Estimation of Survival Distributions of Treatment Policies in Two-Stage Randomization Designs in Clinical Trials

Jared K. Lunceford,1,* Marie Davidian,2 and Anastasios A. Tsiatis²

1Clinical Biostatistics, Merck Research Laboratories, P.O. Box 2000, Rahway, New Jersey 07065-0900, U.S.A.
2Department of Statistics, North Carolina State University, Raleigh, North Carolina 27695-8203, U.S.A.

*email: jared.lunceford@merck.com

SUMMARY. Some clinical trials follow a design where patients are randomized to a primary therapy at entry followed by another randomization to maintenance therapy contingent upon disease remission. Ideally, analysis would allow different treatment policies, i.e., combinations of primary and maintenance therapy if specified up-front, to be compared. Standard practice is to conduct separate analyses for the primary and follow-up treatments, which does not address this issue directly. We propose consistent estimators for the survival distribution and mean restricted survival time for each treatment policy in such two-stage studies and derive large-sample properties. The methods are demonstrated on a leukemia clinical trial data set and through simulation.

KEY WORDS: Induction therapy; Intent to treat; Inverse weighting; Maintenance therapy; Potential outcomes; Survival analysis.

1. Introduction

Cancer therapy frequently is implemented via a two-stage approach, where an initial treatment is given with the intent of inducing disease remission and a follow-up treatment is given to prolong the period before relapse and disease progression. In many cases, this second-stage maintenance therapy is given only to those patients who show a complete or partial remission in response to the induction therapy. Similarly, HIV patients may receive an initial therapy, which may be modified contingent on response to the initial treatment.

Cancer clinical trials studying combinations of induction and maintenance therapies are common (e.g., Tummarello et al., 1997). After enrollment, patients are randomized to induction therapy, followed by a subsequent randomization to maintenance therapy contingent on their remission status and consent. Protocol 8923 was a double-blind, placebo controlled two-stage trial conducted by the Cancer and Leukemia Group B (CALGB) and reported by Stone et al. (1995) examining the effects of infusions of granulocyte-macrophage colony-stimulating factor (GM-CSF) after initial chemotherapy in 388 elderly patients with acute myelogenous leukemia (AML). Standard chemotherapy for AML has a myelosuppressive effect, placing patients at increased risk of death due to infection or bleeding-related complications. As a hematopoietic cytokine, GM-CSF administered after chemotherapy might assist patient recovery by allowing more rapid reconstitution of bone marrow-derived lineages, thus reducing the number of deaths due to such complications. Patients were randomized initially to GM-CSF or placebo following standard chemotherapy. Later, patients meeting the complete remission criteria and consenting to further participation were randomized to one of two intensification treatments.

A main interest is to compare treatment policies, i.e., induction/maintenance therapy combinations if offered to patients up front, and to determine the combination leading to the greatest survival benefit. As is customary, the analysis would focus on comparing treatment policies under the intent-to-treat principle. However, like the randomization scheme used in these trials, data analysis typically is separated into two parts, neither of which addresses this issue directly: (i) estimating survival distribution under different induction therapies and consenting to further participation were randomized to one of two intensification treatments.

A main interest is to compare treatment policies, i.e., induction/maintenance therapy combinations if offered to patients up front, and to determine the combination leading to the greatest survival benefit. As is customary, the analysis would focus on comparing treatment policies under the intent-to-treat principle. However, like the randomization scheme used in these trials, data analysis typically is separated into two parts, neither of which addresses this issue directly: (i) estimating survival distribution under different induction therapies and consenting to further participation were randomized to one of two intensification treatments.

A main interest is to compare treatment policies, i.e., induction/maintenance therapy combinations if offered to patients up front, and to determine the combination leading to the greatest survival benefit. As is customary, the analysis would focus on comparing treatment policies under the intent-to-treat principle. However, like the randomization scheme used in these trials, data analysis typically is separated into two parts, neither of which addresses this issue directly: (i) estimating survival distribution under different induction therapies and consenting to further participation were randomized to one of two intensification treatments.

A main interest is to compare treatment policies, i.e., induction/maintenance therapy combinations if offered to patients up front, and to determine the combination leading to the greatest survival benefit. As is customary, the analysis would focus on comparing treatment policies under the intent-to-treat principle. However, like the randomization scheme used in these trials, data analysis typically is separated into two parts, neither of which addresses this issue directly: (i) estimating survival distribution under different induction therapies and consenting to further participation were randomized to one of two intensification treatments.

A main interest is to compare treatment policies, i.e., induction/maintenance therapy combinations if offered to patients up front, and to determine the combination leading to the greatest survival benefit. As is customary, the analysis would focus on comparing treatment policies under the intent-to-treat principle. However, like the randomization scheme used in these trials, data analysis typically is separated into two parts, neither of which addresses this issue directly: (i) estimating survival distribution under different induction therapies and consenting to further participation were randomized to one of two intensification treatments.

A main interest is to compare treatment policies, i.e., induction/maintenance therapy combinations if offered to patients up front, and to determine the combination leading to the greatest survival benefit. As is customary, the analysis would focus on comparing treatment policies under the intent-to-treat principle. However, like the randomization scheme used in these trials, data analysis typically is separated into two parts, neither of which addresses this issue directly: (i) estimating survival distribution under different induction therapies and consenting to further participation were randomized to one of two intensification treatments.
2. Model Framework and Proposed Estimators

Let treatment $A$, at levels $A_1$ and $A_2$, and treatment $B$, at levels $B_1$ and $B_2$, be the induction and postremission treatments, respectively. In the two-stage trial, patients are randomized initially to one of the $A$ treatment levels. If a patient achieves remission and consents to further participation, he is then randomized to a level of $B$. An objective of the study is to estimate and compare survival probabilities and other relevant quantities associated with the treatment policies “treat with $A_j$ followed by $B_k$ if remission and consent,” which we write as $A_j B_k$, $j, k = 1, 2$. Comparisons among these four policies could be carried out directly with a design that randomizes patients up front to four groups, each group corresponding to a different policy. Under the intent-to-treat paradigm, comparisons would be made among the four groups without regard to differential remission and consent rates. Thus, as with any intent-to-treat analysis, the overall policy survival distributions of interest reflect the combined influences of treatments, remission, and consent, and no attempt is made to separate these effects. With the four-group design, data for patients who do not remit and consent under $A_j B_1$ and $A_j B_2$ would be used only to estimate quantities for their assigned policies, even though such patients never receive maintenance therapy. The methods we now propose allow inference on the four policies based on data from the two-stage design while making more efficient use of the information from patients who do not remit and consent.

A convenient way to conceptualize this problem is through potential outcomes or counterfactuals (Holland, 1986). Assume that each subject $i$ has an associated set of random variables $(R_{1i}, R_{2i}, T_{11i}, T_{12i}, T_{21i}, T_{22i})$, where $R_{1i}$ is the remission/consent status that subject $i$ would achieve were he assigned to one of the policies $A_j B_k$, $k = 1, 2$, and similarly for $R_{2i}$. Implicit is the assumption that potential remission/consent status is a function only of the $A$ treatment level for a given policy and not on the subsequent $B$ treatment that the policy would dictate. The $T_{jki}$, $j, k = 1, 2$, represent the potential survival times $i$ would achieve if assigned to policy $A_j B_k$. We make the reasonable assumption that the potential survival times for subjects whose disease would not remit or who would not consent under $A_j$ would be the same whether the subject would receive $B_1$ or $B_2$ under the assigned policy. Formally, this implies that the distribution of $(R_j, T_{j1}, T_{j2})$ obeys the constraint

$$T_{j1} = T_{j2} \text{ if } R_j = 0, \quad j = 1, 2.$$  

(1)

The full set of variables $(R_1, R_2, T_{11}, T_{12}, T_{21}, T_{22})$ are not all observable for each individual. Rather, these variables represent what potentially might occur under policies possibly contrary to that to which the individual might actually be exposed in the trial.

This framework allows comparisons among policies to be characterized clearly. The distributions of the $T_{jk}$ represent survival for the population were all patients to be assigned to $A_j B_k$, so that inference on features of these distributions addresses directly the intent-to-treat question of interest. Thus, our approach is to estimate the survival distributions and means for the $T_{jk}$ from the observed data from a two-stage trial. It is critical to recognize that our use of counterfactuals does not imply a focus on causal treatment effects (Holland, 1986); rather, we use them as a device with which to address the intent-to-treat question.

If there were no censoring, then the data actually observed for subject $i$ are $(X_i, R_i, R_{zi}, T_i)$. Here, $X_i$ is the treatment assignment indicator, $X_i = j - 1$ if $i$ is randomized to $A_j$, so that $X_i = 0$ if $i$ is randomized to $A_2$; $R_i$ is observed remission/consent status, $R_i = 0$ if no remission or consent, $R_i = 1$ if remission and consent; $Z_i$ is the $B$ treatment assignment indicator, $Z_i = 0$ if randomized to $B_1$ ($B_2$), defined only if $R_i = 1$; and $T_i$ is observed survival time. We assume that observed and counterfactual remission/consent statuses are related by $R_i = R_{zi} (1 - X_i) + R_{zi} X_i$, i.e., observed status is equal to potential status under the $A$ treatment actually assigned. Observed survival time $T_i$ is assumed to be related to the counterfactual survival times as follows. For definiteness, consider $A_1$, i.e., $X_i = 0$. For $i$ such that $X_i = 0$, $R_i = R_{zi}$, and we assume $T_i$ satisfies

$$T_i = (1 - R_i)T_{11i} + R_i (1 - Z_i)T_{11i} + R_i Z_i T_{12i}.$$  

Under assumption (1), we also have that

$$T_i = (1 - R_i)T_{21i} + R_i (1 - Z_i)T_{11i} + R_i Z_i T_{12i}.$$  

The relationship for $A_2$ patients with $X_i = 1$ is analogous.

To account for right censoring, let $C_i$ be the time to censoring. We allow the censoring distribution to differ by $A$ treatment and let $K_i(t) = P(C_i < t \mid X_i = j - 1), j = 1, 2$. Because for most clinical trials total follow-up of patients is limited, of necessity we can only consider restricted lifetime, that is, survival up to some time $L$, where $L$ is smaller than the maximum follow-up time. Restricted lifetime is defined formally as $T_{jk}^L = \min(T_{jk}, L)$, where $K_j(L) > 0, j = 1, 2$. In the sequel, we regard $T_{jk}$ as restricted survival and write $T_{jk}^L$ as $T_{jk}$. Under these conditions, we observe i.i.d. vectors $(R_i, R_i Z_i, V_i, \Delta_i)$, where $\Delta_i = I(T_i < C_i); V_i = \min(T_i, C_i)$ is the observed death or censoring time; and, if $i$ is censored but has not yet remitted, we take $R_i = 0$. Assume that $C_i$ is conditionally independent of $(R_i, R_i Z_i, T_{11i}, T_{21i})$ given $X_i = j - 1, j = 1, 2$, and, conditional on $R_i = 1$ and $X_i = j - 1, Z_i$ is independent of $T_{11i}$ and $T_{21i}$, which is trivially satisfied by randomization of $B$ treatments. We allow the $B$ randomization probability $\pi_{Zj} = P(Z_i = 1 \mid R_i = 1, X_i = j - 1)$ to differ by $A$ treatment, although such differential randomization would be unlikely in practice.

Because patients randomized to $A_1$ and $A_2$ are two independent samples, for simplicity, we henceforth restrict attention to the $A_1$ data here and in the Appendices and suppress the subscript $j$ on $K, \pi_Z$, and other quantities defined below; the development would be identical for $A_2$. In the sequel, $i = 1, \ldots, n$ thus refers to the $n$ patients randomized to $A_1$, e.g., for $30-50 A$ randomization, the total number of patients is roughly $2n$.

We consider estimation of $S_{1k}(t) = 1 - P(T_{1k} \leq t) = 1 - F_{1k}(t)$, the survival function for policy $A_1 B_k$, $k = 1, 2$, from the observed data. Consider $A_1 B_1$. Ideally, if all subjects were assigned to $A_1 B_1$ and there were no censoring, then $V_i = T_i = T_{11i}$, and the natural estimator for $F_{11}(t)$ is $n^{-1} \sum_{i=1}^n I(V_i \leq t)$. With censoring and randomization to $B$ contingent on remission/consent status, only a subset of the $n$ patients have an observed (uncensored) survival time and have actual treatment consistent with the policy $A_1 B_1$. The
proposed estimators incorporate two forms of inverse weighting (e.g., Robins, Rotnitzky, and Zhao, 1994) to weight the observations from this subset in such a way that, probabilistically, the distribution of the weighted observations mimics that in the ideal case. Every uncensored value of \( T_i \) represents \( K^{-1}(T_i) \) individuals, so that, roughly speaking, the response for an uncensored individual counts for him/herself and \( K^{-1}(T_i) - 1 \) similar, censored individuals. Let \( Q_{1i} = 1 - R_i + (1 - \pi_Z)^{-1}R_i(1 - Z_i) \). Note that \( Q_{1i} = 0 \) if \( i \) received \( B \) treatment inconsistent with \( A_1B_1 \), i.e., \( R_i = 1 \) and \( Z_i = 1 \). Otherwise, \( Q_{1i} = 1 \) if \( R_i = 0 \), and \( Q_{1i} = (1 - \pi_Z)^{-1} \) if \( R_i = 1 \) and \( Z_i = 0 \). Thus, in these cases where \( i \)'s treatment is consistent with \( A_1B_1 \), \( Q_{1i} \) acts as a weight. Nonremitters consistent with \( A_1B_1 \) represent themselves and hence receive a weight of one, while if \( i \) achieves remission and consents, then \( i \) represents \( (1 - \pi_z)^{-1} \) remitting/consenting subjects who could have potentially been assigned to \( B_1 \). With \( Q_{2i} = 1 - R_i + \pi_Z R_i Z_i \), an analogous argument may be made for policy \( A_1B_2 \).

These considerations motivate the estimator

\[
\hat{F}_{1k}(t) = n^{-1} \sum_{i=1}^{n} \frac{\Delta_i Q_{ki}}{\hat{K}(V_i)} I(V_i \leq t) \tag{2}
\]

for \( k = 1, 2 \), where \( \hat{K}(t) = \Pi_{u \leq t} \left\{ 1 - d N^C(u)/Y(u) \right\} \) is the Kaplan–Meier estimate of the censoring survivor curve, with \( N^C(u) = \sum_{i=1}^{n} I(V_i \leq u, \Delta_i = 0) \) and \( Y(u) = \sum_{i=1}^{n} I(V_i \geq u) \). If the true \( K(t) \) is substituted in (2), \( \hat{F}_{1k}(t) \) is unbiased for \( F_{1k}(t) \), for \( k = 1 \),

\[
E \left\{ \frac{\Delta_i Q_{1i}}{\hat{K}(V_i)} I(V_i \leq t) \right\} = E \left\{ \frac{I(T_{11i} < C_i) Q_{1i}}{\hat{K}(T_{11i})} I(T_{11i} \leq t) \right\} = E \left\{ \frac{I(T_{11i} \leq t)}{\hat{K}(T_{11i})} \hat{K}(T_{11i}) Q_{1i} \right\} = E \left\{ I(T_{11i} \leq t) Q_{1i} \right\} = E \left\{ I(T_{11i} \leq t) \hat{K}(T_{11i}) Q_{1i} \right\} = E \left\{ I(T_{11i} \leq t) \hat{K}(V_i) \right\} = F_{11}(t),
\]

which follows by noting that \( E(Q_{1i} \mid R_i, T_{11i}) = 1 - R_i + (1 - \pi_Z)^{-1}E(R_i(1 - Z_i) \mid R_i, T_{11i}) = 1 \) from considering the cases \( R_i = 0 \) and \( R_i = 1 \) in turn. The argument for \( k = 2 \) is identical.

A second estimator may be obtained by averaging using a probabilistically adjusted sample size, i.e.,

\[
\hat{F}_{1k}(t) = \left( \sum_{i=1}^{n} \frac{\Delta_i Q_{ki}}{\hat{K}(V_i)} \right)^{-1} \sum_{i=1}^{n} \frac{\Delta_i Q_{ki}}{\hat{K}(V_i)} I(V_i \leq t). \tag{3}
\]

Assuming that there are no ties among the censored observations and that at least one uncensored restricted lifetime is equal to \( L \), which should be true with high probability if \( \hat{K}(L) > 0 \), \( \sum_{i=1}^{n} \Delta_i / \hat{K}(V_i) = n \), and (2) may be written alternatively as the solution to

\[
\sum_{i=1}^{n} \frac{\Delta_i}{\hat{K}(V_i)} [Q_{ki} I(V_i \leq t) - F_{1k}(t)] = 0. \tag{4}
\]

Similarly, (3) may be written as the solution to the estimating equation

\[
\sum_{i=1}^{n} \frac{\Delta_i}{\hat{K}(V_i)} [Q_{ki} I(V_i \leq t) - F_{1k}(t)] = 0. \tag{5}
\]

Combining (4) and (5), both estimators may be written as solutions to equations of the form

\[
\sum_{i=1}^{n} \frac{\Delta_i}{\hat{K}(V_i)} [Q_{ki} I(V_i \leq t) - F_{1k}(t) - \alpha_{1k} (Q_{ki} - 1)] = 0, \tag{6}
\]

where \( \alpha_{1k} = 0 \) yields the equation for \( \hat{F}_{1k}(t) \) and \( \alpha_{1k} = -\hat{\alpha}_{1k} \) gives that for \( \hat{F}_{1k}(t) \). This suggests a third estimator, \( \hat{F}_{1k}''(t) \), say, where \( \alpha_{1k} \) is chosen to have minimum variance among all estimators solving equations of the form (6). It is shown in Appendix A that

\[
\hat{F}_{1k}''(t) = n^{-1} \sum_{i=1}^{n} \frac{\Delta_i Q_{ki}}{\hat{K}(V_i)} I(V_i \leq t) - \hat{\alpha}_{1k} n^{-1} \sum_{i=1}^{n} \frac{\Delta_i}{\hat{K}(V_i)} (Q_{ki} - 1), \tag{7}
\]

where

\[
\hat{\alpha}_{1k} = \left[ n^{-1} \sum_{i=1}^{n} \Delta_i Q_{ki} (1 - Q_{ki})^{-1} I(V_i \geq u) \frac{\hat{K}(V_i)}{\hat{K}(V_i)} \right] + \int_{0}^{L} dN^C(u) \{ \hat{K}(u)Y(u) \}^{-1} \hat{E} \left\{ L_{1k}''(t, u) \right\} \frac{L_{1k}''(t, u)}{L_{1k}''(t, u)}
\]

\[
+ \left[ n^{-1} \sum_{i=1}^{n} (Q_{ki} - 1)^2 \right] + \int_{0}^{L} dN^C(u) \{ \hat{K}(u)Y(u) \}^{-1} \hat{E} \left\{ G_k''(u) \right\},
\]

with

\[
\hat{E} \{ L_{1k}''(t, u) \} = n^{-1} \sum_{i=1}^{n} \Delta_i \{ Q_{ki} I(V_i \leq t) - \hat{G}_{1k}(t, u) \} \times \{ Q_{ki} - 1 - \hat{G}_{Qk}(u) \} I(V_i \geq u) \frac{\hat{K}(V_i)}{\hat{K}(V_i)},
\]

\[
\hat{E} \{ G_k''(u) \} = n^{-1} \sum_{i=1}^{n} \Delta_i \{ Q_{ki} - 1 - \hat{G}_{Qk}(u) \}^2 I(V_i \geq u) \frac{\hat{K}(V_i)}{\hat{K}(V_i)},
\]

\[
\hat{G}_{Qk}(u) = \{ n\hat{S}(u) \}^{-1} \sum_{i=1}^{n} \Delta_i \{ Q_{ki} - 1 \} I(V_i \geq u) \frac{\hat{K}(V_i)}{\hat{K}(V_i)},
\]

\[
\hat{G}_{1k}(t, u) = \{ n\hat{S}(u) \}^{-1} \sum_{i=1}^{n} \Delta_i Q_{ki} I(V_i \leq t) I(V_i \geq u) \frac{\hat{K}(V_i)}{\hat{K}(V_i)},
\]

and \( \hat{S}(u) \) is the Kaplan–Meier estimator for \( P(T > u) \).
we denote estimators are given in Appendix A and yield variance estimators for (2), (3), and (7), leading to variance estimators for $S_1(t) = 1 - \hat{F}_1(t)$, $S'_1(t) = 1 - \hat{F}'_1(t)$, and $S''_1(t) = 1 - \hat{F}''_1(t)$ given by

$$
\widehat{\text{var}} \{ \hat{S}_1(t) \} = n^{-1} \left[ \sum_{i=1}^{n} \frac{\Delta_i Q_{ki} I(V_i \leq t) - \hat{F}_1(t)^2}{K(V_i)} \right. \\
\left. + \int_0^L \frac{dN^c(u)}{K(u)Y(u)} \hat{E} \{L_{ki}(t,u)\}^2 \right),
$$

(9)

$$
\widehat{\text{var}} \{ \hat{S}'_1(t) \} = n^{-1} \left[ \sum_{i=1}^{n} \frac{\Delta_i Q_{ki} I(V_i \leq t) - \hat{F}'_1(t)^2}{K(V_i)} \right. \\
\left. + \int_0^L \frac{dN^c(u)}{K(u)Y(u)} \hat{E} \{L'_{ki}(t,u)\}^2 \right),
$$

(10)

and

$$
\widehat{\text{var}} \{ \hat{S}''_1(t) \} = n^{-1} \left[ \sum_{i=1}^{n} \frac{\Delta_i I(V_i \leq t)}{K(V_i)} \right. \\
\times \{ Q_{ki} I(V_i \leq t) - \hat{F}''_1(t)^2 \} \\
\left. + \int_0^L \frac{dN^c(u)}{K(u)Y(u)} \hat{E} \{L''_{ki}(t,u)\}^2 \right],
$$

(11)

where

$$
\hat{E} \{L_{ki}(t,u)\}^2 = n^{-1} \sum_{i=1}^{n} \Delta_i \{ Q_{ki} I(V_i \leq t) - \hat{G}_{1i}(t,u) \}^2 \\
\times \{ I(V_i \geq u) - \hat{G}'_{1i}(t) \} \\
\hat{E} \{L'_{ki}(t,u)\}^2 = n^{-1} \sum_{i=1}^{n} \Delta_i \{ Q_{ki} I(V_i \leq t) - \hat{F}'_1(t) \} \\
- \hat{G}'_{1i}(t) \{ I(V_i \geq u) \} \\
\hat{G}'_{1i}(t,u) = \{ n\hat{S}(u) \}^{-1} \\
\times \sum_{i=1}^{n} \Delta_i Q_{ki} \{ I(V_i \leq t) - \hat{F}'_1(t) \} \{ I(V_i \geq u) \} \\
\hat{E} \{L''_{ki}(t,u)\}^2 = n^{-1} \sum_{i=1}^{n} \Delta_i \{ Q_{ki} I(V_i \leq t) - \hat{G}_{1i}(t,u) \}^2 \\
- \hat{G}_{1i}(t,u) \{ I(V_i \geq u) \} \\
\times \{ I(V_i \geq u) - \hat{G}_{1i}(t,u) \}.
$$

Variance estimators for $\hat{\mu}_{1i}$, $\hat{\mu}'_{1i}$, and $\hat{\mu}''_{1i}$ are obtained by replacing $I(V_i \leq t)$ by $V_i$ and the relevant survival distributions at specific time points. For example, to compare survival distributions at 1 year, with $j = 1, 2$ corresponding to GM-CSF and placebo and $k = 1, 2$ corresponding to intensification treatments I and II, the 3-d.f. Wald tests of $H_0$: $S_{11}(365) = S_{12}(365) = S_{21}(365) = S_{22}(365)$ based on $S_{12}(365)$, $S_{21}(365)$, and $\hat{S}_{21}(365)$ yield approximate chi-square test statistics of 1.24 ($p = 0.74$), 1.26 ($p = 0.74$), and 0.91 ($p = 0.82$), respectively, suggesting no evidence of a difference.

Table 1 shows estimates of mean restricted survival time for each treatment policy. Three-degree-of-freedom tests for $H_0$: $\mu_{11} = \mu_{12} = \mu_{21} = \mu_{22}$ based on $\hat{\mu}_{ijk}$, $\hat{\mu}'_{ijk}$, and $\hat{\mu}''_{ijk}$ yield test statistics and $p$-values of $2.67 (p = 0.44)$, 2.73 ($p = 0.44$), and 2.76 ($p = 0.43$), respectively, providing no evidence of overall mean restricted survival differences. All point estimates
suggest GM-CSF infusion treatment leads to a decrease in mean restricted survival time; however, two-tailed Wald tests of \( H_0: (\mu_{11} + \mu_{12})/2 = (\mu_{21} + \mu_{22})/2 \) using \( \hat{\mu}_{jk} \), \( \hat{\mu}'_{jk} \), and \( \hat{\mu}''_{jk} \) give z-scores of \(-1.57\) \((p = 0.12)\), \(-1.56\) \((p = 0.12)\), and \(-1.53\) \((p = 0.13)\), respectively, indicating no strong evidence for a GM-CSF effect. Tests of \( H_0: (\mu_{11} + \mu_{21})/2 = (\mu_{12} + \mu_{22})/2 \), contrasting the mean intensification treatment effect over induction therapy arms based on \( \hat{\mu}_{jk} \), \( \hat{\mu}'_{jk} \), and \( \hat{\mu}''_{jk} \), yield z-scores and p-values of \(-0.37\) \((p = 0.71)\), \(-0.29\) \((p = 0.77)\), and \(-0.10\) \((p = 0.92)\), indicating no evidence of a difference. Similarly, two-tailed tests for differences between intensification treatment policies within induction arms are not significant; for \( H_0: \mu_{11} = \mu_{12} \), the corresponding z-values are \(-0.36\) \((p = 0.72)\), \(-0.33\) \((p = 0.74)\), and \(-0.03\) \((p = 0.97)\); for \( H_0: \mu_{21} = \mu_{22} \), the z-values are \(-0.91\) \((p = 0.36)\), \(-0.93\) \((p = 0.35)\), and \(-0.60\) \((p = 0.55)\). The test of no interaction between A and B, i.e., \( H_0: \mu_{11} - \mu_{12} = \mu_{21} - \mu_{22} \), yielded test statistics \(0.44\) \((p = 0.66)\), \(0.51\) \((p = 0.61)\), and \(0.67\) \((p = 0.50)\).

### 4. Simulation Results

To evaluate performance of the methods, we carried out several simulations. Because data from \( A_1 \) and \( A_2 \) are independent, we need simulate only for \( A_1 \), say. Similar to CALGB 8923, all simulations were based on a 2.5-year study for \( n = 200 \) and \( 500 \) subjects. For each individual, censoring time \( C \) was generated as uniform \((0, 2.5)\) independent of all other variables. Remission/consent and \( B \) treatment indicators, \( R = R_1 \) and \( Z \), were sampled from Bernoulli \((\pi_R)\) and Bernoulli \((\pi_Z = 0.5)\) distributions, respectively. When \( R_1 = 0 \), a survival time \( T_\lambda^* \) was drawn from exponential \((\lambda)\) with mean \( 1/\lambda \). When

### Table 1

<table>
<thead>
<tr>
<th>Induction</th>
<th>Intensification</th>
<th>( \hat{\mu} )</th>
<th>( \hat{\mu}' )</th>
<th>( \hat{\mu}'' )</th>
</tr>
</thead>
<tbody>
<tr>
<td>GM-CSF</td>
<td>I</td>
<td>274.7 (43.2)</td>
<td>275.6 (26.7)</td>
<td>276.4 (20.4)</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>271.1 (42.4)</td>
<td>270.2 (26.4)</td>
<td>269.4 (20.1)</td>
</tr>
<tr>
<td>Placebo</td>
<td>I</td>
<td>292.8 (43.4)</td>
<td>305.0 (24.4)</td>
<td>309.5 (21.1)</td>
</tr>
<tr>
<td></td>
<td>II</td>
<td>337.9 (46.6)</td>
<td>325.0 (26.2)</td>
<td>319.0 (22.3)</td>
</tr>
</tbody>
</table>
Survival Distributions of Treatment Policies

53

Table 2
Monte Carlo coverage and relative efficiency for estimation of survival probabilities, 1000
data sets; entries for estimators are coverage expressed as a percentage (nominal level
is 95%); values in parentheses are relative efficiencies with respect to \( \hat{S}_{11}(t) \) or \( \hat{S}_{12}(t) \)

<table>
<thead>
<tr>
<th>( t )</th>
<th>( \tau_R )</th>
<th>( E(T^*_1)/L )</th>
<th>( S_{11}(t) )</th>
<th>( S_{12}(t) )</th>
<th>( \hat{S}_{11}(t) )</th>
<th>( \hat{S}_{11}'(t) )</th>
<th>( \hat{S}_{12}(t) )</th>
<th>( \hat{S}_{12}'(t) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.2</td>
<td>0.3</td>
<td>0.390</td>
<td>0.412</td>
<td>94.4</td>
<td>94.6 (0.96)</td>
<td>94.4 (1.06)</td>
<td>95.6</td>
</tr>
<tr>
<td>0.5</td>
<td>0.3</td>
<td>0.397</td>
<td>0.559</td>
<td>94.9</td>
<td>94.8 (1.01)</td>
<td>95.3 (1.06)</td>
<td>93.4</td>
<td>94.3 (0.97)</td>
</tr>
<tr>
<td>1.0</td>
<td>0.2</td>
<td>0.3</td>
<td>0.481</td>
<td>0.534</td>
<td>95.0</td>
<td>94.7 (0.94)</td>
<td>94.9 (1.11)</td>
<td>95.4</td>
</tr>
<tr>
<td>0.5</td>
<td>0.3</td>
<td>0.537</td>
<td>0.626</td>
<td>94.4</td>
<td>93.5 (1.00)</td>
<td>94.1 (1.09)</td>
<td>94.7</td>
<td>94.6 (1.02)</td>
</tr>
<tr>
<td>0.5</td>
<td>0.3</td>
<td>0.219</td>
<td>0.300</td>
<td>92.9</td>
<td>92.6 (1.17)</td>
<td>92.5 (1.59)</td>
<td>95.0</td>
<td>92.6 (0.94)</td>
</tr>
<tr>
<td>0.5</td>
<td>0.296</td>
<td>0.378</td>
<td>95.9</td>
<td>92.5 (1.21)</td>
<td>93.7 (1.32)</td>
<td>94.9</td>
<td>95.7 (1.14)</td>
<td>95.6 (1.25)</td>
</tr>
<tr>
<td>0.5</td>
<td>0.2</td>
<td>0.3</td>
<td>0.390</td>
<td>0.412</td>
<td>95.6</td>
<td>94.6 (0.92)</td>
<td>95.6 (1.06)</td>
<td>95.9</td>
</tr>
<tr>
<td>0.5</td>
<td>0.3</td>
<td>0.481</td>
<td>0.534</td>
<td>95.0</td>
<td>94.3 (0.98)</td>
<td>94.4 (1.05)</td>
<td>94.7</td>
<td>95.3 (0.94)</td>
</tr>
<tr>
<td>1.0</td>
<td>0.2</td>
<td>0.3</td>
<td>0.481</td>
<td>0.534</td>
<td>95.0</td>
<td>94.6 (1.25)</td>
<td>94.8 (1.40)</td>
<td>94.5</td>
</tr>
<tr>
<td>0.5</td>
<td>0.3</td>
<td>0.277</td>
<td>0.309</td>
<td>94.0</td>
<td>94.7 (1.14)</td>
<td>94.4 (1.21)</td>
<td>93.5</td>
<td>93.1 (1.05)</td>
</tr>
<tr>
<td>0.5</td>
<td>0.3</td>
<td>0.219</td>
<td>0.300</td>
<td>92.9</td>
<td>92.6 (1.17)</td>
<td>92.5 (1.59)</td>
<td>95.0</td>
<td>92.6 (0.94)</td>
</tr>
<tr>
<td>0.5</td>
<td>0.296</td>
<td>0.378</td>
<td>95.9</td>
<td>92.5 (1.21)</td>
<td>93.7 (1.32)</td>
<td>94.9</td>
<td>95.7 (1.14)</td>
<td>95.6 (1.25)</td>
</tr>
</tbody>
</table>

Table 2 shows coverage and relative efficiency results for estimating survival probabilities, 1000
Monte Carlo data sets; entries for estimators are coverage expressed as a percentage (nominal level
is 95%); values in parentheses are relative efficiencies with respect to \( \hat{S}_{11}(t) \) or \( \hat{S}_{12}(t) \).

For each of 1000 Monte Carlo data sets, \( P(T_{1k} > t) \), \( k = 1, 2 \), was estimated using (2), (3), and (7) at \( t = 0.5 \) and \( t = 1.0 \), representing time points early and later in the study. The
parameters \( \alpha, \beta_1, \) and \( \beta_2 \) were selected by specifying \( E(T^*_1)/L, E(T^*_1)/L \), and \( E(T^*_2)/L \) to be 0.1, 0.5, 1.0, respectively, yielding \( \alpha = 6.67, \beta_1 = 0.29, \) and \( \beta_2 = -0.67. \) The two levels of \( E(T^*_1)/L \)
were 0.3 and 0.5, corresponding to \( \mu_1 = 0.3 \) and \( \mu_2 = 0.5 \), to the treatment
policies and estimate the quantities of interest for each policy using only the data in the corre-
sponding group. This approach would yield estimates of the counterfactual survival distributions directly but uses information from nonremitting patients inefficiently. In fact, the
methods we propose could also be used with this design to increase efficiency; the important issue is not the two stages of randomization but the way in which our estimators make
use of the available data, by borrowing information on non-
remitting patients across treatment policies.

We have focused on estimation of and associated tests for
survival distributions and mean restricted survival. An
important open problem is finding optimal tests for treatment
policy differences against various alternatives. We are investi-
gating adaptation of the weighted Kaplan-Meier strategy of
Pepe and Fleming (1989) to this setting.

We take an intent-to-treat perspective, considering treat-
ment policies, so we handle data from patients who do not
Monte Carlo coverage and relative efficiency for estimation of restricted survival means based on 1000 data sets; entries for estimators are coverage expressed as a percentage (nominal level is 95%); values in parentheses are relative efficiencies with respect to $\hat{\mu}_{11}$ or $\hat{\mu}_{12}$

<table>
<thead>
<tr>
<th>$\pi_R$</th>
<th>$E(T_{\alpha}^s)/L$</th>
<th>$\mu_{11}$</th>
<th>$\mu_{12}$</th>
<th>$\hat{\mu}_{11}$</th>
<th>$\hat{\mu}_{12}$</th>
<th>$\hat{\mu}_{11}'$</th>
<th>$\hat{\mu}_{12}'$</th>
<th>$\hat{\mu}_{11}''$</th>
<th>$\hat{\mu}_{12}''$</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.3</td>
<td>0.502</td>
<td>0.537</td>
<td>92.8</td>
<td>92.6 (1.66)</td>
<td>92.6 (2.03)</td>
<td>93.0</td>
<td>92.6 (1.62)</td>
<td>90.8 (1.94)</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.673</td>
<td>0.709</td>
<td>92.9</td>
<td>93.5 (1.72)</td>
<td>93.7 (1.84)</td>
<td>93.7</td>
<td>94.2 (1.81)</td>
<td>93.8 (1.95)</td>
</tr>
<tr>
<td>0.5</td>
<td>0.3</td>
<td>0.604</td>
<td>0.693</td>
<td>93.8</td>
<td>93.1 (2.17)</td>
<td>93.3 (2.52)</td>
<td>93.7</td>
<td>94.0 (2.31)</td>
<td>93.3 (2.78)</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.711</td>
<td>0.800</td>
<td>93.1</td>
<td>94.0 (2.25)</td>
<td>93.2 (2.42)</td>
<td>94.8</td>
<td>93.7 (2.35)</td>
<td>93.3 (2.57)</td>
</tr>
<tr>
<td>n = 500</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.2</td>
<td>0.3</td>
<td>0.502</td>
<td>0.537</td>
<td>94.5</td>
<td>94.1 (1.60)</td>
<td>93.1 (1.87)</td>
<td>94.5</td>
<td>94.5 (1.69)</td>
<td>94.3 (2.22)</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.673</td>
<td>0.709</td>
<td>94.5</td>
<td>95.8 (1.78)</td>
<td>94.9 (1.87)</td>
<td>93.6</td>
<td>93.3 (1.92)</td>
<td>93.5 (2.26)</td>
</tr>
<tr>
<td>0.5</td>
<td>0.3</td>
<td>0.604</td>
<td>0.693</td>
<td>95.1</td>
<td>95.7 (2.22)</td>
<td>95.2 (2.57)</td>
<td>94.0</td>
<td>94.1 (2.32)</td>
<td>93.1 (3.02)</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>0.711</td>
<td>0.800</td>
<td>93.3</td>
<td>94.3 (2.32)</td>
<td>93.4 (2.50)</td>
<td>94.6</td>
<td>95.4 (2.78)</td>
<td>95.7 (3.21)</td>
</tr>
</tbody>
</table>

consent to maintenance therapy according to this principle. Our approach does not attempt to deduce causal treatment effects of regimens of the form $A_j$ followed by $B_k$ if remission, taking into account the confounding role of consent to the second-stage therapy. We are investigating methods to address this issue.

Acknowledgements

This work was supported by NIH grants R01-AL31789, R01-CA85848, and R01-CA51692. The authors are grateful to Steve George, Richard Dodge, and the CALGB 8923 team for providing the data used in the example.

Résumé

Dans certains essais cliniques, les patients sont randomisés une première fois dans un groupe de traitement à leur entrée dans l’étude; puis s’ils font partie des patients pour lesquels on constate une rémission de la maladie, ils sont rerandomisés dans un autre groupe de traitement, en phase de maintenance. Idéalement, l’analyse devrait permettre de comparer différentes stratégies thérapeutiques globales, c’est-à-dire des combinaisons définies à l’avance—des groupes de traitement inclus en phase de maintenances. En réalité, la pratique la plus courante est de mener des analyses séparées pour les traitements à l’inclusion et en phase de maintenance, ce qui constitue un moyen bien indirect d’aborder la question. Pour analyser ces essais en deux phases, nous proposons ici, pour chaque stratégie thérapeutique combinée, des estimateurs convergents de la distribution de la survie et du temps moyen de survie restreint, estimateurs dont nous étudions la propriété asymptotique. Nous illustrons l’utilisation et les propriétés de ces méthodes sur les données d’un essai clinique conduit dans la leucémie, ainsi qu’au travers de simulations.

References


Appendix A

Large-Sample Properties

We sketch arguments to show consistency and asymptotic normality for the proposed estimators of $\hat{F}_{1k}(t)$; arguments for mean estimators are analogous. Consistency of $\hat{F}_{1k}(t)$ and $\hat{F}_{1k}'(t)$ may be deduced straightforwardly by writing

$$\hat{F}_{1k}(t) - F_{1k}(t) = n^{-1} \sum_{i=1}^{n} \left\{ \frac{\Delta_i Q_{ki}}{K(T_i)} I(T_i \leq t) - F_{1k}(t) \right\}$$

$$- n^{-1} \sum_{i=1}^{n} \left( \Delta_i Q_{ki} I(T_i \leq t) \right) \times \left\{ \frac{K(T_i)}{K(T_i)} - K(T_i) \right\} , \quad (A.1)$$

References


That the right-hand side of each of (A.1) and (A.2) is $o_p(1)$ may be established using arguments similar to those in the appendix of Zhao and Tsiatis (1997).

To derive the large-sample distributions of $n^{1/2}\{\hat{F}_{1k}(t) - F_{1k}(t)\}$ and $n^{1/2}\{\hat{F}_{1k}(t) - F_{1k}(t)\}$, define the filtration $\mathcal{F}_n(t)$ as the increasing sequence of sub-$\sigma$ algebras $\sigma\{I(C_i \leq u), u \leq t; R_i, Z_i, T_{1i}, T_{2i}, i = 1 \cdots n\}$ containing all observed censoring and survival information up to time $t$ and all information on the potential survival times. Letting $\lambda^c(u)$ be the hazard function for the censoring distribution, the corresponding martingale process is $M^c(t) = N^c(t) - \int_0^t \lambda^c(u) Y_i(u) du$, where $N^c(t) = I(Y_i = t, \Delta_i = 0)$ and $Y_i(u) = I(V_i \geq u)$. With $M^c(u) = \sum_{i=1}^n M^c_i(u)$ and $Y(u) = \sum_{i=1}^n Y_i(u)$, (A.4)–(A.6) of Zhao and Tsiatis (1997) yield

$$n^{1/2}\{\hat{F}_{1k}(t) - F_{1k}(t)\} = n^{-1/2} \sum_{i=1}^n \psi_{ki} + o_p(1),$$

where $\hat{F}_{1k}(t) = \left\{\int_0^t \frac{dM^c_i(u)}{K(u)} \{Q_{ki}I(T_i \leq t) - \hat{G}_{1k}(t,u)\}\right\}^{-1} X \left\{\int_0^t \frac{dM^c_i(u)}{K(u)} \{Q_{ki}I(T_i \leq t) - \hat{G}_{1k}(t,u)\}\right\}$

and

$G_{1k}(t,u) = \frac{E\{I(T_{1ki} \leq t)I(T_{1ki} \geq u)|T_{1ki} > u\}}{P(T_{1ki} > u)}.$

(A.3) follows by arguments similar to those in Zhao and Tsiatis (1997). Thus, from (A.3),

$$n^{1/2}\{\hat{F}_{1k}(t) - F_{1k}(t)\} = n^{-1/2} \sum_{i=1}^n \psi_{ki} + o_p(1),$$

where

$$\psi_{ki} = Q_{ki}I(T_i \leq t) - F_{1k}(t) - \int_0^t \frac{Q_{ki}I(T_i \leq t) - G_{1k}(t,u)}{K(u)} dM^c_i(u).$$

Thus, $\hat{F}_{1k}(t)$ is an asymptotically linear estimator, i.e., $n^{1/2}$ times the estimator minus $F_{1k}(t)$ is equal to $n^{-1/2}$ times the sum of i.i.d. mean-zero random variables plus a term of $o_p(1)$, where a term in the sum is referred to as the influence function of the estimator. It follows that the influence function for $\hat{F}_{1k}(t)$ is given by (A.4). The asymptotic variance of the estimator is the variance of the influence function, a fact we make use of momentarily.

Because $Q_{ki}I(T_i \leq t) - F_{1k}(t)$ has mean zero and is $\mathcal{F}(0)$ measurable and the second component of (A.4) has mean zero and is uncorrelated with the first, it follows that $n^{1/2}\{\hat{F}_{1k}(t) - F_{1k}(t)\}$ converges to a mean-zero normal distribution with variance

$$E(\psi_{ki})^2 = E\{Q_{ki}I(T_i \leq t) - F_{1k}(t)\}^2 + \int_0^L E\{L_{1k}(t,u)\}^2 \lambda^c(u) du,$$

where $L_{1k}(t,u) = \{Q_{ki}I(T_i \leq t) - G_{1k}(t,u)\} I(T_i > u)$. This variance may be estimated by (9). An entirely similar argument may be used to show that

$$n^{1/2}\{\hat{F}_{1k}(t) - F_{1k}(t)\} = n^{-1/2} \sum_{i=1}^n \psi_{ki} + o_p(1),$$

where the influence function is given by

$$\psi_{ki} = Q_{ki}I(T_i \leq t) - F_{1k}(t) - \int_0^L \frac{Q_{ki}I(T_i \leq t) - G_{1k}(t,u)}{K(u)} dM^c_i(u),$$

(A.5)

with

$$G_{1k}(t,u) = \frac{E\{I(T_{1ki} \leq t) - F_{1k}(t)\}I(T_{1ki} \geq u)}{P(T_{1ki} > u)}.$$

This leads to the variance

$$E(\psi_{ki})^2 = E\{Q_{ki}I(T_i \leq t) - F_{1k}(t)\}^2 + \int_0^L E\{L_{1k}(t,u)\}^2 \lambda^c(u) du,$$

where

$L_{1k}(t,u) = \{Q_{ki}I(T_i \leq t) - F_{1k}(t)\} - G_{1k}(t,u) I(T_i > u)$. This variance may be estimated by (10).

To derive (7), the influence function corresponding to (6) for fixed $a_{1k}$, with $G_{Q}(u) = E\{(Q_{ki} - 1)I(T_i \geq u)\} / P(T_i > u)$ and using (A.4) and (A.5), is given by

$$\psi_{ki} - a_{1k} \left\{Q_{ki} - 1 - \int_0^L \frac{Q_{ki} - 1 - G_{Q}(u)}{K(u)} dM^c_i(u)\right\};$$

(A.6)

thus, $a_{1k}$ should be chosen to minimize the variance

$$E\left[\psi_{ki} - a_{1k} \left\{Q_{ki} - 1 - \int_0^L \frac{Q_{ki} - 1 - G_{Q}(u)}{K(u)} dM^c_i(u)\right\}ight]^2,$$
$$\alpha_{1k} = \left[ E\{Q_{ki}(Q_{ki} - 1)I(T_i \leq t)\} \right. \\
\left. + \int_0^L \lambda^c(u)K(u)^{-1}E\{L_{1k}^\alpha(t, u)\}du \right] \\
\left. + \left[ E(Q_{ki} - 1)^2 + \int_0^L \lambda^c(u)K(u)^{-1}E\{G_{ki}^\alpha(u)\}du \right] \\
\right),$$

where

$$E\{L_{1k}^\alpha(u)\} = E\{(Q_{ki}I(T_i \leq t) - G_{1k}(t, u)) \times (Q_{ki} - 1 - G_{Q_k}(u))I(T_i \geq u)\}$$

and

$$E\{G_{ki}^\alpha(u)\} = E\{(Q_{ki} - 1 - G_{Q_k}(u))^2I(T_i \geq u)\},$$

leading to (A.7).

Consistency of $\hat{F}_{1k}''(t)$ follows by noting

$$\hat{F}_{1k}''(t) - F_{1k}(t) = F_{1k}(t) - F_{1k}(t) \\
- \hat{\alpha}_{1k} \left\{ n^{-1} \sum_{i=1}^n \frac{\Delta_i}{K(T_i)}(Q_{ki} - 1) \right\} \\
= \hat{\alpha}_{1k} \left\{ n^{-1} \sum_{i=1}^n \frac{\Delta_i}{K(T_i)}(Q_{ki} - 1) \right\} + o_p(1).$$

(A.8)

The expectations in (A.7) can be consistently estimated as in (8), and the term in braces may be shown to converge in probability to zero by arguments similar to those used to show the right-hand sides of (A.1) and (A.2) are $o_p(1)$. Moreover,

$$n^{1/2}\left\{ \hat{F}_{1k}''(t) - F_{1k}(t) \right\} \\
= n^{1/2}\left\{ \hat{F}_{1k}(t) - F_{1k}(t) \right\} \\
- \alpha_{1k} \left\{ n^{-1/2} \sum_{i=1}^n \frac{\Delta_i}{K(T_i)}(Q_{ki} - 1) \right\} \\
+ (\alpha_{1k} - \hat{\alpha}_{1k}) \left\{ n^{-1/2} \sum_{i=1}^n \frac{\Delta_i}{K(T_i)}(Q_{ki} - 1) \right\}. $$

The techniques used to obtain (A.3) can be used to show the term in braces is an asymptotically linear estimator of zero with influence function equal to the bracketed term in (A.6). Because $(\alpha_{1k} - \hat{\alpha}_{1k}) \to 0$, estimating $\alpha_{1k}$ has no effect asymptotically. Hence, $\psi_{ki}''$, the influence function of $\hat{F}_{1k}''(t)$, is given by (A.6) substituting (A.7) and has variance

$$E(\psi_{ki}''^2) = E\{Q_{ki}I(T_i \leq t) - \hat{F}_{1k}(t) - \alpha_{1k}(Q_{ki} - 1)\}^2 \\
+ \int_0^L E\{L_{1k}''(t, u)\}^2 \lambda^c(u)du,$$

where

$$L_{1k}''(t, u) = [Q_{ki}I(T_i \leq t) - G_{1k}(t, u)$$

and

$$- \alpha_{1k}(Q_{ki} - 1 - G_{Q_k}(u))]I(T_i \geq u).$$

This variance may be estimated by (11).

**APPENDIX B**

**Covariance of the Estimators**

Large-sample covariances for the proposed estimators for $S_{11}(t)$ and $S_{12}(t)$ are given by the expectation of the product of the corresponding influence functions and are given by

$$E(\psi_{11}\psi_{21}) = E\{Q_{11}Q_{21}I(T_i \leq t) - F_{11}(t)F_{12}(t)$$

$$+ \int_0^L [E\{L_{11}''(t, u)\}K(u)]\lambda^c(u)du,$$

$$E(\psi_{12}\psi_{21}) = E\{Q_{11}Q_{21}I(T_i \leq t) - F_{11}(t)$$

$$\times \{I(T_i \leq t) - F_{12}(t)\}$$

$$+ \int_0^L [E\{L_{11}''(t, u)\}K(u)]\lambda^c(u)du,$$

and

$$E(\psi_{11}\psi_{22}') = E\{Q_{11}I(T_i \leq t) - F_{11}(t) - \alpha_{11}(Q_{11} - 1)$$

$$\times \{I(T_i \leq t) - F_{12}(t) - \alpha_{12}(Q_{21} - 1)\}$$

$$+ \int_0^L [E\{L_{11}''(t, u)\}K(u)]\lambda^c(u)du.$$

Estimators for these covariances are

$$\hat{\text{cov}}\{\hat{S}_{11}(t), \hat{S}_{12}(t)\} \\
= n^{-1} \left\{ n^{-1} \sum_{i=1}^n \frac{\Delta_iQ_{11}Q_{21}}{K(V_i)}I(V_i \leq t) - \hat{F}_{11}(t)\hat{F}_{12}(t) \\
+ \int_0^L \frac{d\gamma^c(u)}{K(u)Y(u)}E\{L_{11}''(t, u)\}K(u)\lambda^c(u)du, \right\}$$

$$\hat{\text{cov}}\{\hat{S}_{11}(t), \hat{S}_{12}(t)\} \\
= n^{-1} \left\{ n^{-1} \sum_{i=1}^n \frac{\Delta_iQ_{11}Q_{21}}{K(V_i)}I(V_i \leq t) - \hat{F}_{11}(t) \\
\times \{I(V_i \leq t) - \hat{F}_{12}(t)\} \\
+ \int_0^L \frac{d\gamma^c(u)}{K(u)Y(u)}E\{L_{11}''(t, u)\}K(u)\lambda^c(u)du, \right\},$$

and

$$\hat{\text{cov}}\{\hat{S}_{11}(t), \hat{S}_{12}(t)\} \\
= n^{-1} \left\{ n^{-1} \sum_{i=1}^n \frac{\Delta_i}{K(V_i)} \\
\times \{Q_{11}I(V_i \leq t) - \hat{F}_{11}(t) - \alpha_{11}(Q_{11} - 1) \\
\times \{Q_{21}I(V_i \leq t) - \hat{F}_{12}(t) - \alpha_{12}(Q_{21} - 1)\} \\
+ \int_0^L \frac{d\gamma^c(u)}{K(u)Y(u)}E\{L_{11}''(t, u)\}K(u)\lambda^c(u)du, \right\}. $$
Survival Distributions of Treatment Policies

where

\[
\hat{E}\{L_{11i}(t, u)L_{12i}(t, u)\} = n^{-1} \sum_{i=1}^{n} \Delta_i \{Q_{1i}I(V_i \leq t) - \hat{G}_{1i}(t, u)\} \\
\times \{Q_{2i}I(V_i \leq t) - \hat{G}_{12}(t, u)\} \frac{I(V_i \geq u)}{K(V_i)},
\]

and

\[
\hat{E}\{L_{11i}(t, u)L_{12i}''(t, u)\} = n^{-1} \sum_{i=1}^{n} \Delta_i \{Q_{1i}I(V_i \leq t) - \hat{G}_{1i}'(t, u)\} \\
\times \{Q_{2i}I(V_i \leq t) - \hat{G}_{12}'(t, u)\} \\
\times \{Q_{2i}I(V_i \leq t) - \hat{G}_{22}'(t, u)\} \\
\times \{Q_{2i}I(V_i \leq t) - \hat{G}_{22}(t, u)\} \\
\times \frac{I(V_i \geq u)}{K(V_i)}.
\]

Substitution of \(V_i\) for \(I(V_i \leq t)\) and replacing estimates of \(F_{1k}(t)\) with the appropriate estimates of \(\mu_{1k}\) yields covariance estimators for mean restricted survival.