STATISTICS IN MEDICINE, VOL. 12, 835–842 (1993)

EVALUATING THE ROLE OF CD4-LYMPHOCYTE COUNTS AS SURROGATE ENDPOINTS IN HUMAN IMMUNODEFICIENCY VIRUS CLINICAL TRIALS

D. Y. LIN

Department of Biostatistics, SC-32, University of Washington, Seattle, WA 98195, U.S.A.

M. A. FISCHL

Department of Medicine, R-60A, University of Miami School of Medicine, P.O. Box 016960, Miami, FL 33101, U.S.A.

AND

D. A. SCHOENFELD

Biostatistics Center, Massachusetts General Hospital, Fruit Street, Boston, MA 02114, U.S.A.

SUMMARY

In human immunodeficiency virus clinical trials, the CD4-lymphocyte count has been regarded as a promising surrogate endpoint for clinical efficacy measures such as time to opportunistic infection and survival time. In the present paper, we test this hypothesis according to a criterion proposed by Prentice. This criterion requires the surrogate variable to capture the entire effect of treatment on the clinical endpoint, and it is satisfied if the hazard rate of the clinical endpoint is not affected by treatment among patients with the same preceding history of the surrogate variable. We analyse data from two completed zidovudine trials using the Cox regression model with the CD4-lymphocyte count as a time-varying covariate. The results indicate that the CD4-lymphocyte count captures part of the relationship between zidovudine and time to a first critical event but does not fulfil the Prentice criterion.

1. INTRODUCTION

The conventional approach to evaluating the efficacy of therapeutic agents is to conduct clinical trials using clinical endpoints that represent the quality or the length of life as outcome measures. For human immunodeficiency virus type 1 (HIV) infection and acquired immunodeficiency syndrome (AIDS), such endpoints include time to opportunistic infection and survival time. Unfortunately, conventional clinical trials not only entail high expense, but may also require a long time for completion. Researchers and patients wish to assess the effectiveness of new drugs as quickly as possible. For HIV infection and AIDS, this desire has taken on added urgency and has led investigators to follow the path taken by researchers in cancer and cardiovascular diseases to explore laboratory variables that could serve as 'surrogate' endpoints in clinical trials. Substitution of a short-term or frequent biological marker for a rare or late clinical endpoint can lead to substantial reduction in sample size and trial duration.

The CD4-lymphocyte count has been regarded as one of the most promising surrogate endpoints for clinical efficacy measures such as time to opportunistic infection or survival time.

0277–6715/93/090835–08$09.00

Received November 1990
Revised July 1992

There is biological rationale and empirical evidence for this proposition. Infection with HIV is associated with a progressive increase in viral replication and a progressive depletion of CD4-lymphocytes.[1] Thus, a treatment that suppresses HIV replication and results in an increase in CD4-lymphocytes should improve clinical outcome. Various studies have demonstrated that the CD4-lymphocyte level has prognostic value for predicting the development of AIDS in HIV-infected people and for predicting mortality resulting from complications of HIV infection.[2-4] Such evidence, however, is insufficient to establish the CD4-lymphocyte count as a surrogate endpoint in clinical trials. A drug that increases CD4-lymphocytes may have little influence on opportunistic infections or survival if it does not suppress HIV replication. On the other hand, a drug may improve survival without causing a substantial increase in the number of CD4-lymphocytes, such as in the case with *Pneumocystis carinii* pneumonia prophylaxis.

A series of papers (Ellenberg and Hamilton;[5] Wittes *et al.*;[6] Hillis and Seigel;[7] Prentice[8]) on surrogate endpoints appeared in the April 1989 issue of *Statistics in Medicine*. In particular, Prentice[8] presented a formal methodology for evaluation and analysis of surrogate endpoints. By the Prentice definition, a surrogate endpoint for a clinical endpoint is a variable that yields a valid test of the null hypothesis of no association between the treatment and the clinical endpoint. This criterion essentially requires the surrogate variable to 'capture' any relationship between the treatment and the clinical endpoint. We can operationalize this notion by requiring that the hazard rate of the clinical endpoint at any follow-up time is independent of treatment conditional on the preceding history of the surrogate variable.

The objective of this paper is to evaluate the role of the CD4-lymphocyte count as a surrogate endpoint in HIV clinical trials according to the Prentice criterion. We analyse data from two completed zidovudine studies (one on patients with AIDS or advanced AIDS-related complex and one on patients with mildly symptomatic HIV infection). We describe these two studies in the next section. In Section 3, we present the statistical methods for our evaluations. The main results of data analysis appear in Section 4. The final section presents further results and some discussions.

2. STUDIES

Zidovudine is a potent inhibitor of HIV replication *in vitro*.[9] The clinical benefit of this agent was first demonstrated in a placebo-controlled trial in adults with AIDS or advanced AIDS-related complex (ARC), in which mortality and the frequency and severity of opportunistic infections (OIs) were significantly decreased among the recipients of zidovudine.[10] The study was sponsored by the Burroughs Wellcome Company and we hereafter refer to it as the BW 02 study.

The BW 02 study enrolled 281 patients, among whom 160 had AIDS and 121 had advanced ARC. There were 144 patients assigned to the zidovudine group and 137 to the placebo group. When the trial ended, 27 patients had completed 24 weeks of the study, 152 had completed 16 weeks, and the remainder had completed at least 8 weeks. Sixty-one patients (21 in the zidovudine group and 40 in the placebo group) had been withdrawn from study medication. The main reasons for early withdrawals were the occurrence of OIs, toxicities and the patient's request. Patients withdrawn from study medication were followed for survival.

Nineteen placebo recipients and 1 zidovudine recipient died during the study. OIs developed in 51 patients receiving placebo, as compared with 25 receiving zidovudine. All deaths occurred after the first OIs or study medication withdrawal. Both mortality and OIs constituted primary endpoints in this study. In our formal analysis, we will determine the extent to which the level of CD4-lymphocytes is a surrogate endpoint for time to the first OI. We treat the individual

withdrawal time as the censoring time with respect to the first OI. There are two main reasons for using OI rather than survival. First, CD4-lymphocytes were rarely evaluated after the occurrence of first OIs; therefore, the covariate information is insufficient for an adequate assessment of the effect of CD4-lymphocytes on survival time. Second, there were many more first OI events than deaths. Most notably, there was only one death among patients treated with zidovudine, which precludes any reliable estimate for the hazard of death among these patients.

Most patients had measurements of CD4-lymphocytes about every four weeks. The median numbers of baseline (that is, pretreatment) CD4-lymphocytes were 50 and 160 cells/mm^3 for the AIDS and advanced ARC patients, respectively. Patients who received zidovudine had increases in CD4-lymphocytes for the first 12 weeks of therapy, whereas a decline was most frequently observed in those who received placebo during the study. The increases in the number of CD4-lymphocytes in patients receiving zidovudine were significant for both the AIDS and ARC patients. After 12 weeks of therapy, however, a gradual decline in the number of CD4-lymphocytes was noted among patients with AIDS who received zidovudine. In contrast, the increases in CD4-lymphocytes among patients with ARC who received zidovudine persisted throughout most of the study.

In addition to demonstrating the benefit of zidovudine, the BW 02 study suggested that patients with less advanced HIV disease may experience greater clinical benefit and less toxicity than those with more advanced disease.[11] These considerations prompted the AIDS Clinical Trials Group (ACTG) Protocol 016 study, which was a placebo-controlled trial on the safety and efficacy of zidovudine in the treatment of patients with mildly symptomatic HIV disease.[12]

Three hundred and fifty-one patients were assigned to placebo and 360 to zidovudine in the ACTG 016 study. The median duration of actual follow-up for all patients was 11 months. Patients withdrawn from study medication for any reason were followed for the development of AIDS or advanced ARC and survival. Upon trial termination, 51 patients (36 in the placebo and 15 in the zidovudine groups) had developed AIDS, advanced ARC or death as a first critical event.

Advanced ARC was defined as the presence of two or more specified symptoms and a CD4-lymphocyte count of less than 200 cells/mm^3 on two consecutive occasions at least 14 days apart.[12] To avoid partial circularity of 'true' and 'surrogate' endpoints in our analysis, we exclude advanced ARC from our definition of a critical event. With this modification, 22 first critical events occurred in the placebo group, as compared with 6 in the zidovudine group.

CD4-lymphocytes were measured about every 4 weeks for the first 16 weeks. After the 16th week, measurements were repeated at weeks 24, 40, 52 and 64. The median numbers of baseline CD4-lymphocytes were about 400 cells/mm^3 for both the zidovudine and placebo groups. Significant differences between the treatment groups in the CD4-lymphocyte count occurred in patients with less than 500 pretreatment CD4 cells/mm^3 after 4 weeks of therapy, which persisted through week 52. Less prominent changes occurred in patients with 500 or more pretreatment CD4 cells/mm^3.

3. METHODS

Let t denote the time from enrolment in a clinical trial and let T denote the true endpoint, which is a first critical event in our setting. The treatment indicator x takes the value 1 if a patient is on zidovudine and takes the value -1 otherwise. Let $S(t) = \{z(u); 0 \leqslant u < t\}$ denote the history prior to time t of the surrogate (CD4-lymphocyte count) process $z(\cdot)$, and let $\lambda_T(t; \mathcal{A})$ denote the hazard rate of the true endpoint T at time t conditional on $\mathcal{A}$.

We can express the notion that a surrogate for T should capture the dependence of T on treatment x as

$$\lambda_T\{t; S(t), x\} = \lambda_T\{t; S(t)\}. \tag{1}$$

Criterion (1) requires that the hazard rate for T is independent of treatment conditional on the surrogate process. For departure from the null hypothesis that the surrogate response is independent of treatment to imply departure from the null hypothesis that the true endpoint is independent of treatment, it is necessary for the surrogate variable to have some prognostic value for the true endpoint; that is,

$$\lambda_T\{t; S(t)\} \neq \lambda_T(t). \tag{2}$$

Following Prentice,[8] we take (1) and (2) as our operational criteria for surrogate endpoint definition.

The Cox regression model[13] with time-varying covariates provides a convenient way to verify conditions (1) and (2). One advantage of this model is that it allows patients to contribute to 'risk sets' only when valid CD4-lymphocyte measurements are available. This is important in our setting because the history of CD4-lymphocyte counts was rather incomplete.

We define the 'proximate' value for $z(t)$ as the nearest prior CD4-lymphocyte measurement, provided that this prior measurement occurred no more than, say, q days before t. We say that a CD4-lymphocyte count at time t, $z(t)$, is 'censored' if there are no proximate values available; otherwise, we say $z(t)$ has a 'valid' measurement. We will use the baseline CD4-lymphocyte count $z(0)$ and the current CD4-lymphocyte count $z(t)$ to represent the history of CD4-lymphocyte measurements up to time t. We say that an event is 'valid' if the patient who experiences a first critical event at time t has valid measurements on both $z(0)$ and $z(t)$. The foregoing definitions are similar to those of Gail.[14]

The risk set for a valid event occurring at time t consists of patients under observation at time t who are event-free and who have valid measurements on both $z(0)$ and $z(t)$. Note that a patient may fail to be at risk at time t_1 and later re-enters the risk set at time $t_2 > t_1$ if additional CD4-lymphocyte data become available. The counting process approach[15] to the Cox regression indicates that the usual asymptotic results hold in this setting provided that the timing of CD4-lymphocyte measurements is unrelated to the patient's status.

The width of the time window $[t - q, t]$ is a trade-off between the closeness of the proximate values to the true CD4-lymphocyte counts and the size of risk sets. If q is too large, then the proximate CD4-lymphocyte counts may differ substantially from the true values. Of course, the accuracy of the approximation depends on the rate of the CD4-lymphocyte change. On the other hand, if q is too small, then there are few valid measurements and the variances of the regression estimates are large. We wish to choose a small value of q for which proximate CD4-lymphocyte values are available on most patients in each risk set.

For the BW 02 study, we let q be 5 weeks since most patients have CD4-lymphocyte measurements less than 5 weeks apart. This width is too small for the ACTG 016 study since after the 16th week CD4-lymphocytes were only evaluated at weeks 24, 40, 52 and 64. Schedules of CD4-lymphocyte measurements and rates of CD4-lymphocyte changes in the ACTG 016 study suggested that 18 weeks is a better choice. Even though $q = 18$ weeks for the entire study, most of the proximate values for the first 20 weeks of the study were prior measurements of less than 5 weeks since CD4-lymphocyte measurements occurred about every 4 weeks during the first 16 weeks.

Table I. Cox regression analysis of time to the first OI for the BW 02 Study

Covariate	Estimate	Model 1	Model 2
Treatment	Coefficient	− 0·459	− 0·316
	Standard error	0·123	0·132
	Coeff./S.E.	− 3·75	− 2·39
Status	Coefficient	− 0·408	− 0·266
	Standard error	0·128	0·141
	Coeff./S.E.	− 3·20	− 1·89
Baseline CD4 count	Coefficient	—	0·121
	Standard error	—	0·184
	Coeff./S.E.	—	0·66
Current CD4 count	Coefficient	—	− 0·576
	Standard error	—	0·173
	Coeff./S.E.	—	− 3·33

4. RESULTS

We investigated the role of the CD4-lymphocyte count as a surrogate endpoint in the two zidovudine trials described in Section 2 using the methods developed in Section 3. We represented the history of CD4-lymphocytes by two covariates, baseline CD4 count and current CD4 count, both of which were expressed on the natural logarithmic scale. We report here the main results.

For the BW 02 study, we first fit a model with two covariates, treatment and status, where status took the value 1 if the patient had advanced ARC at study entry and the value − 1 if the patient had AIDS. Both covariates were highly significant. To check criteria (1) and (2), we fit a second model with treatment, status, baseline CD4 count and current CD4 count. By criterion (1), treatment should become non-significant in this model. The results in Table I indicate that the inclusion of CD4-lymphocyte variables reduces the (two-sided) p-value for treatment from 0·0002 (model 1) to 0·017 (model 2). Thus, criterion (1) is not met at the 5 per cent level. The extremely high prognostic value of current CD4 count shown in model 2 implies that criterion (2) is satisfied.

We considered the question of whether the change in CD4 count or the current CD4 count is a better predictor for time to progression. Of course, when the baseline CD4 is included, the model with the change in CD4 is equivalent to that with the current CD4. The results from model 2 indicate that baseline CD4 count is not significant if we include current CD4 count. When we deleted baseline CD4 count from the model, the results for the remaining covariates changed very little. In comparison, when we replaced current CD4 count in model 2 by CD4 change, the latter being the difference of current CD4 count from baseline CD4 count, both baseline CD4 count and CD4 change were highly significant. Thus, the more parsimonious model would include the current value rather than the change. The positive coefficient for the baseline value in model 2, if it is not due to chance, might indicate that patients who dropped faster to a low current value had worse prognosis.

The results for the ACTG 016 study appear in Table II. The inclusion of CD4-lymphocyte variables only slightly reduces the significance of treatment (compare models 1 and 2). The covariate current CD4 count is again highly predictive of a first critical event (see model 2).

Table II. Cox regression analysis of time to the first critical event for the ACTG 016 Study

Covariate	Estimate	Model 1	Model 2
Treatment	Coefficient	− 0·665	− 0·752
	Standard error	0·230	0·277
	Coeff./S.E.	− 2·89	− 2·71
Baseline CD4 count	Coefficient	—	0·973
	Standard error	—	0·637
	Coeff./S.E.	—	1·53
Current CD4 count	Coefficient	—	− 1·622
	Standard error	—	0·374
	Coeff./S.E.	—	− 4·34

5. DISCUSSIONS AND FURTHER RESULTS

We conclude from this investigation that the (proximate) CD4-lymphocyte count captures part of the relationship between zidovudine and time to a first critical event but does not fulfil the stringent definition of a surrogate endpoint. Specifically, a patient treated with zidovudine might have a considerably better prognosis than an untreated patient who had the same proximate value of the CD4-lymphocyte count.

The proximate value of the CD4-lymphocyte count is a variable that one can use to determine whether or not the patient is responding to a therapy because it is the measurement on CD4-lymphocytes that patients and physicians actually observe. The proximate value of CD4-lymphocytes used in our analysis, however, may be quite different from the real CD4-lymphocyte process that accounts for the effect of zidovudine on the clinical endpoint. To see why this might be the case, assume that the development of a clinical event takes several weeks. Therefore, a patient's hazard at any given day should be a function of his CD4-lymphocyte counts averaged over the past few weeks. We denote this average by $z^*(t)$. Alternatively, we can think of $z^*(t)$ as a 'real' CD4-lymphocyte count about which the observed CD4-lymphocyte count $z(t)$ varies. The function $z^*(t)$ was not observed in the studies we analysed. First, there were measurement errors and short-term variations in a single CD4-lymphocyte measurement. Using data from the ACTG 016 study, we compared the variation in two measurements taken less than a week apart with the variation in measurements taken 12 weeks apart. We found that about 60 per cent of the variations was accounted for by short-term variations. Thus, the observed CD4-lymphocyte count at time t could differ appreciably from $z^*(t)$. Secondly, the proximate value of the CD4-lymphocyte count on a given day might be a measurement taken several weeks earlier and therefore could differ considerably from the CD4-lymphocyte count on that day. Using methods proposed by Prentice,[16] one can show that, even if the relative risk at time t is independent of treatment conditional on $z^*(t)$, the relative risk given the proximate CD4-lymphocyte count will involve a term multiplying the treatment indicator.

Upon one referee's suggestion, we carried out a second analysis with time-varying indicator functions of CD4-lymphocytes. The indicator function approach might be biologically more meaningful if the virus is susceptible to the drug being tested in a fraction of patients, but will be less meaningful if responses to therapy form a continuum. In our analysis, the CD4 covariate took the value 1 if the proximate CD4-lymphocyte count was above a given cutoff and the value 0

o herwise. We chose 50 and 200 cells/mm^3 as the cutoffs for the BW 02 and ACTG 016 studies, r spectively, although we also tried various neighbouring cutoffs. We found that the indicator v riables were poorer surrogates than the actual values of the proximate CD4-lymphocyte counts ir that they explained even less of the treatment effect.

Because death was the other primary endpoint in the BW 02 study, we also investigated the role o' the CD4-lymphocyte count as a surrogate endpoint for survival time. The standardized p rameter estimates for treatment were $-3{\cdot}00$ and $-2{\cdot}72$ under models 1 and 2, respectively. The effect of the CD4-lymphocyte count on survival time turned out to be much weaker than on C I time. Incidentally, A. A. Tsiatis and his Harvard colleagues (personal communications) p rformed a similar analysis with imputed values of CD4-lymphocytes. They carried out their i putation by modelling the marker process with a mixed-effects model. They concluded that the C D4-lymphocyte count did not absorb the treatment effect on survival at all. We note that the s atistical results on survival time are not reliable since there was only one death in the placebo g oup and since most deaths occurred considerably after stopping therapy.

In our analysis, we assumed that the timing of CD4-lymphocyte measurements was unrelated t the patient's status. This assumption might not be satisfied if the measurements were sought b cause of clinical determination. Unfortunately, there was insufficient information to verify this a sumption, and it is difficult to avoid such an assumption in statistical analysis.

Both the BW 02 and the ACTG 016 studies were terminated prematurely due to early evidence f r the zidovudine success. The early stoppings affect our analysis in two ways. First, the z dovudine effect might have been exaggerated. This may account for some of the zidovudine e fect that was in excess of the effect caused by increases in CD4-lymphocytes. Secondly, there was li mited information on the long-term effect of zidovudine. Thus our conclusions apply only to s ort-term zidovudine administrations.

The CD4-lymphocyte count could be a better surrogate endpoint for agents which have a more p tent and sustained inhibitory activity against HIV *in vivo*. The methodology developed in S ction 3 has sufficient generality to assess surrogate roles of any agents. One can use a surrogate e dpoint identified for a particular agent in later studies only if the agents tested have similar m chanisms of action. Therefore, it is essential to understand therapeutic mechanisms at a funda- m ental pathogenesis level before one could recommend any surrogate endpoints for practical use.

Many statisticians feel that the Prentice criterion is too stringent. One cannot realistically e pect a surrogate to explain completely the relationship between the treatment and the true e dpoint. It is therefore desirable to quantify how much information concerning the effect of t eatment the surrogate captures. Some simple measures include changes in the (standardized) p rameter estimate or in the p-value for the treatment indicator when 'surrogate' variables are a dded in the model. In fact, these measures were the basis for our conclusion that only a small f action of the zidovudine effect on clinical outcome was explained by the CD4-lymphocyte c unt.

The CD4-lymphocyte count can still be used and is being used as endpoints in phase I and II s udies, as its changes reflect drug activities. A drug may be approved for use if it increases C D4-lymphocytes and inhibits HIV replication, though its long-term effects on clinical endpoints n ed to be carefully monitored after the approval. There are several possible ways to use the C D4-lymphocyte count as an endpoint. The simplest measure is the CD4 level at a specified time p int. An alternative approach is to represent the CD4 profile with a summary measure such as t e area under the curve. One may also use time to a drop of the CD4-lymphocyte count below s me arbitrary level (for example, 100 cells/mm^3) or time to a 50 per cent reduction from the b seline CD4-lymphocyte count. Owing to measurement errors and short-term variations, it is a visory to confirm such an event with a second measurement.

ACKNOWLEDGEMENTS

The authors are deeply grateful to the Burroughs Wellcome Company for providing the BW 02 data and to the AIDS Clinical Trials Group for providing the ACTG 016 study data. They would also like to thank co-editor T. Colton and two referees for helpful comments and suggestions. This research was supported by the National Institutes of Health grants R01-AI-24643 and R01-AI-29168 (for Lin), U01-AI-27675 (for Fischl) and N01-AI-95030 (for Schoenfeld).

REFERENCES

1. Redfield, R. R. and Burke, D. S. 'HIV infection: the clinical picture', *Scientific American*, **259**, 90–98 (1988).
2. Polk, B. F., Fox, R., Brookmeyer, R., Kanchanaraksa, S., Kaslow, R., Visscher, B., Rinaldo, C. and Phair, J. 'Predictors of the acquired immunodeficiency syndrome developing in a cohort of seropositive homosexual men', *New England Journal of Medicine*, **316**, 62–66 (1987).
3. Goedert, J. J., Biggar, R. J., Melbye, M., Mann, D. L., Wilson, S., Gail, M. H., Grossman, R. J., DiGivia, R. A., Sanchez, W. C., Weiss, S. H. and Blattner, W. A. 'Effect of T4 count and cofactors on the incidence of AIDS in homosexual men infected with human immunodeficiency virus', *Journal of the American Medical Association*, **257**, 331–334 (1987).
4. Safai, B., Johnson, K. G., Myskowski, P. L., Koziner, B., Yang, S. Y., Cunningham-Rundles, S., Godbold, J. H. and Dupont, B. 'The natural history of Kaposi's sarcoma in the acquired immunodeficiency syndrome', *Annals of Internal Medicine*, **103**, 744–750 (1985).
5. Ellenberg, S. S. and Hamilton, J. M. 'Surrogate endpoints in clinical trials: cancer', *Statistics in Medicine*, **8**, 405–413 (1989).
6. Wittes, J., Lakatos, E. and Probstfield, J. 'Surrogate endpoints in clinical trials: cardiovascular diseases', *Statistics in Medicine*, **8**, 415–425 (1989).
7. Hillis, A. and Seigel, D. 'Surrogate endpoints in clinical trials: ophthalmologic disorders', *Statistics in Medicine*, **8**, 427–430 (1989).
8. Prentice, R. L. 'Surrogate endpoints in clinical trials: definition and operational criteria', *Statistics in Medicine*, **8**, 431–440 (1989).
9. Mitsuya, H., Weinhold, K. J., Furman, P. A., Clair, M. H., Lehrman, S. N., Gallo, R. C., Bologneri, D., Barry, D. W. and Broder, S. '3'-Azido-3'-deoxythymidine (BW A509U): an antiviral agent that inhibits the infectivity and cytopathic effect of human T-lymphotropic virus type III/lymphadenopathy-associated virus *in vitro*', *Proceedings of National Academy of Science, U.S.A.*, **82**, 7096–7100 (1985).
10. Fischl, M. A., Richman, D. D., Grieco, M. H., Gottlieb, M. S., Volberding, P. A., Laskin, O. L., Leedom, J. M., Groopman, J. E., Mildvan, D., Schooley, R. T., Jackson, G. G., Durack, D. T., King, D. and the AZT Collaborative Working Group. 'The efficacy of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex', *New England Journal of Medicine*, **317**, 185–191 (1987).
11. Richman, D. D., Fischl, M. A., Grieco, M. H., Gottlieb, M. S., Volberding, P. A., Laskin, O. L., Leedom, J. M., Groopman, J. E., Mildvan, D., Hirsch, M. S., Jackson, G. G., Durack, D. T., Lehrman, S. N. and the AZT Collaborative Working Group. 'The toxicity of azidothymidine (AZT) in the treatment of patients with AIDS and AIDS-related complex', *New England Journal of Medicine*, **317**, 192–197 (1987).
12. Fischl, M. A., Richman, D. D., Hansen, N., Collier, A. C., Carey, J. T., Para, M. F., Hardy, W. D., Dolin, R., Powderly, W. G., Allan, J. D., Wong, B., Merigan, T. C., McAuliffe, V. J., Hyslop, N. E., Rhame, F. S., Balfour, H. H., Spector, S. A., Volberding, P., Pettinelli, C., Anderson, J. and the AIDS Clinical Trials Group. 'The safety and efficacy of zidovudine (AZT) in the treatment of subjects with mildly symptomatic human immunodeficiency virus type 1 (HIV) infection', *Annals of Internal Medicine*, **112**, 727–737 (1990).
13. Cox, D. R. 'Regression models and life tables' (with discussion), *Journal of the Royal Statistical Society, Series B*, **34**, 187–220 (1989).
14. Gail, M. H. 'Evaluating serial cancer marker studies in patients at risk of recurrent disease', *Biometrics*, **37**, 67–78 (1981).
15. Andersen, P. K. and Gill, R. D. 'Cox's regression model for counting processes: a large sample study', *Annals of Statistics*, **10**, 1100–1120 (1982).
16. Prentice, R. L. 'Covariate measurement errors and parameter estimation in a failure regression model', *Biometrika*, **69**, 331–342 (1982).