This choice was made because it is conceivable that a cessation of menses in the first month after the end of the treatment period could be caused by lingering effects of treatment. Also, the timing of the monthly assessment was not perfect, and when taken in combination with the timing of the menstrual cycle it warranted being a bit conservative.

Censoring time, C, was defined to be the last date at which a patient reported data to the IBCSG, unless the patient experienced one of the non-protocol intervening events, in which case C was set to the date of that event. A total of 306 patients experienced one of the intervening events. These patients include 21 on the no-treatment arm, 111 on the goserelin arm, 93 on the CMF arm, and 81 on the CMF+goserelin arm.

In the spirit of an intent-to-treat analysis we considered all patients as having received the assigned treatment duration in our analysis, regardless of actual compliance.

We then defined each patient's observed data configuration by observing where the events first cessation, recovery, and second cessation were observed. A summary of the resulting configurations is given in Table I, by treatment arm.

3.1. Estimation

To allow for a more precise estimation of the parameters μ_M, σ_M, μ_Z and σ_Z, a separate likelihood model was fitted to obtain estimates from the marginal distribution of age across all treatment arms conditional on study entry, $f_Z(z \mid Z < M)$. We inserted these global estimates into the observed data likelihood and maximized it with respect to the remaining parameters. We obtained the unadjusted variance estimates for these parameters from the Hessian matrix, which was calculated using the OPTIM function in *R*. The global estimates and standard

errors (SE) obtained were 51.02 (0.92) for μ_M, 8.44 (1.86) for σ_M^2, 49.90 (2.44) for μ_Z, and 67.65 (13.82) for σ_Z^2.

Figure 2 (right) displays the estimated bivariate normal distribution for Z and M prior to entry into the study. The subset of the curve above the line $Z = M$ represents patients who are eligible to enter the study. The left panel in the same figure displays the graph of the estimated density $f_Z(z \mid Z < M)$. Recalling that M is a measure of the pre-randomization age at menopause, it should be noted that the obtained estimate of the mean age at menopause of 51.02 years is very consistent with other estimates as those found in Reference [14].

We now discuss a sequential estimation procedure in which the model was fitted separately for each treatment arm. The purpose of this was to keep the number of parameters within a reasonable number to facilitate the estimation procedure.

The full model was first fitted to the goserelin-only arm. Given the almost complete success of the treatment in inducing TIA, the model was fitted holding $\alpha(z)$ fixed equal to 1 for all patients, and minimizing (after transformation by natural logarithm) the Weibull parameters in the distributions for A_1 and A_2. The model uses the global parameter estimates for the distribution of Z in the maximization. Since no acceleration of the time to menopause is expected in the premenopausal patients on this treatment arm, we fitted the model with k fixed equal to 1. (We did also fit the accelerated model to this arm as a sensitivity analysis. The estimated value of k was 1.02 with a standard error of 0.0072, confirming the lack of acceleration.) The parameter estimates (reported in Table II) for this arm did not vary much when the accelerated menopause component of the model was added, further confirming the goodness of fit of the model with no acceleration. The obtained estimate of the potential mean age at menopause for this treatment was 50.07, which seems reasonable given the global estimate obtained across all arms. This was also done as a sensitivity analysis, taking

Table II. Results of the model fitting, by treatment arm. (Refer to text for definitions.) The symbol '—' indicates a parameter value that was held fixed in the likelihood maximization. For the goserelin arm, k and α were both fixed equal to 1. For the CMF and CMF+goserelin arms, the parameters for the distributions of M (potential age at menopause) and Z (age at study entry) were fixed equal to the global estimates obtained from the marginal distribution of age across all treatment arms, conditional on study entry. The model was fitted separately for each treatment arm.

	Parameter estimates and standard errors (SE)		
Parameter	Goserelin ($n = 304$)	CMF ($n = 316$)	CMF+goserelin ($n = 302$)
μ_M	50.07 (0.02)	—	—
σ_M^2	17.92 (0.05)	—	—
α_1	—	−10.19 (1.74)	−4.90 (3.59)
α_2	—	0.29 (0.04)	0.23 (0.10)
β_1	1.17 (0.38)	−3.65 (0.59)	−3.75 (1.02)
β_2	2.77 (1.38)	3.97 (0.74)	4.06 (1.06)
β_3	1.19 (0.82)	0.66 (0.78)	2.39 (1.07)
$\log(c_1)$	0.99 (0.04)	1.21 (0.06)	0.85 (0.05)
$\log(\gamma_1)$	−1.32 (0.02)	−0.94 (0.02)	−0.79 (0.03)
$\log(c_2)$	0.08 (0.06)	−0.09 (0.18)	0.09 (0.11)
$\log(\gamma_2)$	−0.47 (0.08)	−0.03 (0.29)	−0.09 (0.16)
k	—	0.75 (0.01)	0.69 (0.01)

Table III. Estimated conditional probabilities of potential recovery and associated 95 per cent confidence intervals, by age group and treatment arm.

Age at entry (years)	Estimated probability of recovery (95 per cent CI)		
	Goserelin ($n=304$)	CMF ($n=316$)	CMF+goserelin ($n=302$)
0–39	0.98 (0.79, 0.99)	0.58 (0.36, 0.77)	0.58 (0.44, 0.71)
40–44	0.91 (0.71, 0.98)	0.05 (0.02, 0.12)	0.20 (0.12, 0.32)
$\geqslant 45$	0.76 (0.61, 0.87)	0.03 (0.01, 0.08)	0.02 (0.003, 0.15)

advantage of the smaller number of parameters for this arm; because there were no parameters for $\alpha(z)$, it was easier to estimate the parameters from the distribution for M from the data.

Fitting the model separately to the CMF and CMF/goserelin arms produced the results shown in Table II. For these arms we used the global parameter estimates for the distributions of Z and M in the maximization. Estimates of $\alpha(z)$ were obtained using continuous age, but $\beta(z)$ was fitted using categorical age, as described in Section 2. Note how for these two arms the acceleration parameter k was estimated as 0.75 and 0.69, respectively, indicating the presence of an acceleration of menopause.

Estimates of the parameters for the cure rate for TIA illustrate a direct relationship between age and the probability of TIA within the CMF and CMF+goserelin arms. A graphical display of $\alpha(z)$ *versus* z shows a higher probability of TIA in the combination therapy arm than in the CMF arm until age 45, at which point both curves plateau at 1. Therefore, there is a higher proportion of 'cured' patients who do not experience TIA on the CMF arm. Table III shows the inverse relationship existing between the probability of potential recovery (β) from TIA and age for the three treatment arms. Older patients have a lower probability of recovery than younger patients, and patients on the goserelin arm have the highest chance of recovery, in general. Plots (not shown) of the estimated density for A_1 for each treatment arm show that the goserelin arm induces TIA earlier than the other two arms. Additionally, plots (not shown) of the estimated densities of A_2 for each arm are similar in that most patients will recover within a year after treatment end if they experience TIA and are going to recover at all.

The goodness of fit of the models for each treatment regimen was explored by deriving the density functions for X_1, X_2, and X_3 using the parameter estimates obtained, and comparing the plots of the densities to the histograms of the observed values of the random variables.

We explored the sensitivity of our results to the definitions of the events of interest during the first 36 months of observation. We modified the definitions first by more strictly redefining a cessation to allow for only no menses (instead of scanty, none, nonresponse, or censored) in the second month, and then by redefining a recovery to allow for only normal menses (instead of normal, scanty, nonresponse, or censored) in the second and third months. The resulting menses patterns did not change much except for a shift of 6 patients in the goserelin arm under the new definition of a recovery. We checked the results by fitting the model for these new data, and the estimates were very similar to those presented above.

3.2. Further extensions

It is of great clinical interest to estimate additional model-based quantities using the parameter estimates from the full model. One of these quantities is the time of the last cessation of

menses either due to menopause or due to TIA, but *never* followed by a recovery, given age and measured from start of treatment. Although one could derive Kaplan–Meier plots for the time to cessation of menses due to any cause, these curves would not be adjusted for age, nor would they allow for the possibility that patients may recover from cessations caused by TIA.

To account for both these issues, we can derive a formula for the probability that the final cessation from which a patient will never recover, denoted by W, occurs beyond a particular time from start of treatment, w. It is possible to construct plots of $P(W > w \mid Z = z)$ *versus* w using the expression

$$\begin{aligned} P(W > w \mid Z = z) &= P(A_1 = +\infty) f_{M_{PR}}(w + z \mid Z = z, Z < M) \\ &+ P(A_1 \in (0, Txend)) P(A_2 \in (0, w - Txend) \mid A_1 < +\infty) f_{M_{PR}}(w + z \mid Z = z, Z < M) \\ &+ P(A_1 > w) f_{M_{PR}}(w + z) + \alpha P(M_{PR} > w + z) f_{A_1}(w) \end{aligned}$$

These plots (shown in Figure 3) highlight this relationship for each treatment arm. The plots are relevant to physicians and patients making treatment decisions since they quantify the probability that a woman will not experience a *permanent* cessation of menses (due to treatment or menopause) at time w from start of treatment, as a function of her age at entry and of the treatment she will receive. One can see from Figure 3 that for all therapies, younger patients experience a much higher probability of any recovery of menses, but the rate at which this probability decreases is much faster for patients on the CMF and the CMF+goserelin arms.

4. DISCUSSION

In this paper, we have proposed a parametric model for the time to event when the type of event that one observes is sometimes uncertain. Our motivating application deals with the times from randomization to menses cessation and recovery in breast cancer patients in clinical trials in which the premenopausal patients receive treatment that may temporarily interrupt menses.

The part of the model that allows for the anticipation of the time to natural menopause appears to fit the data from the CMF and the CMF+goserelin arms well. This is in agreement with recent evidence that suggest that these therapies may cause more permanent damage to ovarian function due to the nature of the CMF regimen [3]. On the other hand, there is little evidence suggesting that goserelin accelerates or delays menopause in any way, and this is also confirmed by our analysis.

It should be noted that our accelerated model assumes for simplicity that the acceleration of menopause is instantaneous at the start of treatment, rather than a cumulative effect over the duration of treatment. Also, note how in our analyses we have followed an intent-to-treat approach, without worrying about compliance issues, and that taking them into account would greatly complicate the problem and the interpretation of the results.

The model that we have described may be sensitive to deviations from the hypothesis of independence between Z and M. Clinical and epidemiological evidence suggests that the age

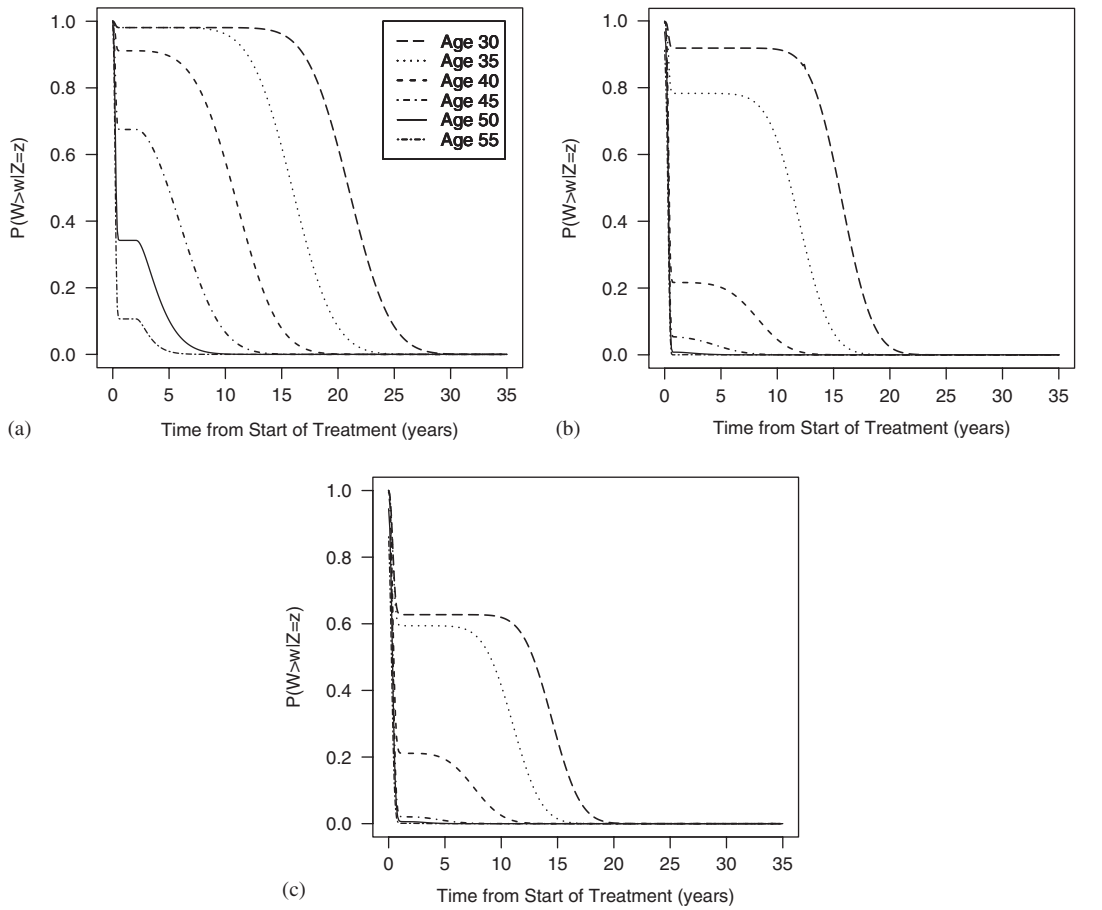

Figure 3. Plots of the function $P(W > w \mid Z = z)$ *versus* the time w from start of treatment for the goserelin (a), the CMF (b), and the CMF+goserelin (c) arms, for various ages at entry (Z) into the study. The quantity W is the time from start of treatment of the last cessation of menses due to menopause or due to TIA not followed by a recovery, given age.

of natural menopause has not varied greatly over time with respect to date of birth, and this would support independence of Z and M in the general population of women [15]. Here, we have considered women who are alive with breast cancer, and therefore additional selection processes exist which may create dependence between Z and M. If one believes, for example, that M tends to increase in this population, then later ages at menopause would occur for younger patients. The correlation between Z and M could be added to the joint distribution of the two variables, but it appears to be hardly identified from the marginal distribution of $Z \mid Z < M$. The model-based estimate for this marginal distribution provides a good fit to the observed IBCSG data, but that marginal distribution could also be consistent with bivariate models that do include a non-extreme correlation. What seems to matter most is the possible influence of such a correlation on the estimation of the remaining parameters in the model. To

explore this in a sensitivity analysis, we have generated data from a distribution that did allow for some correlation between the two variables, and have estimated the remaining parameters from the generated data. The model consisted of exponential distributions for TIA and time to recovery, and age was used as a continuous covariate. We used the parameter values close to the ones estimated from the IBCSG data. The results of this small simulation study indicate only a minor bias in the estimation of the model's parameters when the correlation takes on values between -0.2 and 0.2 (results not shown).

The amount of missing data can make the likelihood maximization difficult. Also, the presence of large amounts of missing data in reports of menstrual status may result in biased estimates for all components of the process, especially if the data is missing in a non-ignorable way, i.e. if the probability that an observation is missing is also a function of unobservable quantities.

The model presented here is unique to these types of data because of the clear differentiation between the processes of treatment-induced amenorrhoea and natural menopause. Should TIA be distinguishable from menopause, competing-risks methods could be applied to the study of the time to the first of the two events [16]. In that case, another possible approach could also be based on multi-state models [17].

The treatment of censoring in these data requires some comments. The data include 35 patients who underwent a hysterectomy or oophorectomy, and 9 patients who became pregnant during the observation period. These patients are likely to suffer from informative censoring, since women censored for these events were likely to be experiencing menses right before being censored. The censoring mechanism for these patients may therefore be a function of the outcome that one is trying to measure (TIA). Because these were less than 5 per cent of the total number of patients in our application, we censored them as we would have censored any patient experiencing one of the other intervening events, i.e. we assumed that censoring was independent. In the situation where more patients are subject to dependent censoring, one might need to complicate the model by explicitly modelling the distribution of censoring [18].

A relevant extension of the present work will include modelling the effect of the menses patterns on the disease outcome. Under the assumption that the underlying reason (TIA *versus* natural menopause) for menses cessation is not important for outcome (so that what matters is the cessation of menses), then existing methods for the joint modelling of longitudinal measures and survival, given covariates such as oestrogen receptor status, could be applied. Clinical applications of such an extension could be the study of early menopause due to treatment of paediatric cancers, and the study of amenorrhoea in patients with anorexia nervosa.

ACKNOWLEDGEMENTS

We thank the patients, physicians, nurses, and data managers who participated in the International Breast Cancer Study Group (IBCSG) Trial VIII. We further acknowledge the IBCSG funding provided by the Swiss Group for Clinical Cancer Research (SAKK), the Frontier Science and Technology Research Foundation, The Cancer Council Australia, the Australian New Zealand Breast Cancer Trials Group, the United States National Cancer Institute (CA-75362), and the Swedish Cancer Society. We also thank Jodi Flaws, Rich Gelber, Shari Gelber, Bob Gray, Anne O'Neill, Ann Partridge, Karen Price, and Molin Wang for their assistance throughout.

REFERENCES

1. Goodwin PJ, Ennis M, Pritchard KI, Trudeau M, Hood N. Risk of menopause during the first year after breast cancer. *Journal of Clinical Oncology* 1999; **17**:2365–2370.
2. Gelber S, Coates AS, Goldhirsch A, Castiglione-Gertsch M, Marini G, Lindtner J, Edelmann DZ, Gudgeon A, Harvey V, Gelber RD. Effect of pregnancy on overall survival after the diagnosis of early-stage breast cancer. *Journal of Clinical Oncology* 2001; **19**:1671–1675.
3. deHaes H, Olschewski M, Kaufmann M, Schumacher M, Jonat W, Sauerbrei W. Quality of life in goserelin-treated versus cyclophosphamide + methotrexate + flourouracil-treated premenopausal and perimenopausal patients with node-positive, early breast cancer: the zoladex early breast cancer research association trialists group. *Journal of Clinical Oncology* 2003; **21**:4510–4516.
4. Partridge AH, Gelber S, Peppercorn J, Sampson E, Knudsen K, Laufer M, Rosenberg R, Przypyszny M, Rein A, Winer EP. Web-based survey of fertility issues in young women with breast cancer. *Journal of Clinical Oncology* 2004; **22**:4174–4183.
5. Thewes B, Meiser B, Rickard J, Friedlander M. The fertility- and menopause-related information needs of younger women with a diagnosis of breast cancer: a qualitative study. *Psycho-Oncology* 2003; **12**:500–511.
6. Peng Y, Dear KBG, Denham JW. A generalized *F* mixture model for cure rate estimation. *Statistics in Medicine* 1998; **17**:813–830.
7. Gray RJ, Tsiatis AA. A linear rank test for use when the main interest is in differences in cure rates. *Biometrics* 1989; **45**:899–904.
8. Laska EM, Meisner MJ. Nonparametric estimation and testing in a cure model. *Biometrics* 1992; **48**:1223–1234.
9. Chen M, Ibrahim JG, Sinha D. A new Bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association* 1999; **94**:909–919.
10. Farewell VT. The use of mixture models for the analysis of survival data with long-term survivors. *Biometrics* 1982; **38**:1041–1046.
11. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical Recipes in C*: *the Art of Scientific Computing* (2nd edn). Cambridge University Press: New York, 1992.
12. Nocedal J, Wright SJ. *Numerical Optimization*. Springer: New York, 1999.
13. International Breast Cancer Study Group. Adjuvant chemotherapy followed by goserelin versus either modality alone for premenopausal lymph node-negative breast cancer: a randomized trial. *Journal of the National Cancer Institute* 2003; **95**:1833–1846.
14. Brambilla DJ, McKinlay SM. A prospective study of factors affecting age at menopause. *Journal of Clinical Epidemiology* 1989; **42**:1031–1039.
15. McKinlay SM. The normal menopause transition: an overview. *Maturitas* 1996; **23**:137–145.
16. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *Journal of the American Statistical Association* 1999; **94**:496–509.
17. Lawless JF. *Statistical Models and Methods for Lifetime Data* (2nd edn). Wiley: New York, 2002.
18. Scharfstein DO, Robins JM. Estimation of the failure time distribution in the presence of informative censoring. *Biometrika* 2002; **89**:617–634.

