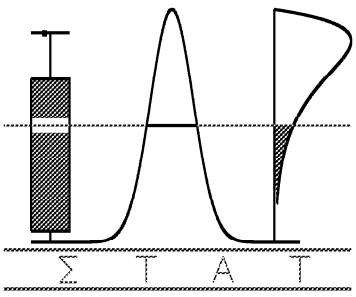
T E C H N I C A L

R E P O R T

0238

Modelling associations between time-to-event responses in pilot cancer clinical trials using a Plackett-Dale model

F. Tibaldi, F. Torres, and G. Molenberghs



I A P S T A T I S T I C S

N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

Modelling associations between time-to-event responses in pilot cancer clinical trials using a Plackett-Dale model

Fabián Tibaldi, Franz Torres Barbosa and Geert Molenberghs

Center for Statistics, Limburgs Universitair Centrum,

Universitaire Campus, B3590 Diepenbeek, Belgium.

Abstract

This work was motivated by the need to find surrogate endpoints for survival of patients in oncology studies. The goal of this article is to determine associations between five time-to-event outcomes coming from three clinical trials for non small cell lung cancer. To this end we propose to use the multivariate Dale model for time-to-event data introduced by Tibaldi *et al.* (2002). We fit the model to these data, using a pseudo-likelihood approach to estimate the model parameters.

We evaluate and compare the performance of different dimensional models and we relate the Dale model association parameter, i.e., the odds ratio, to well known quantities such as Kendall's τ and Spearman's ρ.

Finally, the results are discussed with a perspective on surrogate marker validation. Some suggestions are made regarding further studies in this field.

1 INTRODUCTION

Survival time of patients is one of the most common outcomes when assessing response to treatment in cancer clinical trials. While tumor response or percentage of tumor shrinkage has been used as a surrogate endpoint for cytotoxic drugs, it has been questioned at several occasions [1, 2, 3]. There is a need to detect potential surrogate endpoints to decreases costs, time, and/or to improve the quality of life of cancer patients. Appropriate models, considering the type of response (continuous, binary, time-to-event, etc.) have to be proposed and applied to this effect.

In this contribution, we use a multivariate survival model to estimate associations between time-to-event responses, to explore surrogacy of candidate markers, potentially after adjustment for other factors. The model used here has the advantage that its association parameter, the odds ratio, can be translated without difficulty into quantities that are considered easier to interpret, such as Spearman's rank correlation coefficient ρ or Kendall's τ. Appropriate hypothesis tests can be applied to assess the strength of the association.

Survival-type models using copulas were developed [4] and extended [5] to the multivariate case by using pseudo-likelihood estimation of the parameters.

In Section 2, a pilot study in cancer vaccination for non small cell lung cancer

(NSCLC) patients is described. The focus is the assessment of the association between five time-to-event outcomes, one of which can be considered the true endpoint from a surrogate marker point of view. The statistical model and pseudo-likelihood estimation of its parameters are presented in Section 3. The analysis of the data is presented in Section 4.

2 CLINICAL TRIALS FOR NON SMALL CELL LUNG CANCER

Three pilot clinical trials were performed, with the aim of testing safety, immunogenicity, and survival of a therapeutic vaccine based on the epidermal growth factor (EGF) molecule in patients with advanced non small cell lung cancer (NSCLC) [6, 7]. A first pilot study tested the vaccine in 20 patients, with NSCLC, randomized to the EGF vaccine with two different adjuvants Alum and Montanide ISA-51. The vaccine was administered in a 5 doses schedule for 51 days. Immunogenicity data were collected weekly during the treatment period and monthly during follow-up. The second pilot trial studied the same vaccines in an additional 20 patients, but with common 3 days pre-treatment with cyclosfosfamide. In the third trial, 21 patients were assigned randomly to two different EGF vaccine doses.

In all three trials, the scope of patients is reduced to very advanced cancer patients at stages III, IIIb or IV without any other alternative of oncospecific treatment, with ECOG performance status less than 3. Survival time was con-

sidered from the day of random treatment assignment until the day of death, regardless of its cause. There were three participating hospitals. The mechanism of vaccine activity ought to induce an anti-tumoral immune response. Time to a good immune response could be an indicator of a possible clinical effect. The quality of the immune response is assessed by its titer and the titer ratio with respect to the baseline value.

We consider five time-to-event outcomes. Time 1 is time to response inmunogenecity 2X; Time 2 is time to response of inmunogenecity 1:2000 and 4X; Time 3 is time to maximum titer; TTP is time to progression and TSV is overall survival time. The latter is the true endpoint whereas the earlier four are potential surrogate endpoints. All times are expressed in months. Available covariate information includes age (on a continuous scale), disease stage (categories III, IIIb, and IV), indicator for patient's previous treatment (e.g., chemotherapy), and, of course, treatment assignment.

In a previous analysis, a relationship between immunological response and survival time was detected [8]. For one of the trials, there was a clear advantage on survival for the group of high immunological responders [9].

3 STATISTICAL MODEL

We will introduce the multivariate Dale model for survival data and combine it with pseudo-likelihood ideas to estimate the parameters. To this end, let us consider a trial involving N subjects with k time–to–event measurements. In

our case study, $k = 5$ with times Time 1, Time 2, Time 3, TTP, and TSV.

Suppose that we also observe a vector of covariates $\boldsymbol{Z}$ and assume a Weibull distribution for each time T_j with λ_{T_j} and p_{T_j} the scale and shape parameters, respectively. While we will focus on Weibull marginals, choosing different univariate marginal survival distributions will not induce additional complexities. The choice of marginal survivorship functions will, of course, impact the fit of the marginal outcomes but is expected to have less impact on the estimated values of the association parameters. Express the observed information on individual i in the format: $(T_{i1}, \ldots, T_{ik}, \Delta_{i1}, \ldots, \Delta_{ik}, z_{i1}, \ldots, z_{in_k})$ so that $\boldsymbol{W}_{ij} = (T_{ij}, \Delta_{ij}, \boldsymbol{Z}_i)$ are the values for a particular subject i and time point j, with $j = 1, \ldots, k$.

Let us briefly introduce the concept of pseudo-likelihood estimation. A full introduction in the context of survival outcomes is given in Tibaldi *et al.* [5]. It is well known that full maximum likelihood estimation can become prohibitive for many (marginal) models. For example, in the framework of a marginally specified odds ratio model [10, 11, 12, 13, 14] for multivariate, clustered binary data, full maximum likelihood estimation is extremely computer intensive, especially with large within-unit sizes. Hence, alternative estimation methods, not requiring full specification of the joint distribution, are in demand. The principal idea behind pseudo-likelihood is to replace a numerically challenging joint density by a simpler function that is a suitable product of ratios of likelihoods of subsets of the variables. For example, when a joint density contains a computationally intractable normalizing constant, one might calculate a suit-

able product of conditional densities which does not involve such a complicated function. A bivariate distribution $f(y_1, y_2)$, for example, can be replaced by the product of both conditionals $f(y_1|y_2)f(y_2|y_1)$. While the method achieves important computational economies by changing the method of estimation, it does not affect model interpretation. Model parameters can be chosen in the same way as with full likelihood and retain their meaning. This method converges quickly with only minor efficiency losses, especially for a range of realistic parameter settings. In our case, the k-variate joint distribution will be replaced by the product of all of its pairwise margins. It was shown [15] that this and related procedures are consistent and asymptotically normally distributed.

The pseudo-likelihood function constructed for the estimation of the parameters of this model is based on considering all (in our case, ten) possible pairs of outcomes on an individual $(\boldsymbol{W}_{ir}, \boldsymbol{W}_{i\ell})$ with $1 \leq r < \ell \leq 5$. These pairs produce $f_{T_r T_\ell}(\boldsymbol{W}_{ir}, \boldsymbol{W}_{i\ell})$, where $f_{T_r T_\ell}$ is the density function of the Plackett-Dale distribution.

The Plackett distribution is obtained for a constant cross-ratio $\theta_{r\ell}(t_r, t_\ell) \equiv \theta$ [11, 16, 17]. Details are given in the Appendix. The joint distribution $F_{T_r T_\ell}$ is defined by means of (1) and (2), as soon as F_{T_r}, F_{T_ℓ} and θ_{rl} are known. In this case, the dependence can be defined using a *global cross-ratio* at (t_r, t_ℓ) given by $\theta_{r\ell}(t_r, t_\ell)$:

$$F_{T_r,T_\ell}(t_r, t_\ell) = \begin{cases} \dfrac{1 + (F_{T_\ell}(t_\ell) + F_{T_r}(t_r))(\theta_{rl} - 1) - H(F_{T_\ell}(t_\ell), F_{T_r}(t_r), \theta_{r\ell})}{2(\theta_{r\ell} - 1)} & \text{if } \theta_{r\ell} \neq 1, \\ F_{T_\ell}(t_\ell) F_{T_r}(t_r) & \text{if } \theta_{r\ell} = 1, \end{cases} \tag{1}$$

where

$$H(F_{T_r}, F_{T_\ell}, \theta_{r\ell}) = \sqrt{((1 + ((\theta_{r\ell} - 1)[F_{T_r}(t_r) + F_{T_\ell}(t_\ell)])^2 + 4\theta_{r\ell}(1 - \theta_{r\ell})F_{T_r}(t_r)F_{T_\ell}(t_\ell)}. \tag{2}$$

Based upon this distribution function, we can derive a bivariate Plackett *density* function, $f_{T_rT_\ell}(t_r, t_\ell)$, for two survival times using (1)–(2) by calculating $\partial F_{T_rT_\ell}(t_r, t_\ell)/\partial t_r \partial t_\ell$, thereby properly accounting for censoring information.

Precisely, we can define the pseudo-likelihood function PL through its logarithm

$$p\ell(\mathbf{\Phi}) = \sum_{i=1}^{N} p\ell_i, \tag{3}$$

with

$$p\ell_i = \sum_{(r,\ell)\in S} \ln f_{T_rT_\ell}(\boldsymbol{W}_{ir}, \boldsymbol{W}_{i\ell}, \mathbf{\Phi}),$$

where $S = \{(1,2),(1,3),(1,4),(1,5),(2,3),(2,4),(2,5),(3,4),(3,5),(4,5)\}$ is the set of all ten possible pairs of outcomes, $f_{T_rT_\ell}$ is the value of the function defined earlier and evaluated in the corresponding outcomes for subject i, and $\mathbf{\Phi}$ is the vector of parameters. Specifically, $\mathbf{\Phi}' = (\boldsymbol{\theta}', \boldsymbol{\beta}_T', \boldsymbol{\lambda}_T', \boldsymbol{p}_T')$ with $\boldsymbol{\theta}$ the subvector of association parameters, $\boldsymbol{\beta}_T$ the subvector of coefficients corresponding to the covariates $\boldsymbol{z}$ and, $\boldsymbol{\lambda}_T$ and $\boldsymbol{p}_T$ subvector of parameters from the Weibull distribution.

The pseudo-likelihood estimator $\widehat{\mathbf{\Phi}}$ is defined as the maximizer of (3). Consistency has been shown [15, 18, 19]. Correct estimates of precision can be obtained using sandwich estimator ideas, not dissimilar to those proposed by Liang and Zeger [20]. A brief outline of the estimation procedure is provided in the Appendix.

The Plackett-Dale model allows us to estimate and interpret the strength of the association between a pair of survival times via global cross ratios (the θ parameters in the model). It is often convenient to consider a transformation of θ that has the interpretational properties of a correlation coefficient, such as Spearman's ρ or Kendall's τ .

Kendall's τ lies in the $[-1, 1]$ interval and a zero value implies independence between T_r and T_l. There exists a relationship between Kendall's τ and θ for any copula $C(t_r, t_l, \theta)$ [21]

$$\tau(\theta) = 4 \int_0^1 \int_0^1 C_{T_r T_l}(t_r, t_l, \theta) C_{T_r T_l}(dt_r, dt_l, \theta) - 1. \tag{4}$$

This relationship is independent of the marginal distributions and only depends on $F_{T_r T_l}$ [22]. Kendall's τ thus measures the association between both time points after adjustment for the covariates used in the model. Estimates and confidence intervals, using the delta method, are accordingly easily obtained. There is no closed form for Kendall's τ in the Plackett-Dale case and an estimate has to be obtained from (4). We have developed a SAS IML 8.02 macro to this effect.

Spearman's ρ is also independent of the margins, and belongs to the unit interval. The relationship between Spearman's ρ and θ is

$$\rho(\theta) = \frac{\theta + 1}{\theta - 1} - \frac{2\theta . \ln \theta}{(\theta - 1)^2}. \tag{5}$$

An estimate follows from $\rho = \rho(\widehat{\theta})$, with variance estimated using a straightforward application of the delta method. This allows estimation of the associations between the five outcomes by fitting a multivariate model and adjusting

for other variables as age of the patients, previous treatment status, stage of the diseases, etc, as we will see in Section 4.

Pseudo-likelihood estimates were obtained using Newton-Raphson with analytical first derivatives and numerical second derivatives, as implemented in SAS IML 8.02 and using routine NLPNRR (SAS Institute Inc. 1999–2001). Standard errors of the parameters were calculated using the inverse of the observed matrix of second derivatives.

This model has important implications in the assessment of surrogacy. In previous studies [4], the validation of a new variable as surrogate was performed on only one surrogate and only one true endpoint. In our case, the model allows to study several surrogates and several true endpoints at the same time. It gives also the possibility of developing new strategies not only to validate already identified candidates, but also to identify new variables that have potential regarding surrogacy.

Both Kendall's τ and Spearman's ρ quantities can serve as an indication of individual level surrogacy in the sense of Buyse *et al.* [23]. In case data are available from a sufficciently large number of trials and/or centers, these authors' meta-analytic perspectives can be adopted as well.

4 ANALYSIS OF THE DATA

We will now fit the proposed model to the data described in Section 2. Even when the association between outcomes is of primary scientific interest, as is

the case here, it is mandatory to appropriately adjust the marginal survival regressions for covariate effects. We have included patients' characteristics: age (as a continuous variable), disease stage (three categories labelled III, IIIb, and IV), whether or not a patient received previous treatment (e.g., chemotherapy), and treatment arm. The time unit for the outcomes was months.

We will use the indices 1, 2, 3, 4 and 5 to identify the outcomes Time 1, Time 2, Time 3, TTP and TSV, respectively. Thus, for example, θ_{15} denotes the association between outcomes Time 1 and TSV. Note that the models of primary interest are those containing the variable TSV, considered to be the true endpoint in this study. Nevertheless, the other models are useful to further insight into the association structure.

In the first part of the analysis we explored the importance of hospital and trial to estimate the pairwise associations. We fitted all possible bivariate models using as covariates *age*, *stage*, *prevtrt* and *group* in four different situations. Firstly, we fitted models with the variables *hospital* and *trial*, secondly with the variable *hospital* only, thirdly with the variable *trial* only, and fourthly with neither of these variables. The results, not shown here, reveal that no large differences were observed between the association parameters across the four choices, so it was decided to retain the simplest model and both *trial* and *hospital* were dropped from further consideration.

We first considered all possible bivariate models (1B to 10B) and all different trivariate models (1T to 10T). The association parameters obtained from these models, as well as those from the five-variate model (1F), are presented in

Table 1. The primary use of the bivariate and trivariate models lies in their comparison with the full 5-variate model. Indeed, given the marginal nature of the models, corresponding associations have the same meaning. While each association occurs only once in the collection of bivariate models, they do so several times in the collection of trivariate models, disallowing their easy use. Similarly, each association is used only once in the full 5-variate model. The most obvious advantage is that all associations feature within a single, integrated model. They are also estimated with increased precision as opposed to their bivariate and trivariate counterparts. The bivariate models are also useful to provide starting values for the 5-variate model. Indeed, the model is not easy to bring to convergence in the absence of reasonable starting values.

Let us zoom in on the comparison of association parameters across models. For example, the association between TTP and TSV, θ_{45}, can be found from Models 10B, 1T, 2T, 3T, and 1F. The results are very similar, as can be seen in most other rows in Table 1, with somewhat exceptional behavior for θ_{12} and θ_{13}. Such behavior is not uncommon for relatively large odds ratios, and the difference is less prominent on the log odds ratio scale.

Full details of the parameter estimates from the 5-variate model are given in Tables 2 and 3. Table 2 described the association parameters. Apart from the original odds-ratio scale (θ parameters), the easier-to-interpret Kendall's τ and the Spearman's ρ coefficient are included, together with asymptotic 95% confidence intervals. θ-confidence intervals not containing one provide evidence for association between the corresponding pair of times, after correction for the

covariates. Note that the covariates and other marginal regression parameters are displayed in Table 3. A corresponding association assessment based on Kendall's τ and Spearman's ρ requires exclusion of the zero value from the corresponding confidence intervals.

Several substantive conclusions can be drawn from the model fits. From Model 1F we see that the highest association is observed between TTP and TSV. TSV is also significantly associated with Time 2 and we further observe a significant association between Time 1 and Time 2. While the first two of these three associations are of direct interest, and may lead to reconsideration of Model 2T (containing Time 2, TTP, and TSV), it is of interest to consider a 4-variate model as well, i.e., a model with outcomes 1, 2, 4, and 5 (i.e., Time 1, Time 2, TTP, and TSV). Indeed, through its association with Time 2, Time 1 may indirectly contribute useful information. In any case, Time 3 appears to have no association with any of the other outcomes. Thus, a 4-variate model as presented in Table 4 will be considered our final model.

In summary, we have some evidence that TTP and Time 2 can be used as surrogates for TSV, with some auxiliary information coming from Time 1. Of course, the evidence apported here is just from three relatively small trials, and is based on an assessment of the association between responses only. Clearly, more exhaustive studies need to be designed in order to evaluate the surrogacy in a more authoritative fashion, preferably in a meta-analytic setting such as the one proposed by Buyse *et al.* [23] or Burzykowski *et al.* [4].

5 CONCLUDING REMARKS

We have proposed the use of a multivariate Plackett-Dale model for estimating associations between, possibly censored, time-to-event outcomes. Specifically, we showed how this methodology can be useful in the context of surrogate marker validation.

Given the difficulties of manipulating the likelihood function in this case, a pseudo-likelihood approach was undertaken as a viable and attractive alternative to maximum likelihood. The computational complexity of the algorithm used for the estimation of the model parameters was overcome by using initial values obtained from the bivariate fitted models. Good numerical results were obtained in most cases.

Kendall's τ and Spearman's ρ coefficients can be used as measures of individual level surrogacy [4]. In spite of the multivariate flavor of this model, the pairwise pseudo-likelihood approach provides only bivariate association measures. Valid confidence intervals for such quantities were constructed using the delta method.

One of the primary purposes of this study was to detect or identify possible new surrogate endpoints for survival time. We are particularly interested in the validation of four different surrogate variables (Time 1, Time 2, Time 3 and TTP). This implies the need of a multivariate model considering all of these surrogates and the true endpoint.

We want to note that the methodology we applied here focuses only on the individual level surrogacy but similar ideas as in Buyse *et al.* [23] for the

meta-analytic framework need to be developed further.

Using the selected 4-variate model a high association between Time 1 and Time 2 can be observed. This evidenced that the time of reaching the double baseline titer, for most of the patients, had a strong relationship with the time to achieve a high titer value (1:2000 and 4X).

None of the times to reach a good immune response seem to have a high association with TTP, neither with survival time. It seems that, with the accumulated evidences in this patient population, time to a good immune response is not a strong surrogate of clinical benefit endpoints such as TTP and survival. Other immune information seems to be more important and this should be the objective of further research.

However, there is evidence that TTP is highly associated with survival time. In practice, this variable is not very convienient given its closeness to the actual survival time. The marginal gain does not justify its use as a surrogate.

Acknowledgements

The first author wishes to thank "Bijzonder Onderzoeksfonds" of the Limburgs Universitair Centrum. The authors thanks Dr. Tania Crombet and Dr. Gisela Gonzáles for providing the data of the pilot study in cancer vaccination for non small cell lung cancer (NSCLC) we used in this paper. We are all grateful to the Interuniversity Attraction Poles Programme P5/24 - Belgian State - Federal Office for Scientific, Technical and Cultural Affairs.

References

[1] Anderson, J.R., Cain, K.C., and Gelber, R.D. (1983) Analysis of survival by tumour response. *Journal of Clinical Oncology*, **1**, 710–719.

[2] Ellenberg, S.S. and Hamilton, J.M. (1989) Surrogate endpoints in clinical trials: cancer. *Statistics in Medicine*, **8**, 405–413.

[3] Buyse, M. and Molenberghs, G. (1998) The validation of surrogate endpoints in randomized experiments. *Biometrics*, **54**, 186–201.

[4] Burzykowski, T., Molenberghs, G., Buyse, M., Geys, H., and Renard, D. (2001). Validation of surrogate endpoints in multiple randomized clinical trials with failure-time endpoints. *Applied Statistics* **50**, 405–422.

[5] Tibaldi, F., Molenberghs, G., Burzykowski, T., Geys, H. (2002) Pseudo-likelihood estimation for a marginal multivariate survival model *Submitted for publication*

[6] González, G., Crombet, T., Catalá, M., Mirabal, V., Hernández, J., González, Y., Marinello, P., Guillén, G., and Lage, A. (1998). A novel cancer vaccine composed of human-recombinant epidermal growth factor linked to a carrier protein: Report of a pilot clinical trial. *Annals of Oncology*, **9**, 431–435.

[7] González, G., Crombet, T., Torres, F., Catalá, M., Alfonso, L., Osorio, M., Neninger, E., García, B., Mulet, A., Pérez, R., Lage, A. (2002). Epidermal

growth factor based cancer vaccine for non small cell lung cancer therapy. *Annals of Oncology*, 000 000.

[8] Torres, F., Gonzàlez, G., Crombet, T., and Lage, A. (2001) Empirical Bayes analysis of early clinical trials for cancer vaccines. *Submitted for publication.*

[9] Torres, F., Tibaldi, F., Cortiñas Abrahantes, J., Geys, H., González, G., and Molenberghs, G. (2002) Surrogacy Evaluation of Immunological Parameters in Pilot Cancer Clinical Trials. *Submitted for publication.*

[10] Lipsitz, S.R., Laird, N.M., and Harrington, D.P. (1991). Generalized estimating equations for correlated binary data: using the odds ratio as a measure of association. *Biometrika*, **78**, 153–160.

[11] Dale, J. R. (1986). Global cross–ratio models for bivariate, discrete, ordered responses. *Biometrics* **42**, 909–917.

[12] Molenberghs, G. and Lesaffre, E. (1994). Marginal modelling of correlated ordinal data using a multivariate Plackett distribution. *Journal of the American Statistical Association* **89**, 633–644.

[13] Glonek, G.F.V. and McCullagh, P. (1995). Multivariate logistic models. *Journal of the Royal Statistical Society, Series B*, **57**, 533–546.

[14] Lang, J.B. and Agresti, A. (1994). Simultaneously modeling joint and marginal distributions of multivariate categorical responses. *Journal of the American Statistical Association*, **89**, 625–632.

[15] Geys, H., Molenberghs, G. and Ryan, L. (1999). Pseudo-likelihood modelling of multivariate outcomes in developmental toxicology. *Journal of the American Statistical Association* **94**, 34–745.

[16] Plackett, R. L. (1965). A class of bivariate distributions. *Journal of the American Statistical Association* **60**, 516–522.

[17] Mardia, K.V. (1970). *Families of Bivariate Distributions.* London: Griffin.

[18] Arnold, B.C. and Strauss, D. (1991). Pseudo likelihood estimation: some examples. *Sankhya B* **53**, 233–243.

[19] Le Cessie, S. and Van Houwelingen, J.C. (1994). Logistic regression for correlated binary data. *Applied Statistics* **43**, 95–108.

[20] Liang, K.Y. and Zeger, S.L. (1986).Longitudinal data analysis using generalized linear models. *Biometrika* **50**, 13–22.

[21] Genest, C. and McKay, J. (1986). The joy of copulas: bivariate distributions with uniform marginals. *American Statistician* **40**, 280–283.

[22] Schweizer, B. and Wolff, E.F. (1981). On nonparametric measures of dependence for random variables. *Annals of Statistics* **9** 879–885.

[23] Buyse, M., Thirion, P., Carlson, R.W., Burzykowski, T., Molenberghs, G., and Piedbois, P. (2000) Tumour response to first line chemotherapy improves the survival of patients with advanced colorectal cancer. *Lancet*, **356**, 373–378.

[24] Nelsen, R.G. (1999). An introduction to copulas. *Lecture Notes in Statistics,* **139**. New York: Springer-Verlag.

[25] Cox, D.R. (1972). Regression models in life-tables (with discussion). *Journal of the Royal Statistical Society, Series B* **34**, 187–220.

[26] Lehmann, E.L. (1983). *Theory of Point Estimation.* New York: John Wiley.

[27] Shih, J.H.and Louis, T.A. (1995).Inferences on association parameter in copula models for bivariate survival data. *Biometrics* **51**, 1384–1399.

[28] Fréchet, M. (1951). Sur les tableaux de corrélation dont les marges sont données. *Annals Université de Lyon, Section A, Series 3*, **14**, 143–153.

[29] Aerts, M., Geys, H., Molenberghs, G., and Ryan, L. (2002). *Topics in Modelling of Clustered Data.* London: Chapman & Hall.

[30] Geys, H., Molenberghs, G., and Lipsitz, S.R. (1998). A note on the comparison of pseudo-likelihood and generalized estimating equations for marginal odds ratio models. *Journal of Statistical Computation and Simulation* **62**, 45–72.

[31] Fleming, T. R., Harrington D. P. (1991). *Counting processes and survival analysis* New York: John Wiley.

[32] Oakes, D. (1989). Bivariate survival models induced by frailities. *Journal of the American Statistical Association* **84**, 487–493.

[33] Molenberghs, G. and Lesaffre, E. (1999). Marginal modelling of multivariate categorical data. *Statistics in Medicine* **18**, 2237–2255.

[34] Joe, H. (1997). *Multivariates Models and Dependence Concepts.* London: Chapman and Hall.

[35] Wang, W. and Wells, M.T. (2000). Model selection and semiparametric inferences for bivariate failure-time data. *Journal of the American Statistical Association*, **95**, 62–76.

[36] Verbeke, G. and Molenberghs, G. (2000) *Linear Mixed Model for Longitudinal Data.* New York: Springer-Verlag.

[37] Burzykowski, T., Molenberghs, G, and Buyse, M. (2001) The validation of surrogate endpoints using data from randomized clinical trials: A case study in advanced colorectal cancer. *Submitted for publication.*

[38] Buyse, M. and Piedbois, P. (1996) On the relationship between response to treatment and survival. *Statistics in Medicine*, **15**, 2797–2812.

[39] Freedman, L.S., Graubard, B.I., Schatzkin, A. (1992) Statistical validation of intermediate endpoints for chronic diseases. *Statistics in Medicine*, **11**, 167–178.

APPENDIX: MODEL DESCRIPTION

Assume that T_1 and T_2 are correlated survival times, then the joint survival function of (T_1, T_2) can be written as

$$S_{T_1T_2}(t_1,\, t_2) = P(T_1 \geq t_1,\, T_2 \geq t_2) = C_{\theta_{12}}\{S_{T_1}(t_1),\, S_{T_2}(t_2)\}, \qquad t_1,\, t_2 \geq 0, \tag{6}$$

where S_{T_1} and S_{T_2} denote marginal survival functions and $C_{\theta_{12}}$ is a copula [24]. An attractive feature of model (6) is that the margins do not depend on the choice of the copula function.

To model the effect of specific covariates on the marginal distributions of T_1 and T_2 in (6) we propose to use the proportional hazard model:

$$S_{T_1}(t_1) = \exp\left\{-\int_0^{t_1} h_{T_1}(x)\exp(\boldsymbol{\beta}_{T_1}\boldsymbol{Z}_1)dx\right\}, \tag{7}$$

$$S_{T_2}(t_2) = \exp\left\{-\int_0^{t_2} h_{T_2}(x)\exp(\boldsymbol{\beta}_{T_2}\boldsymbol{Z}_2)dx\right\}, \tag{8}$$

where h_{T_1} and h_{T_2} are marginal baseline hazard functions and $\boldsymbol{\beta}_{T_1}$ and $\boldsymbol{\beta}_{T_2}$ are vectors of unknown regression parameters corresponding to the covariates $\boldsymbol{Z}$. The classical model proposed by Cox [25] is used for the hazard functions. Estimates of the parameters for joint model (6)–(8) can be obtained using the method of maximum likelihood [26] or the two-stage parametric procedure proposed by Shih and Louis [27]. In this case the dependence can be defined using a *global cross-ratio* at (t_1, t_2) given by

$$\theta_{12}(t_1, t_2) = \frac{F(t_1, t_2)[1 - F_{T_1}(t_1) - F_{T_2}(t_2) + F(t_1, t_2)]}{[F_{T_1}(t_1) - F(t_1, t_2)][F_{T_2}(t_2) - F(t_1, t_2)]}. \tag{9}$$

where F_{T_1} and F_{T_2} are the marginal cumulative density functions. Here, $\theta_{12} =$

$\theta_{12}(t_1, t_2)$ satisfies $0 \leq \theta_{12} \leq \infty$ when $F(t_1, t_2)$ satisfies the Fréchet-Hoeffding [28] bounds. For a constant cross-ratio, $\theta_{12}(t_1, t_2) \equiv \theta$, the Plackett distribution is obtained [16, 17].

The values of the Plackett distribution are found as one of the two solutions of the following second degree polynomial equation if the marginal distribution functions F_{T_1} and F_{T_2}, and the cross-ratio θ_{12} are known:

$$\theta_{12}(F - F_{T_1})(F - F_{T_2}) - F[F - (F_{T_1} + F_{T_2} - 1)] = 0. \tag{10}$$

Dale and Mardia [11, 17] gave an explicit solution for (10) and showed that $F_{T_1,T_2}(t_1, t_2)$ is always a bivariate copula, with θ_{12} in $[0, +\infty]$.

$$F_{T_1,T_2}(t_1, t_2) = \begin{cases} \dfrac{1 + (F_{T_2}(t_2) + F_{T_1}(t_1))(\theta_{12} - 1) - H(F_{T_2}(t_2), F_{T_1}(t_1), \theta_{12})}{2(\theta_{12} - 1)} & \text{if } \theta_{12} \neq 1, \\ F_{T_2}(t_2)F_{T_1}(t_1) & \text{if } \theta_{12} = 1, \end{cases} \tag{11}$$

where

$$H(F_{T_1}, F_{T_2}, \theta_{12}) = \sqrt{(1 + (\theta_{12} - 1)(F_{T_1}(t_1) + F_{T_2}(t_2)))^2 + 4\theta_{12}(1 - \theta_{12})F_{T_1}(t_1)F_{T_2}(t_2)}. \tag{12}$$

Mardia [17] showed that $F_{T_1,T_2}(t_1, t_2)$ is always a bivariate copula, with θ_{12} in $[0, +\infty]$. For our pseudo-likelihood approach the resulting PL function and its log will be denoted by

$$\ln p\ell(\boldsymbol{\Phi}) = \sum_{i=1}^{N} \sum_{(r,\ell) \in S} \ln f_{T_r T_\ell}(\boldsymbol{W}_{ir}, \boldsymbol{W}_{i\ell}, \boldsymbol{\Phi}), \tag{13}$$

where S is the set of indices with all possible pairs of outcomes of interest, $f_{T_r T_\ell}$

is the value of the function that results by derivating the expression (11) and $\mathbf{\Phi}$ is the vector of parameters.

Maximizing (13) we obtain the pseudo-likelihood estimator $\widehat{\mathbf{\Phi}}$. Geys, Molenberghs, and Ryan; Arnold and Strauss; Le Cessie and Van Houwelingen [15, 18, 19] showed that it converges in probability to $\mathbf{\Phi}_0$, the true parameter value and $\sqrt{N}(\widehat{\mathbf{\Phi}} - \mathbf{\Phi}_0)$ converges in distribution to $N_q(\mathbf{0}, J(\mathbf{\Phi}_0)^{-1}K(\mathbf{\Phi}_0)J(\mathbf{\Phi}_0)^{-1})$ with $J(\mathbf{\Phi})$ and $K(\mathbf{\Phi})$ defined by

$$J_{rl} = \sum_{(s,t)\in S} E_{\mathbf{\Phi}}\left(\frac{\partial^2 \ln f_{T_sT_t}(t_{is}, t_{it})}{\partial\phi_r \partial\phi_l}\right) \tag{14}$$

$$K_{rl} = -\sum_{(s,t)\in S} E_{\mathbf{\Phi}}\left(\frac{\partial \ln f_{T_s}(t_{is}, t_{it})}{\partial\phi_r}\frac{\partial \ln f_{T_t}(t_{is}, t_{it})}{\partial\phi_l}\right). \tag{15}$$

This result provides an easy way to estimate consistently the asymptotic covariance matrix. Indeed, the matrix J is found from evaluating the second derivate of the log $p\ell$ function at the PL estimate. The expectation in K can be replaced by the cross-product of the observed scores. We will refer to J^{-1} as the model based variance estimator, which should not be used as such because it overestimates precision; to K as the empirical correction; and $J^{-1}KJ^{-1}$ as the empirically corrected variance estimator. Several tests as pseudo-likelihood ratio and pseudo-score test statistics [29] can be performed. As discussed by Arnold and Strauss [18] a PL estimator is always less efficient than the corresponding ML estimator but Aerts *et al.* [29] showed that in many realistic settings efficiency losses are minor.

Table 1: *Comparison of the association parameters obtained from bivariate (B), trivariate (T),and five-variate(F) models.*

	Model											
Par.	1T	2T	3T	4T	5T	6T	7T	8T	9T	10T	1F	1B-10B
θ_{12}	–	–	–	11.10(6.89)	–	–	10.96(0.46)	–	–	11.09(6.98)	**8.82(2.79)**	17.52(8.85)
θ_{13}	–	–	–	–	–	5.58(5.04)	–	5.61(5.03)	–	5.81(5.12)	**4.79(2.28)**	8.93(5.14)
θ_{14}	0.84(0.48)	–	–	–	–	–	0.86(0.46)	0.85(0.46)	–	–	**0.86(0.26)**	0.83(0.41)
θ_{15}	0.71(0.30)	–	–	0.71(0.28)	–	0.72(0.29)	–	–	–	–	**0.72(0.19)**	0.68(0.31)
θ_{23}	–	–	–	–	1.55(0.76)	–	–	–	1.56(0.76)	1.64(0.90)	**1.57(0.41)**	1.58(0.68)
θ_{24}	–	1.04(0.60)	–	–	–	–	1.06(0.46)	–	1.07(0.61)	–	**1.05(0.37)**	1.07(0.50)
θ_{25}	–	0.55(0.29)	–	0.51(0.28)	0.53(0.28)	–	–	–	–	–	**0.55(0.17)**	0.51(0.24)
θ_{34}	–	–	1.06(0.42)	–	–	–	–	1.06(0.43)	1.05(0.41)	–	**1.06(0.28)**	1.05(0.39)
θ_{35}	–	–	1.86(0.73)	–	1.90(0.75)	1.94(0.80)	–	–	–	–	**1.90(0.52)**	1.91(0.76)
θ_{45}	11.22(5.09)	11.17(5.10)	11.16(5.06)	–	–	–	–	–	–	–	**10.56(3.31)**	11.93(4.67)

Table 2: *Pseudo-likelihood estimates of the association parameters (confidence intervals) of the five-variate model, with outcomes Time1, Time2, Time3, TTP, and TSV. Apart from the original odds ratio scale, Kendall's τ and Spearman's ρ are presented.*

(i,j)	θ_{ij}	Kendall's τ_{ij}	Spearman's ρ_{ij}
$(1,2)$	8.821 (3.363;14.280)	0.454 (0.426;0.482)	0.628 (0.497;0.759)
$(1,3)$	4.790 (0.325;9.255)	0.337 (0.290;0.384)	0.483 (0.238;0.727)
$(1,4)$	0.857 (0.356;1.358)	-0.034 (-0.067;-0.002)	-0.051 (-0.246;0.143)
$(1,5)$	0.716 (0.348;1.083)	-0.074 (-0.103;-0.046)	-0.111 (-0.280;0.058)
$(2,3)$	1.565 (0.766;2.363)	0.099 (0.071;0.127)	0.148 (-0.019;0.315)
$(2,4)$	1.045 (0.311;1.779)	0.010 (-0.029;0.049)	0.015 (-0.219;0.249)
$(2,5)$	0.545 (0.209;0.881)	-0.134 (-0.168;-0.100)	-0.200 (-0.398;-0.002)
$(3,4)$	1.060 (0.521;1.599)	0.013 (-0.015;0.041)	0.019 (-0.150;0.189)
$(3,5)$	1.896 (0.882;2.910)	0.141 (0.112;0.171)	0.210 (0.039;0.381)
$(4,5)$	10.567 (4.088;17.046)	0.487 (0.460;0.514)	0.665 (0.544;0.785)

Table 3: *Pseudo-likelihood estimates (standard errors) of the survival regression parameters in the five-variate model with outcomes Time1, Time2, Time3, TTP, and TSV.*

Parameters	k				
	1	2	3	4	5
age_k	0.404 (0.077)	0.143 (0.060)	-0.103 (0.072)	-0.106 (0.057)	-0.220 (0.097)
$stage1_k$	0.746 (0.209)	-0.196 (0.216)	0.235 (0.230)	-0.220 (0.146)	-0.143 (0.194)
$stage2_k$	-0.789 (0.252)	-0.903 (0.241)	-0.472 (0.270)	0.122 (0.180)	-0.007 (0.241)
$prvtrt_k$	0.001 (0.158)	-0.065 (0.137)	-0.326 (0.124)	-0.420 (0.124)	0.004 (0.156)
trt_k	0.538 (0.165)	1.251 (0.162)	0.310 (0.142)	-0.208 (0.118)	-0.039 (0.141)
p_k	1.230 (0.053)	0.903 (0.039)	1.184 (0.039)	1.085 (0.041)	1.638 (0.066)
λ_k	-2.659 (0.438)	-2.901 (0.551)	-0.665 (0.412)	-0.599 (0.284)	-1.539 (0.335)

Table 4: *Pseudo-likelihood estimates (standard errors) of the survival regression and associaton parameters in four-variate model with outcomes Time1, Time2, TTP, and TSV.*

Par.	Time1-Time2-TTP-SVT
θ_{12}	9.441 (5.417)
θ_{14}	0.856 (0.245)
θ_{15}	0.712 (0.184)
θ_{24}	1.041 (0.558)
θ_{25}	0.543 (0.189)
θ_{45}	10.820 (3.559)
age_1	0.424 (0.112)
age_2	0.141 (0.118)
age_4	-0.109 (0.071)
age_5	-0.209 (0.100)
$stage1_1$	0.826 (0.261)
$stage1_2$	-0.164 (0.341)
$stage1_4$	-0.216 (0.201)
$stage1_5$	-0.140 (0.196)
$stage2_1$	-0.758 (0.287)
$stage2_2$	-0.904 (0.333)
$stage2_4$	0.133 (0.216)
$stage2_5$	-0.022 (0.247)
$prvtrt_1$	0.001 (0.199)
$prvtrt_2$	-0.072 (0.227)
$prvtrt_4$	-0.422 (0.153)
$prvtrt_5$	0.008 (0.171)
trt_1	0.531 (0.185)
trt_2	1.240 (0.272)
trt_4	-0.203 (0.143)
trt_5	-0.057 (0.155)
p_1	1.231 (0.053)
p_2	0.881 (0.057)
p_4	1.081 (0.055)
p_5	1.630 (0.074)
λ_1	-2.784 (0.697)
λ_2	-2.928 (1.119)
λ_4	-0.593 (0.391)
λ_5	-1.561 (0.336)