A joint model for the dependence between clustered times to tumour progression and deaths: A meta-analysis of chemotherapy in head and neck cancer
Virginie Rondeau, Jean-Pierre Pignon and Stefan Michiels
Stat Methods Med Res published online 23 October 2011
DOI: 10.1177/0962280211425578

The online version of this article can be found at:
http://smm.sagepub.com/content/early/2011/10/22/0962280211425578

Published by:

SAGE
http://www.sagepublications.com

Additional services and information for Statistical Methods in Medical Research can be found at:

Email Alerts: http://smm.sagepub.com/cgi/alerts
Subscriptions: http://smm.sagepub.com/subscriptions
Reprints: http://www.sagepub.com/journalsReprints.nav
Permissions: http://www.sagepub.com/journalsPermissions.nav

>> Proof - Oct 23, 2011

What is This?
A joint model for the dependence between clustered times to tumour progression and deaths: A meta-analysis of chemotherapy in head and neck cancer

Virginie Rondeau,1,2 Jean-Pierre Pignon3 and Stefan Michiels4 on behalf of the MACH-NC collaborative Group

Abstract
The observation of time to tumour progression (TTP) or progression-free survival (PFS) may be terminated by a terminal event. In this context, deaths may be due to tumour progression, and the time to the major failure event (death) may be correlated with the TTP. The usual assumption of independence between the TTP process and death, required by many commonly used statistical methods, can be violated. Furthermore, although the relationship between TTP and time to death is most relevant to the anti-cancer drug development or to evaluation of TTP as a surrogate endpoint, statistical models that try to describe the dependence structure between these two characteristics are not frequently used. We propose a joint frailty model for the analysis of two survival endpoints, TTP and time to death, or PFS and time to death, in the context of data clustering (e.g. at the centre or trial level). This approach allows us to simultaneously evaluate the prognostic effects of covariates on the two survival endpoints, while accounting both for the relationship between the outcomes and for data clustering. We show how a maximum penalized likelihood estimation can be applied to a nonparametric estimation of the continuous hazard functions in a general joint frailty model with right censoring and delayed entry.

The model was motivated by a large meta-analysis of randomized trials for head and neck cancers (Meta-Analysis of Chemotherapy in Head and Neck Cancers), in which the efficacy of chemotherapy on TTP or PFS and overall survival was investigated, as adjunct to surgery or radiotherapy or both.

Keywords
cancer, heterogeneity, joint frailty models, meta-analysis, penalized likelihood, time to tumour progression

1INSERM, CR897 (Biostatistic), Bordeaux, F-33076, France.
2Université Bordeaux Segalen, Bordeaux, F-33076, France.
3Department of Biostatistics and Epidemiology, Institut Gustave-Roussy, Villejuif, F-94805, France.
4Institut Jules Bordet, Université Libre de Bruxelles, Brussels, Belgium.

Corresponding author:
Virginie Rondeau, INSERM CR897 (Biostatistic), Université Bordeaux Segalen, 146 rue Léo Saignat, 33076 Bordeaux Cedex, France
Email: Virginie.Rondeau@isped.u-bordeaux2.fr
I Introduction

Dependent and informative censoring is a general issue that often arises in the analysis of censored data. And yet, most standard survival analyses are still based on the assumption of independence between time to event endpoints and the terminal event. When the independence assumption is questionable, the inference based on standard methodologies may be biased and possibly misleading. This dependence has been previously described in the analysis of progression-free survival (PFS) in oncology trials and Food and Drug Administration’s recent Draft Guidance for Industry on Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics recommends that patients who stop taking randomized therapy prior to documented progression are censored at the time when randomized treatment was stopped. The rationale for this recommendation is not provided explicitly, but seems to be related to a concern that PFS times may be over-estimated or under-estimated otherwise. A patient’s death in the absence of documented progression remains an event, irrespective of whether the death occurred while the patient was still receiving randomized therapy or some time after stopping it. This approach ignores the issue of informative censoring. Patients who stop taking randomized therapy prior to documented progression frequently do so due to either toxicity of the drug or to a deterioration in the status of their disease. In such cases, the treating physician often judges that immediate intervention, commonly the introduction of a new cancer treatment, is in the best interest of the patient without necessarily waiting for confirmatory, radiographic evidence of progressive disease. PFS therefore can be censored informatively by death or dropout. Under those circumstances, if the prevalence of censoring differs between treatment arms, naive censoring could lead to extremely biased results and, ultimately, incorrect decisions. In total, there can be many reasons to use joint models of two survival endpoints, including to give a general description of the data, to correct for bias in survival analysis due to dependent dropout or censoring and to improve efficiency of survival analysis due to the use of auxiliary information.

Time to tumour progression (TTP) is defined as the time from randomization to the first event, between loco-regional progression or distant metastases, whereas PFS is generally defined as the time from randomization to the first event, between disease progression or death without documented progression, i.e. PFS = min(TTP, death). PFS is generally considered as a better surrogate for clinical benefit than other endpoints such as TTP alone, as the PFS endpoint includes overall mortality, so that unanticipated effects of treatment on survival are included in the endpoint. For the definitions of PFS as well as TTP, a careful and precise definition of tumour progression is crucial. To assess whether PFS could be used as surrogate endpoints for overall survival (OS), research has focussed so far on empirical investigations of the correlation between estimates for PFS and OS. Censoring by death may be associated to worsening of symptoms, i.e. the patient is at higher risk of progression, while there could be situations where in fact censored patients are at less risk of disease progression, due for instance to a higher treatment benefit. Considering the definition of PFS (= the minimum of TTP and OS), it is clear that PFS and OS can hardly be fully independent since some information about OS is already contained in PFS. However, this dependency is less clear between TTP and OS, as deaths may be due to tumour progression. Although understanding the relationship between TTP and OS is highly relevant to the anti-cancer drug development, statistical models that try to describe the dependence structure between these two characteristics are not frequently used. We are aiming at filling this gap by considering a joint statistical model of TTP or PFS and OS. Fleischer et al. have recently proposed a statistical model for the dependence between PFS and OS. However, their approach was completely parametrical and using separate hazard functions for the different times to events, in addition, the value of the correlation coefficient between PFS and OS itself was of limited (at best descriptive) value. This approach can only be a starting point for more sophisticated models that...
lead to a realistic description of the dependence structure between TTP/PFS and OS. Burzykowski et al.\textsuperscript{9} proposed to use copulas models for the bivariate survival modelling. A progressive parametric multi-state model was also recently proposed by Dejardin et al.\textsuperscript{10} assuming that progression always precedes death.

Meta-analytical approaches are clearly superior to single-trial approaches to study surrogate endpoints. Meta-analyses for the validation of surrogate endpoints have become a widely accepted and applied method in oncology research.\textsuperscript{11} Without multi-trial data, it is almost impossible to make any direct inference about the association between the treatment effects and the surrogate and clinical endpoints, because one set of data cannot provide sufficient evidence of any association. Several authors have also indicated that a large sample size is required to evaluate potential surrogate endpoints. Multi-trial data readily have this capacity. In a large meta-analysis of randomized trials for head and neck cancers (Meta-Analysis of Chemotherapy in Head and Neck Cancers: MACH-NC), the efficacy of chemotherapy on OS was investigated, as adjunct to surgery or radiotherapy or both, the standard means to achieve loco-regional control.\textsuperscript{12,13} The Meta-analysis MACH-NC, based on individual patient data, compared loco-regional treatment with loco-regional treatment plus chemotherapy. The analysis showed a significant benefit on OS (HR = 0.88, [95% confidence interval (CI) 0.85–0.92]) of adding chemotherapy.

Association between PFS (or TTP) and OS depends on the disease and the characteristics of the population included such as tumour stage. For instance in old patients or in patients with major co-morbidities, deaths may be related to inter-current disease. In this case, death may often occur without a previous progression, i.e. a patient can die of inter-current disease while having a stable tumour or a tumour that responded to treatment.

In this article, we propose a general joint frailty model for clustered progression and terminal events. This approach is of interest for several reasons. First, it allows to deal with dependence between recurrent events and death for failure time data. It also gives information on whether TTP or PFS could be used as surrogate endpoints for OS. In addition, it allows a joint analysis of two processes which evolve with time, leading to more accurate estimates. This study extends previous research by dealing with clustered events (in meta-analytical studies) and giving smoothed estimates of the two hazard functions which represent incidence and mortality rates in epidemiology. It is natural in epidemiology to impose a continuous hazard function with small local variations. In this article, we propose a semi-parametric penalized likelihood method for estimating parameters in a joint frailty model for clustered and possibly recurrent data.

This article is organized as follows. In Section 2, we describe the joint frailty models. The construction of the full penalized log-likelihood is explained in Section 3. The model is applied to the joint analysis of TTP and OS in an individual patient data meta-analysis of chemotherapy in head and neck cancers (MACH-NC) in Section 4. Finally, section 5 presents a concluding discussion.

2 Joint model for times to progression and death in a meta-analysis

2.1 The models

To simplify the notation, we will set out only the joint models for one failure time and death; however, models can easily be proposed for the joint analysis of clustered recurrent failure times and death. Take the case of a study with $G$ independent clusters ($i = 1, \ldots, G$). We denote $X_{ij}$ as the survival time for subject $j$ ($j = 1, \ldots, N_i$) from group $i$, $C_{ij}$ the corresponding censoring times (not by death) and $D_{ij}$ the corresponding death times. We first consider $X_{ij}$ as a time to event (or TTP). $T_{ij} = \min(X_{ij}, C_{ij}, D_{ij})$ corresponds to each follow-up time and $\delta_{ij}$ is a binary indicator which is 0 if the
observation is censored or if the subject has died, and 1 if \( X_{ij} \) is observed \((\delta_{ij} = I(T_{ij} = X_{ij}), \text{ where } I() \) denotes an indicator function). Similarly, we note \( T_{ij}^* \) as the last follow-up time for subject \( j \), which is either a time of censoring or a time of death \((T_{ij}^* = \min(C_{ij}, D_{ij}))\) and \( \delta_{ij}^* = I(T_{ij}^* = D_{ij}) \). What we actually observe is \((T_{ij}, \delta_{ij}, \delta_{ij}^*)\).

We assume continuous TTP, terminating, and censoring processes, so that progression and death cannot happen at the same time. We adopt the convention that death happens first in the small interval \([t, t + dt)\). For subjects in the application study who died on the same day as their progression, they only count for terminal events, not as a progression.

Several formulations of the joint modelling of TTP and OS can be proposed:

- In the first setting, we assume that the association between TTP and OS is simply the result of individual associations or individual unmeasured factors. We consider the unobserved individual random effects \( \omega_{ij} \) (defined later). Following the previous model of Rondeau et al., the joint model for the hazard functions for TTP \((r_{ij}(.)\)) and death \((\lambda_{ij}(.)\)) for subject \( j \) from group \( i \) is thus:

\[
\begin{align*}
  r_{ij}(t|\omega_{ij}) &= \omega_{ij}r_0(t)\exp(\beta_i'Z_{ij}(t)) = \omega_{ij}r_{ij}(t) \\
  \lambda_{ij}(t|\omega_{ij}) &= \omega_{ij}^2\lambda_0(t)\exp(\beta_i^2Z_{ij}(t)) = \omega_{ij}^2\lambda_{ij}(t)
\end{align*}
\]

(2.1)

with \( Z_{ij}(t) \) the covariate process to simplify the notations, we assume that the covariates vector \((Z_i(t))\) are the same for the TTP and OS (the program proposed in Section 3.1 allows different covariates vectors). The effect of the explanatory variables is assumed to be different for TTP and for death times.

The random effects \( \omega_{ij} \) (frailties) are assumed independent. Mainly for reasons of mathematical convenience, the frailty terms are often assumed to follow a gamma distribution. The gamma frailty density is adopted here with unit mean and variance \( \eta \). This choice and other possibilities such as log-normal, positive stable distributions are discussed in several papers. The dependence between \( T_{ij}^* \) and \( T_{ij} \) conditional on \( Z_{ij}(t) \) is here solely due to the fact that the unobserved \( \omega_{ij} \) affects both the TTP and the death times. The common frailty parameter \( \omega_{ij} \) will take into account the heterogeneity in the data, associated with unobserved covariates.

In the traditional model, the assumption is that \( \zeta = 0 \) in (2.1), that is \( \lambda_{ij}(t) \) does not depend on \( \omega_{ij} \) and thus death (or the terminal event process) is not informative for the failure rate \( r_{ij}(t) \), i.e the two rates \( \lambda_{ij}(t) \) and \( r_{ij}(t) \) are not associated, conditional on covariates. When \( \zeta = 1 \), the effect of the frailty is identical for the TTP and for the terminating event. When \( \zeta > 0 \), the failure and the death rates are positively associated; higher frailty will result in a higher risk of progression and a higher risk of death. Inversely, when \( \zeta < 0 \), higher frailty will result in a higher probability to develop a failure but in a lower survival. The interpretation of the value of \( \zeta \) makes sense only when heterogeneity is present, i.e. when the variance of the random effects is significantly different from zero. However, in this model, we assume that the intra-cluster correlation is not present anymore after having taken into account prognosis factors and after adjustment for a subject-specific random effect term.

- In the second setting, we assume that the association between TTP and OS is the result of a clustered association (in our application, an intra-trial correlation). The second joint model is thus:

\[
\begin{align*}
  r_{ij}(t|\mu) &= u_ir_0(t)\exp(\beta_i'Z_{ij}(t)) = u_ir_{ij}(t) \\
  \lambda_{ij}(t|\mu) &= u_i^2\lambda_0(t)\exp(\beta_i^2Z_{ij}(t)) = u_i^2\lambda_{ij}(t)
\end{align*}
\]

(2.2)
The random effects \( u_i \) are assumed independent and gamma-distributed with unit mean and variance \( \theta \). The dependence between \( T_{ij}^* \) and \( T_{ij} \) conditional on \( Z_i(t) \) is here solely due to the fact that the unobserved \( u_i \), at the cluster-level affects both the TTP and the death times. The common frailty parameter \( \theta \) will also take into account the heterogeneity in the data associated with unobserved covariates at the cluster level (or intra-cluster correlation). We link the two models via the frailties \( u_i \), which can be thought of as a surrogate for the underlying process that links the two responses. The effect of \( u_i \) is different on \( \lambda_i(t) \) and \( r_i(t) \) due to the inclusion of \( \alpha \) in (2.2). However, in this model, the value of \( \alpha \) makes sense only if the variance of the random effects \( u_i \) is significantly different from zero. Zero value of \( \alpha \) implies that the dependence between TTP and OS can be fully explained by the (observed) covariates. Would the variance component be zero, TTP and time to death would also be independent even if \( \alpha \) is nonzero. On the other hand, if the variance component is nonzero and \( \alpha \) is also nonzero, then the variance component not only accounts for the intra-clusters correlations, but also represents the dependence between the progression time and the terminal event.

The Kendall’s tau can be used as a measure of dependency between the two outcomes of interest. Kendall’s \( \tau \) is the difference between the probability of concordance and the probability of discordance of two realizations of \( (T_{ij}, T_{ij}^*) \). It belongs to the interval \([-1, 1]\) and assumes a zero value when \( T_{ij} \) and \( T_{ij}^* \) are independent. Using the expression (2.2) of the joint gamma frailty model, with trial-specific random effects, one can show (see details in the Appendix) that \( \tau \) is given by:

\[
\tau = 2 \int_0^\infty \int_0^\infty \frac{(u_i^{a+1} + u_r^{a+1})}{(u_i + u_r)(u_i + u_r^{a+1})} \frac{u_i^{(1/\theta - 1)} \exp(-u_i/\theta)}{\Gamma(1/\theta) \theta^{1/\theta}} du_i \frac{u_r^{(1/\theta - 1)} \exp(-u_r/\theta)}{\Gamma(1/\theta) \theta^{1/\theta}} du_r - 1
\]

Model (2.2) can be generalized to:

\[
\begin{align*}
\{ & r_{ij}(t | u_i, \omega_{ij}) = u_i \omega_{ij} r_{i0}(t) \exp(\beta_i^o Z_i(t)) = u_i \omega_{ij} r_{ij}(t) \\
& \lambda_{ij}(t | u_i, \omega_{ij}) = u_i^2 \omega_{ij}^2 \lambda_{i0}(t) \exp(\beta_i^o Z_i(t)) = u_i^2 \omega_{ij}^2 \lambda_{ij}(t) 
\}
\end{align*}
\]

This generalization allows the baseline hazard functions used in (2.2) to be split into cluster- and subject-specific components. We assume here that the association between TTP and OS is both due to a clustered association and an individual association. We define here two random effects \( u_i \) and \( \omega_{ij} \) and assume that the cluster-level random effects \( u_i \) and the individual random effects \( \omega_{ij} \) are independent and gamma-distributed random effects \( \Gamma(1/\theta; 1/\theta) \) and \( \Gamma(1/\eta; 1/\eta) \) with \( E(u_i) = 1 \), \( \text{var}(u_i) = \theta \) and \( E(w_{ij}) = 1 \), \( \text{var}(\omega_{ij}) = \eta \). The association or dependency between TTP and OS is then assumed to be the result of both individual- or cluster-level unobserved factors. These concepts are related to the individual- and trial-level surrogacy measures proposed in the framework of meta-analytic evaluation of surrogate endpoints.\(^{17}\) We have not considered this generalization in the present application, due to the increased computational complexity of the problem.

### 2.2 Inference in the joint frailty model

We show the expression of the full log-likelihood for right-censored data for the model (2.2). The construction of the log-likelihood is detailed in Appendix 1. The expression can be easily extended to left truncated data (for instance if age is chosen as the basic timescale, see Appendix 2). In shared gamma-frailty models, the full log-likelihood for left-truncated and right-censored data takes a simple form with an analytical solution for the integrals.\(^{18}\) This is not the case in the joint frailty model (2.2). The results should not be sensitive to the choice of frailty distribution, as supported for
instance by Pickles and Crouchley’s\textsuperscript{15} simulation results. We denote \( \phi = (r_0(\cdot), \lambda_0(\cdot), \beta, \alpha, \theta) \). To construct the likelihood function, apart from the usual assumption of independent censoring (other than death), the censoring must be non-informative for \( u_i \). We obtain the following expression of the full marginal log-likelihood in the joint gamma frailty model (2.2) for right-censored data:

\[
l(\phi) = \sum_{j=1}^{G} \left\{ \sum_{i=1}^{N_j} \left[ \delta_{ij} \log r_{ij}(T_{ij}) + \delta_{ij} \log \lambda_{ij}(T_{ij}) \right] - \log \Gamma(1/\theta) - \frac{1}{\theta} \log \theta \right. \\
+ \log \int_0^{\infty} \left( m_j + \alpha \delta_{ij} + 1/\theta - 1 \right) \exp \left( -u_i/\theta - u_i \sum_{j=1}^{N_i} R_{ij}(T_{ij}) - u_i^\alpha \sum_{j=1}^{N_i} \Lambda_{ij}(T_{ij}) \right) du_i \right\} \tag{2.4}
\]

with, \( \Lambda_{ij}(t) = \int_0^t \lambda_{ij}(v)dv \) the cumulative hazard function for death, with \( \Lambda_{ij}(\cdot|u_i) = u_i^\alpha \Lambda_{ij}(\cdot) \), and \( R_{ij}(t) = \int_0^t r_{ij}(v)dv \) the cumulative hazard function for recurrent events, with \( R_{ij}(\cdot,u_i) = u_i R_{ij}(\cdot) \).

### 3 The semi-parametric penalized likelihood approach

We introduce a semi-parametric penalized likelihood approach to estimate the different parameters \( \beta, \alpha, \theta \) and the baseline hazard function \( r_0(t) \) for TTP or \( \lambda_0(t) \) for death times.\textsuperscript{4,19}

In most situations, it is reasonable to expect smooth baseline hazard functions, piecewise constant modelling for the hazard functions being often unrealistic. To introduce such knowledge \textit{a priori}, we penalize the likelihood by a term which has large values for rough functions.\textsuperscript{20,21} The roughness penalty function is represented by the sum of two squared norms of the second derivative of the hazard functions.\textsuperscript{20} The penalized log-likelihood is thus defined as

\[
pl(\phi) = l(\phi) - \kappa_1 \int_0^{\infty} \ddot{r}_0^2(t)dt - \kappa_2 \int_0^{\infty} \ddot{\lambda}_0^2(t)dt \tag{3.1}
\]

where \( l(\phi) \) is the full log-likelihood defined in (2.4), \( \kappa_1 \) and \( \kappa_2 \) the positive smoothing parameters which control the trade-off between the data fit and the smoothness of the functions. Maximization of (3.1) defines the maximum penalized likelihood estimators (MPnLE) \( \hat{r}_0(t), \hat{\lambda}_0(t), \hat{\beta}, \hat{\alpha}, \hat{\theta} \) and \( \eta \). We directly use \( \hat{H}^{-1} \) as a variance estimator, where \( H \) is minus the converged hessian of the penalized log-likelihood.

The estimators \( \hat{r}_0(\cdot), \hat{\lambda}_0(\cdot) \) cannot be calculated explicitly but can be approximated on a basis of splines. Splines are piecewise polynomial functions that are combined linearly to approximate a function on an interval. We use cubic M-splines, which are a variant of cubic B-splines (for more details see Ramsay\textsuperscript{22}). M-splines are non-negative and easy to integrate or differentiate. As we use cubic splines (or of order 4), the second derivative of \( r \) or \( \lambda \) is approximated by a linear combination of piecewise polynomial of order 2. This approximation allows flexible shapes of the hazard functions while reducing the number of parameters. If we denote \( \tilde{r}(\cdot) \) as an approximation to the MPnLE \( \tilde{r}(\cdot) \), the approximation error can be made as small as desired by increasing the number of knots. In our approach, although there are two different hazard functions (for TTP and for death), we use the same basis of splines for each function but the spline coefficients are different for the distinct functions.

We have previously shown that to obtain a good estimation of the theoretical hazard function, the more knots we used, the closer the MPnLE was to the true hazard function.\textsuperscript{18} The smoothing parameters can be chosen by maximizing a likelihood cross-validation (LCV) criterion as in
Another approach consists in fixing the number of degrees of freedom to estimate the hazard function, as has been previously described. We thus use the relation linking the model degrees of freedom (mdf) and the smoothing parameter $k$ to evaluate the smoothing parameter:

$$\text{mdf} = \text{trace} \left( \left( \frac{H}{C} \right)^{\frac{1}{2}} \right) I \left( \frac{H}{C} \right)^{-\frac{1}{2}}$$

(with $I$ the hessian matrix of the log-likelihood computed at the MPnLE). Indeed, it is easier to specify a number of degrees of freedom to estimate a given curve, rather than to specify a smoothing parameter.

We propose to maximize directly the observed log-likelihood (3.1) using a modified robust Marquardt optimization algorithm, which is a combination between a Newton–Raphson algorithm and the steepest descent algorithm. This algorithm is more stable than the Newton–Raphson algorithm but preserves its fast convergence property near the maximum. The integrations in the full log-likelihood expression in (2.3) were evaluated using Gaussian quadrature. Laguerre polynomials with 20 points were used to treat the integration $[0,1)$. To test whether the variance of the random effects was different from zero, i.e. $H_0 : \theta = 0$ versus. $H_a : \theta > 0$, the $p$-value from a Wald test was used. Because zero is the lower bound of $\theta$, the unilateral Wald test was used, which is equivalent to a squared Wald test with a half-half mixture of zero and a chi-squared distribution with degree of freedom 1.

### 3.1 Software

The statistical software used was R and the library frailtypack (version 2.2-12) with the function frailtyPenal for the joint frailty models (2.1). The frailtyPenal function was slightly modified to be able to handle joint modelling (2.2) and inserted in the R package ‘frailtypack’. Penalized likelihood maximization was used. In the reduced shared frailty models, $k_1$ and $k_2$ were evaluated using the cross-validation method, thereafter this value was used in the joint model. Furthermore, to deal with the constraint on the variance component ($\theta > 0$ and $\eta > 0$), we used a squared transformation and standard errors of $\theta$ and $\eta$ were computed by the $\Delta$-method. This reparametrization did not have any major adverse effect on the approximation, while being very numerically convenient. Since the initialization of the parameters is very important in this approach, we initialized the parameters with values obtained with simpler frailty models or Cox proportional hazard models.

### 4 Application to the meta-analysis MACH-NC

The meta-analysis MACH-NC, based on individual patient data, compared loco-regional treatment (as radiotherapy and/or surgery) with loco-regional treatment plus chemotherapy. We previously evaluated the heterogeneity between the underlying risks in trials and the heterogeneity of treatment effect between trials, using correlated random effects. Results in the MACH-NC data confirmed a slight benefit for OS of adding chemotherapy to loco-regional treatment. A key issue when analysing these pooled data is to assess the dependency between TTP and OS. We thus aimed to investigate the proposed joint random effects models (2.1,2.2) for this updated MACH-NC data with treatment as a fixed effect (treated by chemotherapy versus not) and independent and gamma-distributed trial-specific or individual-specific random effects.

#### 4.1 Data collection

Head and neck carcinomas (oral cavity, oropharynx, hypopharynx, nasopharynx and larynx) are frequent tumours for which surgery and/or radiotherapy are the main therapeutic agents. In the absence of a large (over 1000–2000 patients) randomized trial, the most reliable way to evaluate
chemotherapy is to do a meta-analysis based on updated individual patient data. The initial Meta-Analysis of Chemotherapy on Head and Neck Cancer (MACH-NC) analysed data from nearly 11 000 patients in 63 randomized trials between 1965 and 1993. The first results showed an absolute benefit of 4% at 5 years in OS in favour of chemotherapy \((p < 0.0001)\). Most of the benefit was seen with concomitant radio-chemotherapy but with a relatively large heterogeneity in this subgroup of trials. The MACH-NC group decided to confirm the results by updating its database with the inclusion of the randomized trials performed between 1994 and 2000. Results were published in 2009. Because some trials had strata that corresponded to different loco-regional treatments or chemotherapies, and because some trials had three-arms or a \(2 \times 2\) design, a few trial arms were used twice, such that the number of comparisons in the meta-analysis was 108 and the number of patients was 17 493. The overall median follow-up was 5.6 years. The description of the dataset included and their references can be found in Ref. 13

As in Ref. 6, we pre-specified exclusion of two trials, which had no registered recurrences at all: one trial of induction chemotherapy (680 patients) and one trial of adjuvant chemotherapy (499 patients). We also excluded two very small trials (58 and 27 patients) of concomitant chemotherapy from the same institution, in which all patients were dead at 2 years. In MACH-NC, trials including only nasopharyngeal carcinomas were excluded.

### 4.2 Endpoints

Duration of TTP was defined as the time from randomization to the first event (loco-regional recurrence, loco-regional progression or distant metastases). Deceased patients were censored at the date of death. Patients without documented evidence of death were censored at the date of last follow-up. In all analyses, we used the locoregional or distant events as provided by the trial investigators, as defined by their own timing and assessment method. These definitions were heterogeneous across the included trials. For example, in some trials that investigated the addition of chemotherapy to radiotherapy as loco-regional treatment, patients who never reached complete remission after radiotherapy (defined by clinical or radiological assessment) were considered as having loco-regional failure at time zero, i.e. randomization.

PFS was defined as the time from randomization to the first event (i.e. locoregional, distant recurrence or death from any cause). Patients without documented evidence of an event were censored at the date of last follow-up. What we defined as PFS was often called disease-free survival in trials that included patients with resectable tumours and PFS in trials that included patients with non-resectable tumours. In the meta-analyses, both types of trials were present. There are also trials that included the mixed population. Therefore, we chose to apply our definition of PFS to all trials.

OS was defined as the time from randomization to death from any cause; patients still alive at the last follow-up visit were censored at the date of last follow-up.

During the data-collection process of the meta-analyses, central analyses of different types of events (i.e. loco-regional or distant relapses and deaths) were performed and sent out to the investigators for approval or modification.

### 4.3 Parameter estimates and interpretation

We analysed 16 099 patients from 103 treatment–control comparisons. The number of patients per trial varied between 25 and 573 with a median of 114 and a mean of 156. A total of 10 538 patients (65.3%) died and the number of deaths over trials ranged from 10 to 493 with a median of 74 and a
mean of 102. A total of 8,314 patients (51.6%) experienced tumour progressions (between 5 and 487 by unit, with a median of 55 and a mean of 81).

The first results are presented in Table 1 comparing frailty models: the first column represents the reduced shared frailty models for the three outcomes TTP, PFS and OS (i.e. considering that those events are independent) and the other columns the joint frailty models (2.1) and (2.2), i.e. with a subject- or a cluster-level dependency between times to progression and death times or between PFS and death times.

Using simple shared frailty models (with cluster random effects) yielded a benefit of adding chemotherapy to loco-regional treatment for the TTP, the PFS and death. The variance of the cluster random effects were significantly different from zero for TTP, \( \theta = 0.48 \) (SE = 0.07), for PFS, \( \theta = 0.09 \) (SE = 0.02) and for OS, \( \theta = 0.18 \) (SE = 0.03). When comparing joint model (2.2) to the reduced shared frailty models, the treatment effect on TTP is exactly the same, while the hazard ratio of chemotherapy on OS is closer to one in the joint model (2.2), but still significantly different from 1. Ignoring the dependence between time to the terminal event and TTP resulted in biases in the independent shared frailty model compared to the joint model. The same tendencies were observed for PFS. The intra-cluster association was lower in the shared frailty models compared to the joint model (2.2) \((\theta = 1.17 \text{ for TTP and } \theta = 1.10 \text{ for PFS})\). This can be explained by the fact that the variance \( \theta \) of the cluster-specific random effects not only takes into account the intra-cluster correlation in the data but also the dependence between TTP and time to death.

In the joint frailty model (2.1), the chemotherapy treatment was also significantly associated with TTP and with OS, with the same tendencies. The chemotherapy treatment was no more significantly associated with PFS, nor with OS in the joint frailty model (2.1). The differences for TTP and OS between models (2.1) and (2.2) can be explained by their different modelling. Model (2.1) takes the individual over-dispersion into account, whereas model (2.2) takes into account the between trials heterogeneity.

In the joint frailty model (2.2), we observed that the treatment was more beneficial for TTP (RR = 0.88 (0.85–0.92)) than for OS (RR = 0.93 (0.89–0.97)). The treatment was also more beneficial for TTP than for PFS (RR = 0.93 (0.90–0.96)).

The joint models (2.1) and (2.2) showed significant variance of the random effects, meaning significant association between TTP and times to death, or between PFS and times to death, due to factors unobserved at the individual-level or at the cluster-level. We observe a positive association between the failure rates \( \lambda_{ij}(t) \) and \( r_{ij}(t) \) (measured by \( \zeta > 0 \) or \( \alpha > 0 \)). This indicates that the incidence of progression is positively associated with death.

The values of Kendall’s tau give moderate positive association between TTP and OS (or between PFS and OS) with \( \tau = 0.37 \) (\( \tau = 0.45 \)) in the joint frailty model (2.1) and \( \tau = 0.31 \) (\( \tau = 0.33 \)) in the joint frailty model (2.2) adjusted for the treatment effect. As expected by definition, the association between PFS and death is higher than the association between TTP and death.

The trials could be divided into three pre-established categories according to the timing of chemotherapy: adjuvant (after loco-regional treatment), induction (before loco-regional treatment) or concomitant (given concomitantly or alternating with radiotherapy). The results (Table 2) indicate that most of the benefit was seen with concomitant radio-chemotherapy but with a larger and significant link between the two outcomes in this subgroup of trials \((\text{var}(u_i) = 0.64 \text{ (SE = 0.12) after adjustement, see Table 3})\). As previously seen in the whole sample, when comparing joint model (2.2) to the reduced shared frailty models (results not shown), the treatment effect on TTP is very similar while the hazard ratio of chemotherapy on OS is closer to one in the joint model (2.2), but still significantly different from 1.
Table 1. Joint analysis of the TTP, PFS and death for the entire meta-analysis MACH-NC (1965–2000)

<table>
<thead>
<tr>
<th></th>
<th>Three reduced shared frailty models</th>
<th>Joint frailty models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>((n=16,099, ,G=103))</td>
<td>((n=16,099, ,G=103))</td>
</tr>
<tr>
<td></td>
<td>((\text{TTP, OS}))</td>
<td>((\text{PFS, OS}))</td>
</tr>
<tr>
<td>RR (95%CI)</td>
<td>Joint model (2.1) RR (95%CI)</td>
<td>Joint model (2.1) RR (95%CI)</td>
</tr>
<tr>
<td>TTP</td>
<td>0.87 (0.83–0.91)</td>
<td>0.92 (0.87–0.96)</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>Control (0)</td>
<td>0.88 (0.85–0.92)</td>
</tr>
<tr>
<td>(pl^a)</td>
<td>–16518.7</td>
<td>0.99 (0.91–1.04)</td>
</tr>
<tr>
<td>TTP or death</td>
<td>0.91 (0.88–0.95)</td>
<td>0.93 (0.90–0.96)</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>Control (0)</td>
<td>0.97 (0.88–1.06)</td>
</tr>
<tr>
<td>(pl^a)</td>
<td>–22,895.6</td>
<td>0.93 (0.90–0.97)</td>
</tr>
<tr>
<td>Time to death</td>
<td>0.89 (0.86–0.92)</td>
<td>0.90 (0.82–0.99)</td>
</tr>
<tr>
<td>Chemotherapy</td>
<td>Control (0)</td>
<td>0.93 (0.89–0.97)</td>
</tr>
<tr>
<td>(\theta = \text{var}(u_i) , (SE))</td>
<td>1.17 (0.15)</td>
<td>0.97 (0.88–1.06)</td>
</tr>
<tr>
<td>(\alpha)</td>
<td>0.88 (0.03)</td>
<td>0.93 (0.90–0.97)</td>
</tr>
<tr>
<td>(\eta = \text{var}(\omega_j) , (SE))</td>
<td>3.51 (0.05)</td>
<td>1.03 (0.01)</td>
</tr>
<tr>
<td>(\zeta)</td>
<td>2.61 (0.04)</td>
<td>0.99 (0.04)</td>
</tr>
<tr>
<td>(pl^a)</td>
<td>–24,315.7</td>
<td>–43,166.0</td>
</tr>
<tr>
<td>LCV(^b)</td>
<td>1.59</td>
<td>1.77</td>
</tr>
</tbody>
</table>

\(^a pl^a\text{ Marginal penalized Log-Likelihood}\)
\(^b\text{Approximate Cross-validation criterion (see Section 4.4)}\)
Figure 1 illustrates the survival functions for the three timing groups and according to the treatment. This confirms the greater survival benefit obtained with chemotherapy in the group of concomitant trials for the times to progression. We also observed a lower baseline survival for concomitant trials than for other timing of chemotherapy. Even after adjusting for the patients’ demographics (sex and age), the stage and the site of the initial tumour, there appeared to be a positive association between TTP and overall mortality due to unknown factors (Table 3). Adjustment for different individual covariates had no effect on the estimation of the overall treatment effect neither on its standard deviation. Very similar effects were observed for TTP and for OS in the joint models. When adjusting for covariates, the heterogeneity between trials decreased (from $\theta=1.17$ (0.15) to $\theta=0.76$ (0.10) in the joint model (2.2)) but remained significantly different from zero, this also indicates a significant association between the progression times and death.

We also performed a sensitivity analysis comparing recent trials, i.e. ending accrual between 1994 and 2000 (a more homogeneous group of trials mainly concomitant, with higher data quality and improved follow-up techniques) to the oldest (1965–1993). The observed percentages of progressions (45.8% vs 56.7%) or death (61.5% vs 70.3%) were lower in the more recent period than in the oldest. This is partly explained by the shorter follow-up for the more recent period, even if the follow-up

| Table 2. Joint analyses (using model 2.2) of the TTP and death or of the PFS and death for the treatment effect according to chemotherapy timing for the meta-analysis MACH-NC (1965–2000) |
|---|---|---|
| | Adjuvant ($n = 2\ 068, G = 11$) | Induction ($n = 4\ 631, G = 33$) | Concomitant ($n = 9\ 400, G = 59$) |
| **TTP** |  |  |  |
| Chemotherapy | Treated | 0.89 (0.77–1.03) | 1.02 (0.94–1.11) | 0.81 (0.76–0.85) |
| | Control |  |  |  |
| **Time to death** |  |  |  |
| Chemotherapy | Treated (1) | 1.06 (0.93–1.20) | 0.97 (0.90–1.04) | 0.88 (0.84–0.93) |
| | Control (0) |  |  |  |
| $\theta = \text{var}(u_i)$ (SE) | 0.28 (0.11) | 0.80 (0.18) | 0.48 (0.08) |
| $\alpha$ | 1.31 (0.08) | 0.72 (0.05) | 0.66 (0.03) |
| Marginal penalized | $-4315.4$ | $-11\ 559.4$ | $-21\ 820.2$ |
| Log Likelihood ($pl$) |  |  |  |
| **TTP** |  |  |  |
| Chemotherapy | Treated | 1.03 (0.92–1.16) | 1.00 (0.93–1.07) | 0.88 (0.84–0.92) |
| | Control |  |  |  |
| **Time to death** |  |  |  |
| Chemotherapy | Treated (1) | 1.05 (0.93–1.20) | 0.98 (0.91–1.05) | 0.87 (0.84–0.93) |
| | Control (0) |  |  |  |
| $\theta = \text{var}(u_i)$ (SE) | 0.29 (0.12) | 0.18 (0.04) | 1.24 (0.20) |
| $\alpha$ | 1.30 (0.08) | 0.88 (0.06) | 0.93 (0.05) |
| Marginal penalized | $-5107.3$ | $-13\ 483.3$ | $-25\ 467.4$ |
| Log Likelihood ($pl$) |  |  |  |
techniques were improved. The magnitude of the benefit of the treatment was more important for the 1994–2000 trials (adjusted RR = 0.80 (0.74–0.86) for TTP and adjusted RR = 0.94 (0.88–1.00) for death) than for the 1965–1993 trials, especially for TTP (adjusted RR = 0.98 (0.91–1.05) for TTP and adjusted RR = 1.01 (0.95–1.08) for death). The difference between periods is explained by the change of proportion of concomitant trials over time. When the analysis is restricted to concomitant trials, the two RR values are similar. Some of the results presented in this article are different than those provided in the medical publications. 12 This is probably due to the fact that the previous simple survival models did not take into account the dependence with death, whereas the new joint

### Table 3.

<table>
<thead>
<tr>
<th></th>
<th>Total (n = 14 977, G = 96)</th>
<th>Adjuvant (n = 1 825, G = 9)</th>
<th>Induction (n = 4 366, G = 30)</th>
<th>Concomitant (n = 8 786, G = 57)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RR (95%CI)</td>
<td>RR (95%CI)</td>
<td>RR (95%CI)</td>
<td>RR (95%CI)</td>
</tr>
<tr>
<td><strong>TTP</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treated</td>
<td>0.89 (0.86–0.94)</td>
<td>0.90 (0.77–1.05)</td>
<td>1.04 (0.96–1.13)</td>
<td>0.82 (0.77–0.87)</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51–60 vs &lt;50</td>
<td>1.11 (1.05–1.17)</td>
<td>1.15 (0.94–1.42)</td>
<td>1.03 (0.93–1.14)</td>
<td>1.07 (0.99–1.15)</td>
</tr>
<tr>
<td>&gt;60 vs &lt;50</td>
<td>1.10 (1.04–1.16)</td>
<td>1.02 (0.83–1.25)</td>
<td>1.04 (0.94–1.16)</td>
<td>1.07 (0.99–1.14)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female vs Male</td>
<td>0.89 (0.84–0.95)</td>
<td>0.87 (0.70–1.07)</td>
<td>0.81 (0.71–0.92)</td>
<td>0.92 (0.85–0.99)</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III vs I + II</td>
<td>2.04 (1.85–2.26)</td>
<td>1.39 (1.12–1.72)</td>
<td>1.49 (1.24–1.79)</td>
<td>2.37 (2.00–2.80)</td>
</tr>
<tr>
<td>IV vs I + II</td>
<td>3.06 (2.79–3.36)</td>
<td>1.73 (1.41–2.13)</td>
<td>2.13 (1.78–2.54)</td>
<td>3.50 (2.99–4.10)</td>
</tr>
<tr>
<td>Site</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Larynx vs others</td>
<td>0.84 (0.79–0.89)</td>
<td>0.79 (0.64–0.96)</td>
<td>0.86 (0.75–0.98)</td>
<td>0.88 (0.81–0.95)</td>
</tr>
<tr>
<td><strong>Time to death</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemotherapy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treated</td>
<td>(1)</td>
<td>0.93 (0.90–0.97)</td>
<td>1.06 (0.93–1.21)</td>
<td>0.98 (0.91–1.05)</td>
</tr>
<tr>
<td>Control (0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51–60 vs &lt;50</td>
<td>1.13 (1.08–1.19)</td>
<td>1.20 (0.99–1.44)</td>
<td>1.08(0.99–1.19)</td>
<td>1.12 (1.05–1.20)</td>
</tr>
<tr>
<td>&gt;60 vs &lt;50</td>
<td>1.25 (1.19–1.32)</td>
<td>1.35 (1.12–1.62)</td>
<td>1.19 (1.08–1.30)</td>
<td>1.27 (1.19–1.36)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female vs Male</td>
<td>0.84 (0.79–0.89)</td>
<td>0.74 (0.61–0.90)</td>
<td>0.82 (0.74–0.92)</td>
<td>0.88 (0.82–0.95)</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>III vs I + II</td>
<td>1.36 (1.24–1.48)</td>
<td>1.53 (1.26–1.85)</td>
<td>1.22 (1.06–1.41)</td>
<td>1.29 (1.13–1.48)</td>
</tr>
<tr>
<td>IV vs I + II</td>
<td>1.91 (1.75–2.08)</td>
<td>2.00 (1.66–2.42)</td>
<td>1.75 (1.52–2.01)</td>
<td>1.73 (1.52–1.96)</td>
</tr>
<tr>
<td>Site</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Larynx vs others</td>
<td>0.78 (0.74–0.82)</td>
<td>0.89 (0.75–1.05)</td>
<td>0.72 (0.64–0.81)</td>
<td>0.83 (0.78–0.89)</td>
</tr>
<tr>
<td>( \theta = \text{var}(u_i) ) (SE)</td>
<td>0.76 (0.10)</td>
<td>0.22 (0.10)</td>
<td>0.40 (0.10)</td>
<td>0.64 (0.12)</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>0.75 (0.03)</td>
<td>1.83 (0.15)</td>
<td>0.63 (0.06)</td>
<td>0.69 (0.04)</td>
</tr>
<tr>
<td>Marginal penalized</td>
<td>-35 615.1</td>
<td>-3609.7</td>
<td>-10 980.7</td>
<td>-20 485.9</td>
</tr>
<tr>
<td>Log-Likelihood ((p_l))</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
survival modelling deals with this dependence and assumes that the association between TTP and OS is simply the result of unmeasured factors.

We also compared a penalized likelihood method of estimation with a parametric approach (i.e. using Weibull baseline hazard functions). The two methods gave similar estimations and standard deviations. The variance of the random effects ($\theta = 0.98$ (0.12) in model (2.2) and $\eta = 0.74$ (0.01) in model (2.1)) and their corresponding coefficients ($\alpha = 1.05$ and $\zeta = 3.27$) were slightly under-estimated in the parametric approach compared to the semi-parametric approach (Table 1).

### 4.4 Model choice

A set of models were fitted in our analysis. For the random effects in the joint models, both trial- and individual-specific associations are considered and compared. Having included different types of random effects in the joint models for the two event-time outcomes, we obtained several candidate models from which a best-fitting model needed to be selected. The LCV criterion is adopted to guide

![Figure 1. Survival functions for times to progression or time to death according to the chemotherapy and timing of the chemotherapy (in bold: treated) using joint frailty models (2.2).](image)
the choice of the model used in this analysis. As LCV is particularly computationally demanding when \( n \) is large, an approximate version \( LCV_a \) has been proposed by O’Sullivan\(^{20} \) for the estimation of the hazard function in a survival case and adapted recently by Commenges et al.\(^{30} \) Lower values of \( LCV_a \) indicate a better fitting model. The \( LCV_a \) is then defined as:

\[
LCV_a = \frac{1}{n} ( \text{trace}(H_{pl}^{-1} H_l) - l(.))
\]

with \( H_{pl} \) minus the converged hessian of the penalized log-likelihood and \( H_l \) minus the converged hessian of the log-likelihood and \( l(.) \) is the full log-likelihood.

In the case of a parametric approach, \( \text{trace}(H_{pl}^{-1} H_l) - l(.) \) will represent the number of parameters and \( LCV_a \) will be approximately equivalent to the Akaike information criterion. The \( LCV_a \) for parametric approaches is defined as:

\[
LCV = \frac{1}{n} (np - l(.))
\]

with \( np \) the total number of parameters.

In Table 1, we report the values of \( LCV_a \) for the joint models considered. The smallest \( LCV_a \), and therefore the best fit is achieved by the joint model (2.2) using a penalized likelihood estimation. This indicates that the joint association between TTP and times to death would be the result of unobserved factors at the trial level. Because the type of trial is probably correlated with both the ‘sickness’ of the patients and their survival times, linking the responses at the cluster-level would be more appropriate. This cluster-level link would indicate that trial could be a better surrogate for the unobserved process, and that a cluster-level link could be more likely to result in conditional independence of TTP and of times to death.

### 5 Discussion/Conclusion

In this article, we have developed a joint frailty model for clustered time-to-event data and a dependent terminal event. A penalized likelihood was used for the estimation of parameters in this joint model. Very similar effects were observed for TTP and for OS in the joint models, but with a higher beneficial effect of the treatment for TTP than for death. We observed a cluster- or trial-level link between the TTP and the times to death in the meta-analysis MACH-NC. As expected by definition, the association between PFS and death (measured with the Kendall’s tau) is higher than the association between TTP and death. This link was more pronounced in the concomitant group of trials. This cluster-level link would indicate that trial is a better surrogate for the unobserved process.

The MACH-NC database might not be the most ideal database to answer the question of the impact of dependent terminal event in this specific clinical situation, but may be used for illustrative purposes. Indeed, the quality of the data is highly variable (some trials have a high proportion of patients who died of their cancer without date of progression before death). The trials were performed between 1965 and 2000 and the tools to detect progression have changed with time. After radiotherapy, it is sometimes difficult to know if a patient is in complete response or not, or to know when the loco-regional progression occurs. This problem was handled differently from one trial to the other. In some trials, the absence of complete response was considered as an initial loco-regional ‘progression’/failure, others used the change in tumour size or appearance of new nodes as loco-regional progression. At least, the quality of detection was the same in both arms.

In a meta-analysis combining survival data from different clinical trials, an important issue is the possible heterogeneity between trials. Such inter-trial variation cannot only be explained by the
heterogeneity of the patients’ baseline risk, but also by the heterogeneity of treatment effects across trials. Such a scenario can be accounted for using additive random effects in the Cox model, with a random trial effect and a random treatment by trial interaction, as we have previously proposed.29 Thus, to include a random treatment by trial interaction in the joint model (2.2) would require complex correlation structure. While the model can be extended to other correlation structures, we chose not to do so, due to the increased computational complexity of the problem.

In this article, we have focussed on the cluster-level association between TTP (or PFS) and death in a meta-analysis of clinical trials using model (2.2). However, this association at the cluster-level may also be present in multicentre clinical trials. These two designs have differences: the meta-analysis will more often be international, with a larger sample size and with different protocols for each trial. In multicentre clinical trials, the participating centres may vary from university centres to community hospitals, so there is likely to be variation (heterogeneity) as well. The proposed model and estimation approach are of particular interest for both designs in order to study the heterogeneity when the sample size is sufficient.

In analyses of the natural history of cancer, there is a great interest in dynamic prediction of death, that is, in the computation of the predictive distribution of death at a certain moment in time given, the history of event(s) (ex: local or distant relapse) and covariates until that moment.31,32 The interest is then to produce an accurate estimate of the probability to survive beyond \( t + \Delta t \), conditional on information available at the prediction time point \( t \). The joint models described here could be an interesting prognostic tool to dynamically assess a patient’s prognosis of death using recurrence information. To more accurately guide clinical decision making, the monitoring of recurrences after initial treatment would be facilitated by dynamic powerful prognostic tools that incorporate the complete post-treatment recurrence evolution. It is worth to note that any investigations on the dependency structure between OS and TTP can only be reasonable if the measurement of TTP itself is accurate. Therefore, before thinking about predictions for OS, issues like the determination of progression and interval censoring of TTP need to be taken into account. Furthermore, in our application, we censored patients at their date of last follow-up or at their date of death. We then assumed that patients were actively followed up for their progression until their date of censoring. This assumption can be accurate for some trials but more questionable for other trials.

Another application of these joint frailty models would be to study the specific causes of death, i.e. cancer-death instead of OS. However, from a modelling point of view, this would require the modelling of three functions instead of two: one for TTP, one for deaths related to head and neck cancers and one for non-cancer deaths. This extension could be a little bit computationally challenging. Recent publication on the same meta-analysis has shown that the benefit of chemotherapy was mostly due to its effect on deaths related to head and neck cancers (HR = 0.78 [0.73–0.84], \( p < 0.0001 \)) and had no effect on non-cancer deaths (HR = 0.96 [0.82–1.12], \( p = 0.62 \)).

The likelihood cross-validation criterion (\( LCV_a \)) measures only the relative goodness of fit among a collection of models. Although \( LCV_a \) contributes to the relative identification of the best-fitting model, it provides no information about the absolute adequacy of the models. Thus, other model diagnostic tools are needed to assess the adequacy of the model. A future research topic would be to study model checking, using for instance martingale residuals.

**Acknowledgements**

The authors thank the trialists and research group who agreed to share and update their data and the following institutions for funding the investigators’ meetings or the meta-analysis project: Association pour la Recherche
Trialists and research group

D.J. Adelstein (Cleveland Clinic Foundation), J.P. Armand (Institut Claudius Regaud), C. Amand (Intitut Gustave-Roussy [IGR]), H. Audry (IGR), J.M. Bachaud (Institut Claudius Regaud), H.G. Bartelink (Netherlands Cancer Institute, Amsterdam), J. Bernier (Clinique de Genolier), W.R. Bezwoda (University of the Witwatersrand), A. Buffoli (Istituto di Radioterapia Oncologica, Udine), J. Bourhis (IGR), K.T. Bhowmik (Vardhaman Mahavir Medical College), D. Brizel (Duke University Medical Center), G.P. Browman (Mc Master University), V. Budach (Charité University Medicine Berlin), G. Calais (Centre Hospitalier Universitaire de Tours), B.H. Campbell (Medical College of Wisconsin), A. Carugati (Instituto de Oncología Ángel H Roffo), S. Chalkidou (Hellenic Cooperative Oncology Group), J.R. Clark (Dana Farber Cancer Institute), E. Cohen (Institut Claudius Regaud), L. Collette (European Organization for Research and Treatment of Cancer), O. Dal Canton (Ospedale S Giovanni Vecchio, Torino), D. Dalley (St. Vincent’s Hospital, Sydney), J. Depondt (Centre Hospitalier Universitaire de Bichat), L. Désigné (IGR), A. Deszcz-Thomas (Centre Hospitalier Universitaire de la Pitié-Salpêtrière), B. Di Blasio (Parma University Hospital), W. Dobrowsky (University of Vienna), C. Domenge (IGR), F. Eschwege (IGR), J.F. Evensen (Norwegian Radium Hospital), Radiation Therapy Oncology and Head and Neck Cancer Groups of the European Organisation for Research and Treatment of Cancer, C. Fallai (Università di Firenze), J.J. Fischer (Yale University), A.A. Forastière (Johns Hopkins University), G. Fountzilas (Aristotle University of Thessaloniki), K.K. Fu (University of California San Francisco), D. Gedouin (Centre Eugène Marquis), S. Guérin (IGR), F. Geara (MD Anderson Cancer Centre), R. Giglio (Instituto de Oncología Ángel H Roffo), N.K. Gupta (Christie Hospital, Manchester), E. Haddad (Hôpital Universitaire H Mondor), B.G. Haffty (UMDNJ-Robert Wood Johnson Medical School), Y. Hasegawa (Aichi Cancer Centre), C. Hill (IGR), J.C. Horiot (Centre Georges François Leclerc), M. Hor lux (Tokaï University), J. Houghton (Cancer Research Campaign & University College London Cancer Trials Centre), P. Huguenin (Deceased) (University Hospital Zurich), P. Jan (IGR), Ch. Jaulerry (Institut Curie), B. Jeremic (Kragulevac University Hospital), T. Jouffroy (Institut Curie), A. Jortay (Brumann University Hospital), B. Kapstad (University of Bergen), L.P. Kowalski (A.C. Camargo Hospital, Sao Paulo), S. Kramer (Thomas Jefferson University Hospital), K.S. Krishnamurthi (Madras Cancer Institute), S. Kumar (Sanjay Gandhi Post Graduate Institute of Medical Science), G. Laramore (University of Washington, Seattle), E. Lartigau (Centre Oscar Lambret), J.L. Lefebvre (Centre Oscar Lambret), A. Le Maitre (IGR), T.Leon (Harvard University), F. Lewin (Huddinge University Hospital), B. Luboinski (IGR), M. Luboinski (IGR), E. Maillard (IGR), T. Maipang (Prince of Songkla University), M. Martin (Centre Hospitalier Intercommunal de Créteil), J.J. Mazeron (Centre Hospitalier Universitaire de la Pitié-Salpêtrière), M. Merlano (National Institute for Cancer Research, Genoa), S. Mehta (Mumbai Group), R. Molinari (Istituto Nazionale Tumori, Milano), K. Monson (Cancer Research Campaign & University College London Cancer Trials Centre), K. Morita (Aichi Cancer centre), R.P. Mueller (University of Wuerzburg), A. Murias (Hospital Insular del Gran Canaria, Las Plamas), V. Mosseri (Institut Curie), G. Numico (National Institute for Cancer Research, Genoa), P. Olmi (Università di Firenze), B. O’Sullivan (Princess Margaret Hospital, Toronto University) J. Overgaard (Aarhus University Hospital), A. Paccagnella (SS Giovanni and Paolo Hospital, Venezia), T. Pajak (Radiation Therapy Oncology Group), L.M. Parvinen (Turku University Hospital), N.W. Pearlman (Denver VA Medical Center), J.P. Pignon (IGR), R.S. Rao (Tata Memorial Hospital, Bombay), J.M. Richard (IGR), K. Rufibach (Swiss Group for Clinical Cancer Research), F. Sanchiz (I Policlinico, Barcelona), H. Sancho-Garnier (Centre Val d’Aurelle), D.E. Schuller (Ohio State University Hospital, Columbus) V. Shanta (Madras Cancer Institute), R.J. Simes (NHMRC Clinical Trials Center, Camperdown), L. Smid (University of
Ljubljana), L.A. Stewart (University of York), S. Staar (University of Cologne), P. Strojan (Institute of Oncology Ljubljana), H. Stuetzer (University of Cologne), H. Szpirglas (Centre Hospitalier Universitaire de la Pitié-Salpêtrière), G. Schwaab (IGR), N. Syz (IGR), S. Takaku (Saitama Medical School), S.G. Taylor (Rush Medical Center), J.S. Tobias (Cancer Research Campaign & University College London Cancer Trials Centre), R.J. Toohill (Medical College of Wisconsin), E.E. Vokes (University of Chicago Medical Center), P. Volling (University of Cologne), M.C. Weissler (University of North Carolina, Chapel Hill), T. Wendt (University of Jena), K.D. Wernecke (Charité University Hospital), J. Widder (University of Vienna), G.T. Wolf (University of Michigan) and K. Yoshino (Osaka Medical Center for Cancer and Cardiovascular Disease).

References


Appendix 1

Construction of the full log-likelihood for the joint frailty model for clustered survival data possibly right censored

Note $T_i = (T_{i1}, \ldots, T_{iN_i})$ is the observation times and $T_{i} = (T_{i1}, \ldots, T_{iN_i})$ the last follow-up times for subjects from cluster $i$, $\phi = (r_0(.), \lambda_0(.), \beta, \alpha, \theta)$. The marginal contribution to the likelihood, for right censored in the group $i$ is:

$$L_i(T_i, T_i^*, \phi) = \int_{u_i} L_i(T_i, T_i^* | u_i) f(u_i) du_i$$

In this expression,

- The density probability function for the random effects $u_i$ are:
  $$f(u_i) = \frac{u_i^{\theta - 1} \exp(-u_i/\theta)}{\Gamma(1/\theta)^{m+am_i+1/\theta - 1}}$$

- The conditional distribution of the survival times is the product of the individual contributions:
  $$L_i(T_i, T_i^* | u_i) = \prod_{j=1}^{N_i} r_{ij}(T_{ij}|u_i)^{\delta_{ij}} \exp(-R_{ij}(T_{ij}|u_i)) \prod_{j=1}^{N_i} \lambda_{ij}(T_{ij}^*|u_i)^{\delta_{ij}} \exp(-\Lambda_{ij}(T_{ij}^*|u_i))$$

From the two previous expressions, we can deduce the expression of the $i^{th}$ contribution to the likelihood. We then obtain the expression (2.4) of the full marginal log-likelihood using:

$$l(\phi) = \sum_{i=1}^{G} \left\{ \sum_{j=1}^{N_i} \left[ \delta_{ij} \log r_{ij}(T_{ij}) + \delta_{ij} \log \lambda_{ij}(T_{ij}^*) \right] \right\}$$

$$- \log \Gamma(1/\theta) - \frac{1}{\theta} \log \theta$$

$$+ \log \int_0^{+\infty} u_i^{(m+am_i+1/\theta - 1)} \exp \left( -u_i/\theta - u_i \sum_{j=1}^{N_i} R_{ij}(T_{ij}) - u_i^{\theta} \sum_{j=1}^{N_i} \Lambda_{ij}(T_{ij}^*) \right) du_i$$

Appendix 2

Construction of the full log-likelihood for the joint frailty model for clustered survival data possibly right censored and left truncated

Note $T_i = (T_{i1}, \ldots, T_{iN_i})$ is the observation times and $T_{i} = (T_{i1}, \ldots, T_{iN_i})$ the last follow-up times for subjects from cluster $i$ and $L_i = (t_1, \ldots, t_N)$ the corresponding left truncated times, $\phi = (r_0(.), \lambda_0(.), \beta, \alpha, \theta)$. The marginal contribution to the likelihood, for right-censored and left-truncated data in the group $i$ is:

$$L_i(T_i, T_i^*, \phi) = \int_{u_i} L_i(T_i, T_i^* | u_i) f(u_i) du_i$$

$$S_i(L_i)$$
In this expression,

- The density probability function for the random effects \( u_i \) is: 
  \[ f(u_i) = \frac{u_i^{(1/\theta - 1)} \exp(-u_i/\theta)}{\Gamma(1/\theta)^{1/\theta}}. \]

- The conditional distribution of the survival times is the product of the individual contributions:
  \[
  L_i(T_i, T_i^\ast | u_i) = \prod_{j=1}^{N_i} r_{ij}(T_{ij}|u_i)^{\delta_{ij}} \exp(-R_{ij}(T_{ij}|u_i)) \prod_{j=1}^{N_i} \lambda_{ij}(T_{ij}^\ast | u_i)^{\delta_{ij}} \exp(-\Lambda_{ij}(T_{ij}^\ast | u_i))
  \]

- The marginal survival function \( S_i \) for the time of left truncation \( \tau \) is obtained by integrated out the random effects
  \[
  S_i(L_i) = \int_{u_i} S_i(L_i|u_i) f(u_i) du_i = \frac{(1/\theta)^{(1/\theta)}}{\Gamma(1/\theta)} \times \int_{u_i} u_i^{(1/\theta - 1)} \exp(-u_i/\theta - u_i \sum_{j=1}^{N_i} R_{ij}(L_{ij}) - u_i^\ast \sum_{j=1}^{N_i} \Lambda_{ij}(L_{ij})) du_i
  \]

From the three previous expressions, we can deduce the expression of the \( i \)th contribution to the likelihood. We then obtain the expression of the full marginal log-likelihood using:

\[
I(\phi) = \sum_{i=1}^{G} \left[ \sum_{j=1}^{N_i} \delta_{ij} \log r_{ij}(T_{ij}) + \delta_{ij} \log \lambda_{ij}(T_{ij}^\ast) \right]
\]

\[
+ \log \int_{0}^{\infty} u_i^{(m_i + a_i \gamma + 1/\theta - 1)} \exp\left(-u_i/\theta - u_i \sum_{j=1}^{N_i} R_{ij}(T_{ij}) - u_i^\ast \sum_{j=1}^{N_i} \Lambda_{ij}(T_{ij}^\ast)\right) du_i
\]

\[
- \log \int_{0}^{\infty} u_i^{(1/\theta - 1)} \exp\left(-u_i/\theta - u_i \sum_{j=1}^{N_i} R_{ij}(L_{ij}) - u_i^\ast \sum_{j=1}^{N_i} \Lambda_{ij}(L_{ij}^\ast)\right) du_i
\]

**Appendix 3**

**Construction of the Kendall’s tau for the joint frailty model (2.2)**

\[
\tau = E[\text{sign}((T_{ij} - T_{ij}^\ast)(T_{ij}^\ast - T_{ij}^\ast))] = 4 \int_{0}^{\infty} \int_{0}^{\infty} f_{T_{ij}, T_{ij}^\ast}(x, y) S_{T_{ij}, T_{ij}^\ast}(x, y) dx dy - 1
\]

where \( \text{sign}(x) = -1, 0, 1 \) for \( x < 0, x = 0, x > 0 \).

One can show that \( \tau \) is given by:

\[
\tau = 4 \int_{0}^{\infty} \int_{0}^{\infty} f_{T_{ij}, T_{ij}^\ast}(x, y) S_{T_{ij}, T_{ij}^\ast}(x, y) dx dy - 1
\]

with \( f_{T_{ij}, T_{ij}^\ast}(x, y) = \int_{0}^{\infty} f(x, y|u) du \) and \( S_{T_{ij}, T_{ij}^\ast}(x, y) = \int_{0}^{\infty} S(x, y|u) du \) the marginal density and survival functions. Using the joint frailty model (2.2), this expression is equivalent to:

\[
\tau = 4 \int_{0}^{\infty} \int_{0}^{\infty} f_{T_{ij}, T_{ij}^\ast}(x, y|U_i) S_{T_{ij}, T_{ij}^\ast}(x, y|U_i) + f_{T_{ij}, T_{ij}^\ast}(x, y|U_i) S_{T_{ij}, T_{ij}^\ast}(x, y|U_i) dx dy - 1
\]

\[
= 2 \int_{0}^{\infty} \int_{0}^{\infty} \left(\lambda_{T_{ij}, T_{ij}^\ast}(x, y|U_i) S_{T_{ij}, T_{ij}^\ast}(x, y|U_i) S_{T_{ij}, T_{ij}^\ast}(x, y|U_i) \right)\right. dx dy - 1
\]

\[
\left. + \lambda_{T_{ij}, T_{ij}^\ast}(x, y|U_i) S_{T_{ij}, T_{ij}^\ast}(x, y|U_i) S_{T_{ij}, T_{ij}^\ast}(x, y|U_i) dx dy - 1
\]

\[
= 2 \int_{0}^{\infty} \int_{0}^{\infty} \left( u_i^{\theta - 1} + u_i^{\theta - 1} \right) \frac{u_i^{(1/\theta - 1)} \exp(-u_i/\theta)}{\Gamma(1/\theta)^{1/\theta}} du_i \left( u_i^{(1/\theta - 1)} \exp(-u_i/\theta) / \Gamma(1/\theta)^{1/\theta}\right) du_i - 1
\]