STATISTICS IN MEDICINE
Statist. Med. 2002; **21**:2449–2463 (DOI: 10.1002/sim.1219)

Estimating a treatment effect in survival studies in which patients switch treatment

Michael Branson[1,*,†] and John Whitehead[2]

[1]*Novartis Pharma AG, Lichtstrasse 35, CH-4056, Basel, Switzerland*
[2]*Medical and Pharmaceutical Statistics Research Unit, The University of Reading, P.O. Box 240, Earley Gate, Reading, RG6 6FN, U.K.*

SUMMARY

For disease indications such as Acquired Immune Deficiency Syndrome (AIDS) and various cancers, randomization to a pure control treatment may be scientifically desirable but not ethically acceptable. Clinicians may insist that the experimental treatment be made available, at least as a rescue medication, for all patients in the control arm. A method for estimating a treatment effect in survival data from randomized clinical trials of this type is developed under an accelerated failure time model. This approach retains all patients in the groups to which they were randomized and is not based on an *ad hoc* subgroup analysis. By conditioning on having observed patient switch times, this method avoids the need to model patient switching patterns in the analysis. This new approach is evaluated using simulation studies, and is illustrated through analysing data from a Medical Research Council lung cancer trial. Copyright © 2002 John Wiley & Sons, Ltd.

KEY WORDS: non-compliance; survival data; estimation

1. INTRODUCTION

This paper focuses on the statistical analysis of randomized clinical trials in which patients allocated to the control arm are permitted eventually to receive experimental treatment. For clinical research in disease indications such as Acquired Immune Deficiency Syndrome (AIDS) and various cancers, the active controlled prospective randomized study may be scientifically desirable, but considered unethical because some patients receive control therapy indefinitely.

Delayed administration of the experimental treatment to patients in the control arm may be necessary if clinicians insist that it be made available as a rescue medication. On a more pragmatic note, without the option of experimental treatment for all patients in the study, patient recruitment may be difficult.

*Correspondence to: Michael Branson, Novartis Pharma AG, Lichtstrasse 35, CH-4056, Basel, Switzerland
†E-mail: mps@reading.ac.uk

Contract/grant sponsor: MRC; contract/grant number: G78/5359

Received June 2000
Accepted November 2001

Unbiased statistical inference regarding the comparison of treatment policies is obtained by an intention-to-treat analysis comparing the groups as randomized. Despite allowing a switch from control to experimental treatment, many such studies have the objective of comparing the experimental arm with the control arm, *as if no patients in the control arm had ever switched over to experimental treatment*. The advantage of experimental over control treatment under complete adherence to the original randomized treatment is referred to by Sommer and Zeger [1] as *biological efficacy*, and its statistical evaluation is complicated by patients switching treatment. In the original paper by Sommer and Zeger, the term biological efficacy referred to the treatment effect among patients who complied with treatment. Throughout this paper, the meaning of this term is widened to encompass the treatment effect as if all patients adhered to their original randomized treatment.

Robins and co-workers [2–5] have developed analytical techniques for survival outcomes which relate a patient's observed event time to an event time that would have been observed if no treatment had been administered (T_L), assuming treatment has a multiplicative effect (e^{η}) on a patient's lifetime. In the approach suggested by Robins and Tsiatis [2], e^{η} is estimated under a 'test-based' procedure. This test-based estimation process consists of applying a candidate value for η to the observed survival times and obtaining a value for a rank-test statistic (denoted by $Z(\eta)$), such as the logrank statistic, comparing the groups as randomized. The point estimate for η is given by $\{\hat{\eta} : Z(\hat{\eta}) = 0\}$. Essentially, the point estimate for the multiplicative effect is the value that balances T_L across treatment groups. No systematic estimation algorithm exists for providing $\hat{\eta}$ when employing the Robins and Tsiatis [2] model. In general, a grid search over possible values is used in order to determine the appropriate point estimate (and corresponding confidence interval).

The Robins and Tsiatis method has been applied to data from the AIDS Clinical Trials Group study [2, 3], ACTG 002, comparing high dose to low dose Zidovudine, and to the MRFIT trial [4, 5], in which the effect of smoking cessation on time to death or first myocardial infarction was investigated. White and co-workers [6–8] applied this method to the Concorde trial [6, 7] comparing the use of Zidovudine in patients with asymptomatic HIV infection and to data from a hypertensive trial in the elderly [8].

In this paper a new method for the analysis of survival data from such studies is presented, which builds on the ideas of Robins and Tsiatis [2], to develop an alternative estimator. The proposed method replaces their rank-test approach with a likelihood-based analysis, and provides a systematic and rapid parameter estimation scheme, rather than an exhaustive grid search.

White *et al.* [7] describe Robins' approach as producing a 'randomization-based effect estimator' (RBEE) that is, the treatment estimate is based on the groups as randomized, thus avoiding many of the potential pitfalls and biases introduced with subgroup analyses. In this paper, the parametric approach adopted also leads to an RBEE that allows for patients switching treatment and it may be generalized to incorporate other baseline covariates. The treatment effect is estimated as if patients in the control arm never switched to experimental treatment to produce a biological efficacy estimator. A distributional assumption is imposed on survival times, but the relationship between switching treatment and patient prognosis is not modelled. The properties of this iterative parameter estimation (IPE) method are evaluated using extensive simulation studies, and it is illustrated through the analysis of data from a British Medical Research Council lung cancer study [9].

In Section 2 a clinical trial employing an immediate versus delayed treatment strategy is described in order to motivate what follows. In Section 3, the modelling framework used throughout this paper is introduced. Parameter estimation using IPE is discussed in Section 4 and evaluated using simulation studies in Section 5. Extension of this modelling approach to allow for baseline covariates is given in Section 6. Details regarding censored failure time data and how to implement the estimation algorithm in the presence of censoring is presented in Section 7. The analysis of the data from the MRC trial is presented in Section 8 and concluding remarks are given in Section 9.

2. A TRIAL OF RADIOTHERAPY IN NON-SMALL CELL LUNG CANCER

Kalk *et al.* [9] describe a Medical Research Council randomized clinical trial in non-small cell lung cancer (NSCLC). This trial was designed to investigate and compare immediate versus delayed thoracic radiotherapy in patients with inoperable NSCLC, as there exists some doubt as to the clinical benefit of immediate radiotherapy for those patients with NSCLC (Carroll *et al.* [10], Crook *et al.* [11]). Articles by Phillips and Miller [12] and Cox *et al.* [13] advocate immediate thoracic radiotherapy, suggesting that one may achieve a prolonged survival and may also improve a patient's quality of life. Brashear [14] and Cohen [15] advocate the 'delayed' or 'wait and see' treatment regimen, maintaining that no evidence exists that suggests immediate radiotherapy provides prolonged survival or increased quality of life, thus recommending that radiotherapy be implemented as palliative symptom control. The MRC prospective randomized lung cancer trial was needed to provide evidence from a randomized comparison between immediate and delayed radiotherapy for patients with inoperable NSCLC. Therefore, this study was designed to evaluate the 'immediate' versus 'delayed' treatment policies. As a secondary analysis, it is of interest to evaluate thoracic radiotherapy itself, as if compared to standard treatment. Patients with no or minimal symptoms, in whom there was no compelling indication for immediate radiotherapy and who were not suitable for radical radiotherapy with curative intent, were eligible for this trial.

3. THE MODELLING FRAMEWORK

Suppose that a study is conducted to compare two treatment groups, experimental and control, where the study objective is to compare the experimental arm with the control arm, as if no patients in the control arm had ever switched over to experimental treatment. Assume that the experimental treatment acts multiplicatively on a patient's survival time, that is an accelerated failure time model applies, and denote the magnitude of this multiplicative effect by e^{η}. Define the treatment indicator for the ith patient by X_i, taking the value 0 for control and 1 for the experimental treatment. Suppose that n_E and n_C patients are randomized to the experimental and control groups, respectively, ($n_E + n_C = n$). Also denote the observed survival and switch times for the ith patient by $T_{R,i}$ and U_i, respectively. For patients randomized to the experimental arm, the switch time (U_i) equals zero. Define a latent survival time ($T_{L,i}$) for the ith patient as the survival time that would have been observed if no experimental treatment had been administered.

In order to set the scene, we first consider the simple scenario in which no patients in the control group switch to the experimental treatment. In this case the survival time $T_{R,i}$ of the ith patient satisfies

$$P(T_{R,i} > t) = \begin{cases} S(t) & \text{for } X_i = 0, \\ S(\mathrm{e}^{-\eta} t) & \text{for } X_i = 1, \quad i = 1, \ldots, n \end{cases} \tag{1}$$

The usual interpretation of the effect e^{η} from an accelerated failure time model is that values of $\mathrm{e}^{\eta} > 1$ reflect a beneficial experimental treatment effect, with the expected lifetime equal to e^{η} times that of the control group, and values of $\mathrm{e}^{\eta} < 1$ indicate a detrimental effect of experimental treatment. Equality of expected time to event for the experimental and control groups holds when $\mathrm{e}^{\eta} = 1$.

In this scenario, the latent event time is observed for all uncensored patients in the control arm, and is an abstract quantity for all other patients. The observed and latent event times are related according to,

$$T_{L,i} \stackrel{d}{=} \mathrm{e}^{-\eta X_i} T_{R,i} \tag{2}$$

where $\stackrel{d}{=}$ denotes equality in distribution. If equality in distribution in Equation (2) is replaced by true equality, then this becomes the 'strong version of the accelerated failure time model' of Cox and Oakes [16], and defines a class of survival models frequently referred to as 'structural' or 'causal' models [2–8, 17]. For the purposes of this paper, equality in distribution is sufficient.

To generalize model (2) to the case where patients in the control group may switch treatment, the multiplicative effect e^{η} is regarded as the rate at which a patient's lifetime is 'slowed down' when receiving experimental treatment. In this case, latent event times are observed for only those uncensored patients who never receive experimental treatment, and are an abstract quantity for other patients. Conditional on a patient's switching time (U_i), if the true estimate for the multiplicative effect of experimental treatment e^{η} were known, Equation (2) could be generalized to patients who switch by,

$$T_{L,i} \stackrel{d}{=} U_i + \mathrm{e}^{-\eta}(T_{R,i} - U_i) \tag{3}$$

This is a simplified case of the time-dependent accelerated failure time model originally proposed by Cox and Oakes [16]. Equation (3) demonstrates that if the true value of e^{η} were known, the event times in the delayed group for patients who switch could be back-transformed to the event time that would have been observed if experimental treatment had not been administered: the latent event times.

In order to implement the IPE algorithm, a parametric survival time distribution for the event times is imposed, for example, this might be the Weibull or the exponential distribution. It may be appropriate to use the event times from the experimental therapy arm in identifying an appropriate parametric distribution.

4. PARAMETER ESTIMATION WHEN PATIENTS SWITCH TREATMENT

The IPE algorithm retains all patients in the treatment group to which they were originally allocated (experimental or control) and it does not require any modelling of the distribution of a patient's switch time. Thus, by the definition introduced in Section 1, following reference [7], IPE is an RBEE. The IPE algorithm is an iterative procedure that converges to the point estimate for e^{η} that satisfies Equation (2), while using Equation (3) to back-transform event times for patients who switch treatment in the control group.

The IPE algorithm is implemented as follows: An initial estimate of $e^{-\eta}$ on the right hand side of Equation (2) is obtained by comparing the groups as randomized using a parametric accelerated failure time model. For a given initial point estimate of $e^{-\eta}$, the survival times of patients who switch treatment in the control arm are transformed using the right hand side of Equation (3). Using these transformed times and the original observed survival times for all other patients, the two groups are once again compared using the parametric survival analysis. This analysis leads to another estimate for $e^{-\eta}$. This estimate is used in a second transformation of the survival times of patients who switch treatment in the control group. Once the entire process has been repeated several times, the value of $e^{-\eta}$ used in the transformation will be close (say within 10^{-5}) to the value used in the previous iteration. At that point, the procedure is considered to have converged (see Figure 1 for details).

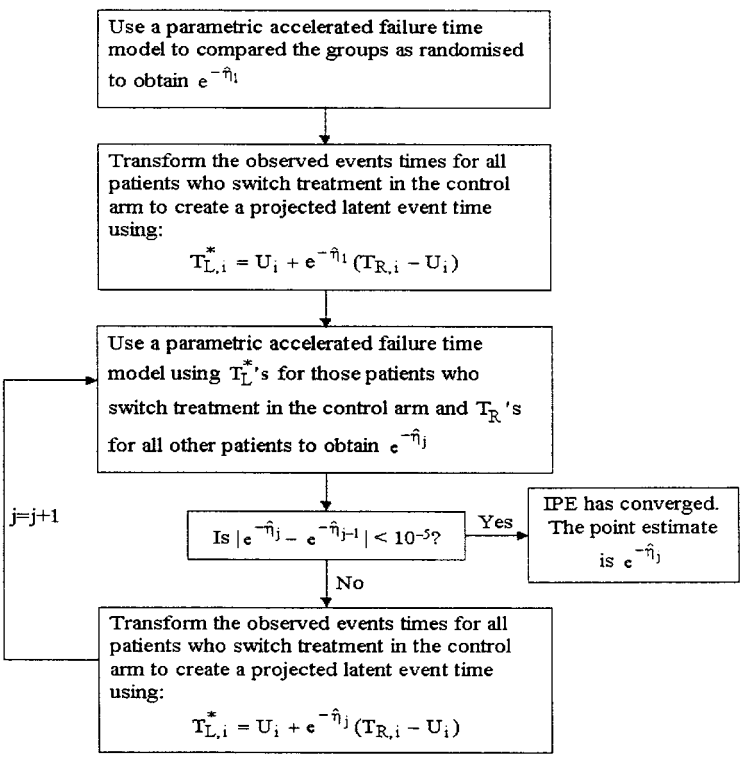

Figure 1. Flow chart of IPE algorithm.

If no patients in the control arm switch treatment, IPE reduces to the usual intention-to-treat parametric survival analysis, as the algorithm will stop after two iterations. If some patients switch treatment, IPE provides the estimate of *biological efficacy* of experimental over control treatment. In the final iteration of IPE, the estimated 'covariance matrix' from the parametric survival analysis is *uncorrected* for the extent to which patients switch treatment in the control arm. These 'estimated variances' are too small and thus cannot be used to provide confidence limits. Since patient switching patterns are not modelled when using the IPE algorithm, estimated variances are obtained using bootstrap techniques [18].

5. PARAMETER ESTIMATION: SIMULATION RESULTS

For the sake of simplicity, consider the case where no survival times are censored. The application to censored survival times is discussed in Section 7.

Latent survival times ($T_{L,i}$) were generated from a Weibull distribution, with survivor function $S(t)=\exp(-e^{-\mu}t^{\gamma})$. The expected time to failure was equal to 500 days, and the data were positively skewed. The value of the shape parameter γ was set to 1.5, leading to the scale parameter $e^{-\mu}$ defined such that $\mu=6.3169$. Extensive simulation studies were undertaken with a variety of switching patterns for patients in the control arm. The extent to which patients switch to experimental treatment was varied in two ways: (i) increasing the proportion of patients who switch treatment and (ii) increasing the proportion of a patient's lifetime spent on experimental treatment, given that they were a 'switcher'. All simulation results were investigated for the following multiplicative effects (e^{η}): 5, 2, 1, 0.5, 0.2, with 200 patients per treatment arm.

In generating the survival data for these simulations, first the underlying latent event times ($T_{L,i}$) for all patients were generated from a Weibull distribution. Having generated these latent event times, we generated the event times that would have been observed had no patients in the control arm switched treatment (these are computed using Equation (2)). We refer to these generated data as 'pure' event times and are clearly unobservable when patients in the control arm are permitted to switch treatment. However, throughout these simulations, analysing these data provide a measure with which the 'IPE' summary statistics may be compared. The summary statistics for these pure times are labelled 'Pure E v C' in the summary tables.

For each patient in the control group, a Bernoulli(p) outcome determined whether the patient was a 'switcher', with p set to 0.3 and 0.7, reflecting small and large expected proportions of patients switching in the control arm, respectively. If a patient was a 'switcher', the switch time (U_i) was calculated as the latent event time ($T_{L,i}$) multiplied by an observation from a beta(a,b) distribution, where parameters a and b were chosen to reflect the expected proportion ($=a/(a+b)$) of the latent event time spent on the control treatment. The observed survival time ($T_{R,i}$) was then computed using the relationship between a patient's latent survival time and observed survival time as defined in Equation (3). For all patients in the experimental group, the switch time was set to zero, and the observed survival time was computed using Equation (3). The Appendix lists the details regarding generating these survival data.

The parameter estimates for all simulation studies were obtained from maximizing a standard likelihood function with two treatment groups, using a Newton–Raphson algorithm with

Table I. Comparison of 'IPE' versus 'Pure E v C' based on 1000 simulations with 200 patients per treatment group.

Beta(a, b)	Multiplicative effect (e^{η})	Estimation method	Mean	Variance	Bias	MSE
Probability of a patient in the control group switching = 0.3						
(4, 2)	2	IPE	1.99901	0.02122	−0.00099	0.02122
		Pure E v C	1.99932	0.01711	−0.00068	0.01711
	1	IPE	0.99950	0.00531	−0.00050	0.00531
		Pure E v C	0.99966	0.00428	−0.00034	0.00428
	0.5	IPE	0.49975	0.00133	−0.00025	0.00133
		Pure E v C	0.49983	0.00107	−0.00017	0.00107
(2, 4)	2	IPE	2.01165	0.03014	0.01165	0.03028
		Pure E v C	2.00990	0.01916	0.00990	0.01926
	1	IPE	1.00583	0.00754	0.00583	0.00757
		Pure E v C	1.00495	0.00479	0.00495	0.00481
	0.5	IPE	0.50291	0.00188	0.00291	0.00189
		Pure E v C	0.50248	0.00120	0.00248	0.00120
Probability of a patient in the control group switching = 0.7						
(4, 2)	2	IPE	2.01504	0.03073	0.01504	0.03095
		Pure E v C	2.01169	0.01797	0.01169	0.01811
	1	IPE	1.00752	0.00768	0.00752	0.00774
		Pure E v C	1.00584	0.00449	0.00584	0.00453
	0.5	IPE	0.50376	0.00192	0.00376	0.00193
		Pure E v C	0.50292	0.00112	0.00292	0.00113
(2, 4)	2	IPE	2.00758	0.06265	0.00758	0.06270
		Pure E v C	2.00542	0.01770	0.00542	0.01773
	1	IPE	1.00380	0.01566	0.00380	0.01568
		Pure E v C	1.00271	0.00442	0.00271	0.00443
	0.5	IPE	0.50190	0.00392	0.00190	0.00392
		Pure E v C	0.50136	0.00111	0.00135	0.00111

back-tracking [19]. The vector of first derivatives and the Hessian matrix used in IPE were obtained using Maple V. The C-code used to evaluate the score functions and the Hessian matrix was obtained using the C-code generator (specifying that the C-code is optimized) in Maple. The first author has written SAS macros for IPE and bootstrap estimation in the SAS programming language which are available on request. The properties of the IPE estimation algorithm were evaluated using 1000 simulated clinical studies per combination of switching patterns.

Table I details the results from the simulation studies using Weibull survival times, presenting the long-run estimate of the multiplicative effect (e^{η}) of experimental treatment, the long-run estimated variance, bias and MSE. Table I details the results for $e^{\eta} = 2, 1$ and 0.5. These results were typical of all the simulation studies conducted, hence, for brevity, this reduced set of results are presented.

These results demonstrate that IPE consistently estimates the multiplicative effect of treatment. This was also the case for the nuisance parameters in the likelihood function, although these results are not given. As the extent of switching treatment was increased, either by increasing the probability of switching treatment from 0.3 to 0.7, or by increasing the

expected proportion of lifetime spent on experimental treatment from beta$(4,2)$ to beta$(2,4)$, the long-run estimate of variance increased for IPE. For all simulation studies investigated, IPE provided accurate estimates of the multiplicative effect of experimental treatment.

If patients in the control group are permitted to switch treatment, one would expect the uncertainty in estimating the effect of treatment to be greater then for the situation in which no patients switch treatment. By comparing the estimated variance of e^{η} for 'IPE' to 'Pure E v C' in Table I, it can be seen that this is indeed the case.

The estimates of variance given in Table I are long-run estimates from the 1000 simulated clinical studies, per permutation of patient switching patterns. In Section 4, the use of the bootstrap, for estimates of variance, was advocated when using IPE. In order to confirm that the bootstrap provides good estimates of variance, a simulated data set which had parameter estimates close to those specified (that is, $\hat{\mu} \approx 6.3169$, $\hat{\gamma} \approx 1.5$, $e^{\hat{\eta}} \approx 5, 2, 1, 0.5$ or 0.2 for Weibull survival times) was generated, for each permutation of patient switching patterns. Bootstrapping was applied to these simulated data sets, with each data set analysed using IPE, in order to obtain empirical estimates of variance. For all simulation studies investigated, the bootstrap estimates of variance (based on 2000 resampled data sets) were in very good agreement with those given in Table I.

6. ADJUSTING FOR BASELINE COVARIATES

To illustrate that IPE, applied with a suitable distributional assumption on survival times, is easily generalized to adjust for baseline covariates, further simulations were conducted. Simulated survival times were generated with one continuous variable (age), and one categorical variable (disease status) taking three levels (None {1}, Moderate {2}, Severe {3}), with each patient being equally likely to be in one of these three categories. Each patient's age (years) was generated from a uniform (40,70) distribution. The latent survival time for the ith patient ($T_{L,i}$) was generated using the following equation:

$$T_{L,i} = T_{0,i} \exp(\beta_1 \text{STAGE}_2 + \beta_2 \text{STAGE}_3 + \beta_3(\text{AGE} - 40)) \tag{4}$$

where $T_{0,i}$, is generated from the Weibull with the same values for the parameters as given in Section 5 ($i = 1, \ldots, n$), and STAGE_j defines indicator variables taking the value 1 if disease stage $= j$, and 0 otherwise.

The effect of disease stage was defined as: no effect for disease stage $= 1$; a 10 per cent reduction in survival time for disease stage $= 2$; and a 40 per cent reduction in survival time for those patients with disease stage $= 3$. The regression effect of age was defined such that at age 70, patients' survival times were half those at age 40. Therefore $\exp(\beta_1) = 0.9$, $\exp(\beta_2) = 0.6$ and $\exp(\beta_3) = 0.977[= \exp\{\ln(0.5)/30\}]$.

Introducing baseline prognostic factors into the IPE algorithm is achieved by entering the effects as one would normally do in an intention-to-treat parametric accelerated failure time model. That is, the prognostic factors (continuous regressor effects and/or indicator variables) are entered into the linear predictor component of the survival model. The IPE algorithm

Table II. Covariate adjustment $P(\text{switch} = 0.3)$: a comparison of 'IPE' versus 'Pure E v C' based on 1000 simulations with 200 patients per treatment group.

Beta(a,b)	Estimation method	Effect	Mean	Variance	Bias	MSE
Probability of a patient in the control group switching $= 0.3$; *multiplicative effect* $(e^{\eta}) = 2$						
(4, 2)	IPE	TRT	2.00704	0.02395	0.00704	0.02400
		$STAGE_2$	0.90309	0.00591	0.00309	0.00592
		$STAGE_3$	0.60177	0.00237	0.00177	0.00238
		AGE	0.97716	0.00001	0.000003	0.00001
	Pure E v C	TRT	2.00649	0.01936	0.00649	0.01940
		$STAGE_2$	0.90310	0.00590	0.00310	0.00591
		$STAGE_3$	0.60181	0.00238	0.00181	0.00238
		AGE	0.97716	0.00001	0.00002	0.00001
(2, 4)	IPE	TRT	1.99853	0.03085	−0.00147	0.03085
		$STAGE_2$	0.90330	0.00552	0.00330	0.00553
		$STAGE_3$	0.60243	0.00242	0.00243	0.00243
		AGE	0.97714	0.00001	−0.00002	0.00001
	Pure E v C	TRT	1.99911	0.01970	−0.00089	0.01970
		$STAGE_2$	0.90338	0.00550	0.00337	0.00552
		$STAGE_3$	0.60246	0.00242	0.00246	0.00242
		AGE	0.97713	0.00001	−0.00003	0.00001

then continues as specified in Section 4, but instead of having only treatment in the linear predictor, this now contains treatment and all prognostic covariates.

All the simulation studies reported in Section 4 were repeated using the latent survival times generated according to Equation (4). Table II and Table III detail only a few of the results from these simulations, using Weibull survival times, as these are indicative of the results obtained for all simulation studies used to investigate adjustment for baseline prognostic factors. The results in Tables II and III demonstrate that IPE provides accurate estimates for the effect of baseline prognostic factors and accurate estimates for the multiplicative effect of experimental treatment. As in Section 5, the bootstrap was employed to obtain estimates of variance, and these agreed very closely with the long-run estimates of variance presented in Tables II and III.

7. ANALYSIS OF CENSORED SURVIVAL DATA

The results presented in the previous sections are based on the analysis of uncensored survival data. IPE has been established as a method for estimating the effect of factors (treatment and baseline factors) in an accelerated failure time model. The extension of IPE to censored survival time data is straightforward and intuitive. Consider a clinical study where patients are followed up for a set length of time. Patients may be censored for two main reasons: they survive beyond the end of the study, or they are lost to follow-up prior to reaching the end of their follow-up time.

The IPE algorithm works by creating a projected latent event time for patients who switch treatment in the control group. Therefore, if there is an end of follow-up time for each patient,

Table III. Covariate adjustment $P(\text{switch} = 0.7)$: a comparison of 'IPE' versus 'Pure E v C' based on 1000 simulations with 200 patients per treatment group.

Beta(a,b)	Estimation method	Effect	Mean	Variance	Bias	MSE
Probability of a patient in the control group switching $= 0.7$; *multiplicative effect* $(e^{\eta}) = 2$						
$(4,2)$	IPE	TRT	2.01313	0.03068	0.01313	0.03085
		$STAGE_2$	0.89968	0.00574	−0.00032	0.00574
		$STAGE_3$	0.60145	0.00236	0.00145	0.00236
		AGE	0.97699	0.00002	−0.00017	0.00002
	Pure E v C	TRT	2.01045	0.01801	0.01045	0.01812
		$STAGE_2$	0.89967	0.00574	−0.00033	0.00574
		$STAGE_3$	0.60141	0.00235	0.00141	0.00236
		AGE	0.97699	0.00002	−0.00017	0.00002
$(2,4)$	IPE	TRT	2.00096	0.05829	0.00096	0.05829
		$STAGE_2$	0.89985	0.00537	−0.00016	0.00537
		$STAGE_3$	0.59843	0.00256	0.00157	0.00256
		AGE	0.97714	0.00001	−0.00002	0.00001
	Pure E v C	TRT	2.00140	0.01654	0.00140	0.01654
		$STAGE_2$	0.89977	0.00536	−0.00023	0.00536
		$STAGE_3$	0.59845	0.00255	0.00155	0.00255
		AGE	0.97714	0.00001	−0.00002	0.00001

IPE should be constrained by this 'upper' limit on observable survival times. If the parameter estimate obtained while implementing IPE projects a patient's latent event time beyond the end of study time, this event would naturally have been censored at this time and should be censored on the projected time scale. The method due to Robins and Tsiatis [2] also requires the implementation of a similar 'recensoring' algorithm, but a principal difference between IPE and their approach is that IPE requires recensoring survival times only if they are projected beyond the end of the study. Consequently, the occurrence of recensoring is limited to those patients who switch treatment in the control arm, and will be required only if experimental treatment is detrimental compared with control. This is because a detrimental treatment effect suggests that patients would survive longer on control leading to the *potential* for IPE to project a patient's latent event time beyond the end of study time, for those patients in the control arm who switched treatment.

White *et al.* [7] report extensive work, undertaken to demonstrate that the recensoring algorithm of Robins and Tsiatis [2] is necessary. They demonstrated that a (small) reduction in bias was achieved when employing the potentially intensive recensoring algorithm when the switch time and latent event time were highly correlated. In the simulation studies used to evaluate IPE, given that a patient switched treatment, the switch time was computed as a beta(a,b) multiple of the latent event time, T_L. It can be shown for the scenarios considered in Sections 4 and 7 that the correlation between the latent event time and the switch time, conditional on switching occurring, is,

$$\text{Corr}(U, T_L) = \mathbf{E}(B)\left\{\mathbf{E}(B^2) + \mathbf{E}(T_L)^2 \frac{\text{Var}(B)}{\text{Var}(T_L)}\right\}^{-1/2} \tag{5}$$

Table IV. Summary of recensoring algorithm for 1000 simulations with 200 patients per treatment group (P(censored by end of study) = 0.1, P(random censored) = 0.0).

Beta(a, b)	Multiplicative effect (e^{η})	Recensored	Mean	Variance	Bias	MSE
Probability of a patient in the control group switching = 0.3						
(4, 2)	2	Yes	2.00006	0.02288	0.00006	0.02288
		No	1.99972	0.02336	−0.00028	0.02336
	1	Yes	1.00019	0.00587	0.00019	0.00587
		No	1.00039	0.00592	0.00039	0.00592
	0.5	Yes	0.49991	0.00138	−0.00010	0.00138
		No	0.50158	0.00139	0.00158	0.00139
(2, 4)	2	Yes	2.01308	0.02948	0.01308	0.02966
		No	2.01255	0.03191	0.01255	0.03207
	1	Yes	1.00706	0.00763	0.00706	0.00768
		No	1.00703	0.00825	0.00703	0.00830
	0.5	Yes	0.50338	0.00184	0.00338	0.00185
		No	0.50346	0.00202	0.00346	0.00203
Probability of a patient in the control group switching = 0.7						
(4, 2)	2	Yes	2.00914	0.03034	0.00914	0.03042
		No	2.00782	0.03397	0.00782	0.03403
	1	Yes	1.00733	0.00743	0.00733	0.00748
		No	1.00798	0.00804	0.00798	0.00810
	0.5	Yes	0.50560	0.00176	0.00560	0.00179
		No	0.50920	0.00192	0.00919	0.00200
(2, 4)	2	Yes	2.00552	0.04084	0.00552	0.04087
		No	1.99978	0.06581	−0.00022	0.06581
	1	Yes	1.00618	0.01035	0.00618	0.01039
		No	1.00583	0.01647	0.00583	0.01650
	0.5	Yes	0.50552	0.00242	0.00552	0.00245
		No	0.50447	0.00408	0.00447	0.00410

where the term B denotes the beta(a, b) random variable. For the simulation studies presented in this paper, the correlation between the switch time and the latent event is high. When switch time is determined by a beta(2, 4) multiple of the latent event time, the conditional correlation is equal to 0.72, and when it is a beta(4, 2) multiple, the conditional correlation is 0.90.

Simulation studies were used to investigate the effect of the inclusion of censored data on estimates computed using the IPE algorithm – convergence of the IPE algorithm was observed in all simulation studies. The end of study marker was chosen such that patients allocated to the treatment offering comparative longevity were censored due to the end of the study with probability 0.1. Table IV details a few of the results (that is, for $e^{\eta} = 2, 1, 0.5$) from the simulation studies used to investigate the recensoring algorithm and are typical of all results obtained from these simulation studies. Also, a set of simulation studies were conducted in which patients were not only censored at the end of the study, but were also randomly censored during follow-up with probability 0.1, creating a total censoring of 20 per cent of survival times. These results are given in Table V.

These results demonstrate that IPE remains accurate in the presence of censoring and that recensoring projected survival times reduces the variance of the parameter estimate relative to

Table V. Summary of recensoring algorithm for 1000 simulations with 200 patients per treatment group (P(censored by end of study) = 0.1, P(random censored) = 0.1).

Beta(a, b)	Multiplicative effect (e^{η})	Recensored	Mean	Variance	Bias	MSE
Probability of a patient in the control group switching = 0.3						
(4, 2)	2	Yes	2.00530	0.02742	0.00530	0.02745
		No	2.00406	0.02800	0.00406	0.02801
	1	Yes	1.00295	0.00716	0.00295	0.00717
		No	1.00321	0.00721	0.00321	0.00722
	0.5	Yes	0.50212	0.00166	0.00212	0.00166
		No	0.50383	0.00165	0.00383	0.00166
(2, 4)	2	Yes	2.01057	0.03100	0.01057	0.03111
		No	2.00642	0.03357	0.00642	0.03361
	1	Yes	1.00445	0.00795	0.00445	0.00797
		No	1.00444	0.00856	0.00444	0.00856
	0.5	Yes	0.50381	0.00184	0.00330	0.00186
		No	0.50352	0.00203	0.00352	0.00204
Probability of a patient in the control group switching = 0.7						
(4, 2)	2	Yes	2.00045	0.03253	0.00045	0.03253
		No	1.99290	0.03619	−0.00710	0.03624
	1	Yes	1.00370	0.00813	0.00370	0.00814
		No	1.00426	0.00872	0.00426	0.00874
	0.5	Yes	0.50554	0.00196	0.00554	0.00199
		No	0.50932	0.00212	0.00932	0.00221
(2, 4)	2	Yes	2.02236	0.04484	0.02236	0.04534
		No	1.99891	0.07123	−0.00109	0.07123
	1	Yes	1.00828	0.01171	0.00828	0.01178
		No	1.00916	0.01822	0.00916	0.01830
	0.5	Yes	0.51116	0.00276	0.01116	0.00288
		No	0.51026	0.00452	0.01026	0.00463

not recensoring. Recensoring does not generally inflate the bias and leads to smaller MSEs. Also, as the extent to which patients switch treatment increases (that is, as the potential for recensoring patients increases), the greater is the reduction in estimated variance (and MSE). This is in accord with intuition, since survival times have to be less than or equal to the end of study time by definition. If a patient's projected survival time goes beyond the end of study time, and recensoring is not employed, this survival time will be large compared with observed survival times. In essence, this (outlying) projected survival time adds noise to the estimation process, thereby increasing variance.

8. DATA ANALYSIS: MRC NSCLC TRIAL

To demonstrate the application of IPE for censored survival data, the analysis of data from a British Medical Research Council (MRC) lung cancer trial (LU17) is now presented. All analyses presented in this section are unadjusted for baseline covariates. In this data set, 206 patients were randomized (104 immediate treatment, 102 delayed treatment). A total of 153

patients experienced the event (75 immediate, 78 delayed). The survival times for patients in the immediate treatment arm were used to establish an appropriate parametric distribution for the survival times. Inspecting the plot of log(–log(S(time))) versus log(time) confirmed that, for these data, an appropriate parametric survival time distribution was Weibull.

An intention-to-treat analysis of these data using a Weibull accelerated failure time model provided a point estimate of 1.111 and a 95 per cent confidence interval of (0.841,1.380). That is, comparing the treatment policies, immediate thoracic radiotherapy is seen to provide some benefit to patients relative to delayed thoracic radiotherapy, but this benefit was not statistically significant at the 5 per cent two-sided level ($p = 0.3956$).

The analysis of these data using the IPE algorithm provided a point estimate of 1.158 and a 95 per cent normal approximation confidence interval based on 2000 bootstrap resampled data sets of (0.753,1.563). The corresponding bootstrap 95 percentile interval was (0.786, 1.611), which is in good agreement with the normal approximation interval. An analysis of these data using the approach of Robins and Tsiatis [2], based on the logrank statistic, provided a point estimate 1.14 with approximate 95 per cent confidence interval (0.69, 1.7). The recensoring algorithm for the latter analysis required 36 patients to be recensored (17.5 per cent) whereas, for the IPE algorithm, no recensoring was required in the primary analysis of these data, and the most extreme case of the 2000 bootstrap samples required recensoring 5 patients (2.4 per cent) and this situation occurred in only three of the samples. When implementing Robins' approach, it is advocated that the resulting point estimate and confidence interval is accompanied by the intention-to-treat p-value. We also recommend that the IPE point estimate and confidence interval is accompanied by the intention-to-treat p-value.

9. CONCLUDING REMARKS

This paper has demonstrated that the estimation of a treatment effect and baseline prognostic factors under an accelerated failure time modelling approach may be achieved by using the IPE algorithm. The results show that, as the extent to which patients switch treatment increases, the variance of the estimate of treatment effect also increases. Irrespective of the extent to which patients switch treatment in the control arm, the IPE algorithm provided accurate point estimates for the treatment effect. It was also shown that by employing a recensoring algorithm as defined in Section 7, the variance (and MSE) of the estimate of treatment effect is reduced. Unlike the method due to Robins and Tsiatis [2], recensoring is limited to patients in the control arm who switch treatment.

It seems intuitive that an analysis of survival data with minimal recensoring should be preferred. The IPE algorithm and recensoring algorithm offers an approach to estimating a treatment effect (and baseline factors), when patients in the control arm are permitted to switch treatment, and limits the extent to which patients are recensored.

Throughout this paper, the iterative parameter estimation (IPE) algorithm has been motivated by the need to adjust a treatment estimate for patients switching treatment in the control arm and has been implemented under the assumption of a parametric distribution for survival times. Conceptually, the IPE, having imposed a distributional assumption on survival times, may be visualized as using the 'best' estimate for the multiplicative effect of experimental treatment (e^{η}) in projecting patient survival times, for those patients in the control arm who switch treatment, to times that would have been observed if they had never been administered

experimental treatment. Thus, it is simple to extend the use of IPE to adjust a treatment estimate in the case where patients in the control arm switch over to experimental treatment *and* patients in the experimental arm switch to control treatment. This is again achieved by using the 'best' estimate for e^{η} in projecting the survival times of those patients who switch in the control arm to the times that would have been observed if they had never been administered experimental treatment, and projecting the survival times of those patients who 'switch' in the experimental arm to the times that would have been observed if they had never stopped taking experimental treatment. Relaxation of the parametric assumption has been investigated and appears to be possible; this will be the subject of a future publication.

The general applicability of IPE can also be thought of in the context of patient non-compliance, that is when patients do not comply with the randomized treatment to which they were assigned. Observed patient non-compliance is certainly commonplace in randomized clinical studies, and IPE offers the data analyst a methodology that respects the randomization and adjusts the treatment estimate for the observed patient non-compliance. Computationally, the IPE algorithm is typically more intensive than a standard intention-to-treat analysis. A degree of pragmatism is required when deciding if the parameter estimate(s) should be adjusted for patient non-compliance. For example, if the extent of non-compliance is less than 10 per cent of all patients in the study, it is unlikely that this would introduce serious bias if the intention-to-treat analysis were used to estimate the biological efficacy of the experimental treatment.

APPENDIX: SIMULATING SURVIVAL DATA IN WHICH PATIENTS SWITCH TREATMENT

Survival data were generated according to the following steps:

1. Generate a patient's latent event time ($T_{L,i}$) as a random number from a Weibull (μ, γ) distribution.
2. Compute the 'Pure' event times ($T_{P,i}$) as $T_{P,i} = e^{\eta X_i} T_{L,i}$, where X_i is an indicator random variable that equals one for those patients in the experimental arm and zero otherwise.
3. Generate a patient's 'switch' time (U_i):
   (i) Set the switch times for all patients in the experimental arm to zero.
   (ii) For a patient in the control arm, generate a Bernoulli(p) observation which takes the value one if the patient is a 'switcher' and zero otherwise. If a patient in the control arm is a 'switcher' define their switch time as $U_i = w \times T_{L,i}$, where w is a random number from a beta(a, b) distribution. Otherwise, set their switch time equal to their latent event time.
4. A patient's observed event time ($T_{R,i}$) is calculated as $T_{R,i} = U_i + e^{\eta}(T_{L,i} - U_i)$.

ACKNOWLEDGEMENTS

During the preparation of this paper, the first author was funded by MRC grant G78/5359. The authors are grateful to the MRC Cancer Trials Unit for provision of the data described in Section 2. We also thank the referees for their valuable contributions that have improved this paper.

REFERENCES

1. Sommer A, Zeger SL. On estimating efficacy from clinical trials. *Statistics in Medicine* 1991; **10**:45–52.
2. Robins J, Tsiatis A. Correcting for non-compliance in randomised trials using rank preserving structure failure time models. *Communications in Statistics – Theory and Methods* 1991; **20**:2609–2631.
3. Robins J, Greenland S. Adjusting for differential rates of prophylaxis therapy for PCP in high-dose versus low-dose AZT treatment arms in an AIDS randomised trial. *Journal of the American Statistical Association* 1994; **89**:737–749.
4. Mark S, Robins J. A method for the analysis of randomised trials with compliance information: an application to the multiple risk factor intervention trial. *Controlled Clinical Trials* 1993; **14**:79–97.
5. Mark S, Robins J. Estimating the causal effect of smoking cessation in the presence of confounding factors using a rank preserving structural failure time model. *Statistics in Medicine* 1993; **12**:1605–1628.
6. White IR, Walker S, Babiker AG, Darbyshire JH. Impact of treatment changes on the interpretation of the Concorde trial. *AIDS* 1997; **11**:999–1006.
7. White IR, Babiker AG, Walker S, Darbyshire JH. Randomization-based methods for correcting for treatment changes: examples from the Concorde trial. *Statistics in Medicine* 1999; **18**:2617–2634.
8. White IR, Goetghebeur EJT. Clinical trials comparing two treatment policies: which aspects of the treatment policies makes a difference? *Statistics in Medicine* 1998; **17**:319–339.
9. Kalk S, White R, Hopwood P, Girling D, Sambrook R, Harvey A, Qian W, Stephens R. S53 immediate versus delayed thoracic radiotherapy (trt) in patients with unresectable advanced non-small cell lung cancer (NSCLC) and minimal symptoms: results of an MRC/BTS randomised trial. *Thorax* 1999; **53** (Supplement 3):A14.
10. Carroll M, Morgan SA, Yarnold JR, Hill JM, Wright NM. Prospective evaluation of a watch policy in patients with inoperable non-small cell lung cancer. *European Journal of Cancer and Clinical Oncology* 1986; **22**: 1353–1356.
11. Crook A, Duffy A, Girling DJ, Souhami RL, Parmar M. Survey on the treatment of non-small cell lung cancer (NSCLC) in England and Wales. *European Respiratory Journal* 1997; **10**:l552–1558.
12. Phillips TL, Miller RJ. Should asymptomatic patients with inoperable bronchogenic carcinoma receive immediate radiotherapy? YES. *American Review of Respiratory Disease* 1978; **117**:405–410.
13. Cox JD, Komaki R, Byhardt RW. Is immediate chest radiotherapy obligatory for any or all patients with limited-stage non-small cell carcinoma of the lung? Yes. *Cancer Treatment Reports* 1983; **67**:327–331.
14. Brashear RE. Should asymptomatic patients with inoperable bronchogenic carcinoma receive immediate radiotherapy? NO. *American Review of Respiratory Disease* 1978; **117**:411–414.
15. Cohen MH. Is immediate radiation therapy indicated for patients with unresectable non-small cell lung cancer? No. *Cancer Treatment Reports* 1983; **67**:333–336.
16. Cox DR, Oakes D. *Analysis of Survival Data.* Monographs on Statistics and Applied Probability. Chapman and Hall: London, 1984.
17. Korhonen P, Laird, NM, Palmgren J. Correcting for non-compliance in randomised trials: an application to the ATBC study. *Statistics in Medicine* 1999; **18**:2879–2897.
18. Efron B, Tibshirini RJ. *An Introduction to the Bootstrap.* Monographs on Statistics and Applied Probability. Chapman and Hall: London, 1993.
19. Press W, Teukolsky S, Vetterling W, Flannery B. *Numerical Recipes in C—The Art of Scientific Computing*. 2nd edn. Cambridge University Press: Cambridge, 1996.

