Supplementary materials for "Inference on treatment effects from a randomized clinical trial in the presence of premature treatment discontinuation: The SYNERGY trial"

MIN ZHANG*

Department of Biostatistics, University of Michigan, Ann Arbor, MI 48109-2029, USA

ANASTASIOS A. TSIATIS, MARIE DAVIDIAN

Department of Statistics, North Carolina State University, Raleigh, NC 27695-8203 USA

KAREN S. PIEPER, KENNETH W. MAHAFFEY

Duke Clinical Research Institute, Durham, NC 27705, USA

**email:* mzhangst@umich.edu

Throughout, notation is as defined in the main paper, and citations are to references given in the main paper.

A. SYNERGY patient population

Patients with acute coronary syndrome (ACS) experience one or more of a set of signs and symptoms of coronary artery disease, including unstable angina, which is chest discomfort that changes or worsens; and two different forms of myocardial infarction (MI), ST-segment elevation MI and non-ST-segment elevation (NTSE) MI. ST-segment is a portion of the pattern of an electrocardiogram, elevation of which relative to a normal pattern is associated with MI. NTSE indicates evidence of a MI based on enzyme elevation but without confirmation from the electrocardiogram. Subjects in SYNERGY were drawn from the NTSE ACS population who were at risk for undergoing an invasive procedure to treat narrowed coronary arteries, such as percutaneous coronary intervention, also known as angioplasty, or coronary

bypass grafting.

B. Justification of weighting scheme

We sketch an argument suggesting that replacing $Y_i^*(u)$ in (3.3) of the main paper by (4.3)(a) as in the estimating equation (4.4) should have the desired effect of appropriately weighting the contributions of subjects who do not optionally discontinue assigned treatment; the argument for (4.3)(b) is similar.

Because $dN_i(u) = dN_i^*(u)$ and $S_i = S_i^*$ when $O_i \geq u$, (4.3)(a) becomes

$$\frac{I(O_i \geq u) dN_i^*(u)}{K\{u, Z_i, Q_i(\cdot), S_i^*\}}. \tag{B.1}$$

We argue that $E\{(4.3)(a)|Z_i, Q_i^*\} = dN_i^*(u)$ for $u \geq 0$, which suggests that the proposed weighting scheme will mimic the contributions of the subjects not optionally discontinuing treatment, as desired. Using (B.1), $E\{(4.3)(a)|Z_i, Q_i^*\}$ is equal to

$$E\left(\frac{I(O_i \geq u) dN_i^*(u)}{\{1 - p_0(Z_i, X_i)\} \exp\left[-\int_0^{u \vee S_i^*} q\{s, Z_i, Q_i(s)\}\, ds\right]} \middle| Z_i, Q_i^* \right). \tag{B.2}$$

By the key missing at random assumption, (B.2) becomes

$$\begin{aligned} &E\left[\frac{I(O_i \geq u) dN_i^*(u)}{\{1 - p_0(Z_i, Q_i^*)\} \exp\left\{-\int_0^{u \vee S_i^*} q(s, Z_i, Q_i^*)\, ds\right\}} \middle| Z_i, Q_i^* \right] \\ &\quad = \frac{E\{I(O_i \geq u)|Z_i, Q_i^*\} dN_i^*(u)}{\{1 - p_0(Z_i, Q_i^*)\} \exp\left\{-\int_0^{u \vee S_i^*} q(s, Z_i, Q_i^*)\, ds\right\}}. \end{aligned} \tag{B.3}$$

It it may be shown that the conditional expectation in the numerator of (B.3) is equal to the denominator, which follows by writing

$$E\{I(O_i \geq u)|Z_i, Q_i^*\} = \{1 - p_0(Z_i, Q_i^*)\} \exp\left\{-\int_0^u q(s, Z_i, Q_i^*)\, ds\right\},$$

and noting that $q(s, Z_i, Q_i^*) = 0$ when $s > S_i^*$. Thus, under the positivity assumption $K\{u, Z, Q(\cdot), S)\} \geq \epsilon > 0$ for all $u \geq 0$, the result follows.

C. Large sample properties and discussion of weights

We argue heuristically that the estimators solving Equation (4.4) of the main paper are consistent and asymptotically normal. For brevity, define $\widetilde{X} = (Z, X^T)^T$, and let

$$\overline{(Z,X)}(u,\beta,\gamma) = \frac{\sum_{j=1}^n \widetilde{X}_j \exp(\beta Z_j + \gamma^T X_j) w(u, Z_j, X_j) I(O_j \geq u) Y_j(u) / K\{u, Z_j, Q_j(\cdot), S_j\}}{\sum_{j=1}^n \exp(\beta Z_j + \gamma^T X_j) w(u, Z_j, X_j) I(O_j \geq u) Y_j(u) / K\{u, Z_j, Q_j(\cdot), S_j\}}.$$

Dividing numerator and denominator by n and using iterated conditional expectations yields

$$\overline{(Z,X)}(u,\beta,\gamma) \xrightarrow{p} \frac{E\{\widetilde{X} \exp(\beta Z + \gamma^T X) w(u, Z, X) Y^*(u)\}}{E\{\exp(\beta Z + \gamma^T X) w(u, Z, X) Y^*(u)\}} = \mu_{ZX}(u,\beta,\gamma),$$

say. Adding and subtracting a common term in (4.4) of the main paper, write this as

$$\begin{aligned} &\sum_{i=1}^n \int \{\widetilde{X}_i - \overline{(Z,X)}(u,\beta,\gamma)\} \frac{I(O_i \geq u) w(u, Z_i, X_i)}{K\{u, Z_i, Q_i(\cdot), S_i\}} dL_i^*(u,\beta,\gamma) \\ &= \sum_{i=1}^n \int \{\widetilde{X}_i - \mu_{ZX}(u,\beta,\gamma)\} \frac{I(O_i \geq u) w(u, Z_i, X_i)}{K\{u, Z_i, Q_i(\cdot), S_i\}} dL_i^*(u,\beta,\gamma) + o_p(n^{1/2}), \quad \text{(C.1)} \end{aligned}$$

where $dL_i^*(u,\beta,\gamma) = dN_i^*(u) - \lambda_0(u) \exp(\beta Z_i + \gamma^T X_i) Y_i^*(u)$. Therefore, under appropriate regularity conditions, the solution in (β, γ) of (4.4), $(\widehat{\beta}_n, \widehat{\gamma}_n)$, say, is asymptotically equivalent to the solution of (C.1) set equal to 0, $(\widetilde{\beta}_n, \widetilde{\gamma}_n)$, say, in the sense that $n^{1/2}(\widehat{\beta}_n - \widetilde{\beta}_n)$ and $n^{1/2}(\widehat{\gamma}_n - \widetilde{\gamma}_n) \xrightarrow{p} 0$. Equation (C.1) is a sum of independent and identically distributed quantities, which we write as $\sum_{i=1}^n m(W_i, \beta, \gamma)$, where

$$m(W,\beta,\gamma) = \int \{\widetilde{X} - \mu_{ZX}(u,\beta,\gamma)\} \frac{I(O \geq u) w(u, Z, X)}{K\{u, Z, Q(\cdot), S_j\}} dL^*(u,\beta,\gamma).$$

The expected value of this estimating function, using an argument analogous to that in Section B above, is equal to $E\left\{\int \{\widetilde{X} - \mu_{ZX}(u,\beta,\gamma)\} w(u, Z, X) dL^*(u,\beta,\gamma)\right\}$. Because this term is a martingale with respect to the counting process $L^*(u)$, it has mean zero. Thus, as $m(W,\beta,\gamma)$ has mean zero, $(\widetilde{\beta}_n, \widetilde{\gamma}_n)$ is an M-estimator (Stefanski and Boos, 2002) and hence is consistent and asymptotically normal with variance that can be consistently estimated using the sandwich method, and the same is true for $(\widehat{\beta}_n, \widehat{\gamma}_n)$.

The foregoing argument holds for any choice of $w(u, Z, X)$ when interest focuses on conditional (on X) inference in the context of (3.2) of the main paper. For the consistency result, $w(u, Z, X)$ must depend directly on u, not $u \vee S$. As discussed in the main paper, possible choices include $w(u, Z, X) \equiv 1$ and $w(u, Z, X) = \{1 - p_0(Z, X)\} \exp\big\{-\int_0^u r(s, Z, X)\, ds\big\}$, where $r(u, Z, X) = \lim_{h \to 0} h^{-1} \Pr(u \leq S_i \leq u + h, E_i = 1 | S_i \geq u, Z, X)$. When interest focuses on β in the model (3.1) of the main paper, so on unconditional inference on treatment effect, as discussed below (4.4), $w(u, Z, X)$ should not depend on X, as noted in the main paper.

At the beginning of Section 3 of the main paper, it is noted that the methods require that subjects who mandatorily discontinue assigned treatment prior to $t_{\max}$ are followed to $t_{\max}$, so that survival/censoring information is available for these subjects after the time of discontinuation to $t_{\max}$, while such information is not required for subjects who are observed to optionally discontinue study treatment. (It is assumed that information on survival/censoring up to $t_{\max}$ is available for all subjects who are not observed to discontinue assigned treatment for any reason, as would be the case in a usual study with no treatment discontinuation.) Inspection of Equation (4.4) reveals why this is so. Subjects who are observed to optionally discontinue assigned treatment have weight $\kappa(u, W_i) = 0$ for all u after the time of optional discontinuation in Equation (4.4) (because $I(O_i \geq u) = 0$ for all such u). Thus, these subjects do not contribute to the integrand after this time, and hence the information on whether or not they die or are administratively censored after the time of optional discontinuation will not be incorporated in the equation. In contrast, this information is required for subjects who are observed to mandatorily discontinue assigned treatment up to $t_{\max}$. This is because, for these subjects, $\kappa(u, W_i) > 0$ for all u subsequent to the time of mandatory discontinuation up to $t_{\max}$, because these subjects are no longer eligible to optionally discontinue (so $I(O_i \geq u) = 1$ for all such u). Accordingly, these subjects

continue to contribute to the integrand after this time, and hence information on whether or not they are at risk and/or die from this point on up to $t_{\max}$ is needed in the equation.

D. Practical implementation

We first outline a series of steps to be carried out to implement the methods in practice, which requires postulating and fitting models for $q\{u, Z, Q(u)\}$ and $p_0(Z, X)$, and hence $K\{u, Z, Q(\cdot), S\}$ (and $r(u, Z, X)$ if applicable).

1. Categorize treatment discontinuation for each subject as mandatory or optional.
2. Fit models used to form $K\{u, Z, Q(\cdot), S\}$ and $w(u, Z, X)$. For $u > 0$, for each $i = 1, \ldots, n$ for whom $O_i > 0$, if i was observed to discontinue treatment for optional reasons prior to $t_{\max}$, set S_i equal to the time of optional discontinuation and $\Lambda_i = 1$ (here, $E_i = 1$); otherwise, set S_i equal to the minimum of i's observed time to mandatory discontinuation or treatment completion prior to $t_{\max}$ and time to failure or censoring, and set $\Lambda_i = 0$ (here, $E_i > 1$). Thus, Λ_i is the "censoring indicator" for optional discontinuation time. For $K\{u, Z, Q(\cdot), S\}$, to model $q\{u, Q(u)\}$, postulate and fit to these data a proportional hazards model conditional on treatment assignment Z_i and baseline and post-randomization covariates X_i and $V_i^H(u)$. Alternatively, separate models conditional on X_i and $V_i^H(u)$ may be fitted for each treatment; see below. If there are optional discontinuations at 0, obtain $p_0(Z, X)$ by defining $H_i = I(O_i = 0)$ and postulate and fit a binary regression model for $\Pr(H_i = 1 | Z_i, X_i)$; e.g., logistic regression.

   If stabilized weights are to be used, if interest focuses on β in (3.2) of the main paper, then one may postulate a proportional hazards model for $r(u, Z, X)$ that depends only on Z and X and fit it to the data (S_i, Λ_i), $i = 1, \ldots, n$; alternatively, one may fit proportional hazards models in X separately by treatment. If instead interest is in β

in (3.1) of the main paper, one should take $w(u, Z, X)$ to not depend on X, which we write as $w(u, Z)$. Here, to form $w(u, Z)$, if there are optional discontinuations at time 0, a separate model $p_0(Z)$ not depending on X is required, which can be estimated by the treatment-specific sample proportions of optionally discontinuing treatment at 0, i.e., the estimate of $p_0(1)$ would be $\sum_{i=1}^{n} H_i Z_i / \sum_{i=1}^{n} Z_i$. The second term in $w(u, Z)$ at time u may be estimated based on proportional hazards model for $r(u, Z)$ with covariate Z or using treatment-specific Kaplan-Meier estimates of surviving to time u.

3. Based on the fitted models, for each i and u equal to every distinct event time across all subjects, estimate $K\{u, Z, Q(\cdot), S\}$, and, if not taking $w(u, Z, X) \equiv 1$, estimate $w(u, Z, X)$ or $w(u, Z)$ as appropriate. This may be accomplished using, for example, Breslow's estimator for survival probability.
4. For each i, create weights at each distinct event time u equal to 0 if s/he optionally discontinued treatment by u or the inverse of the estimate of $K\{u, Z, Q(\cdot), S\}$ from the previous step if s/he did not (multiplied by the estimate of $w(u, Z_i)$ or $w(u, Z_i, X_i)$ if not taking $w(u, Z, X) \equiv 1$).
5. Substitute the weights in (4.4) of the main paper and solve for β and γ, and estimate the variance of the estimators via the sandwich method. This may be implemented, for example, using SAS `proc phreg` (SAS Institute, 2006) with the counting process input format, a `weight` statement, and the `covs(aggregate)` option. The resulting score test for β corresponds to a "logrank" test for $\beta = 0$.

In the next section of this document, we present SAS code implementing these steps.

It is worth noting that, in fitting the proportional hazards models in step (2) above, one may regard the "censoring" indicated by Λ_i as administrative. The partial likelihood methods used to fit these models directly estimate the cause-specific hazards $q\{u, Z, Q(u)\}$ in (4.2) of the main paper and $r(u, Z, X)$, which for our purposes are the quantities of

interest. Ordinarily in routine survival analysis, one is interested in the net-specific hazard, and the assumption of independence of the event and censoring times is needed so that the cause-specific and net-specific hazards are equivalent, as discussed in general in Chapter 8 of Kalbfleisch and Prentice (2002). Because we are interested in cause-specific hazards directly, this assumption is not required.

Asymptotic theory for inverse weighted methods implies that, the more complex the models involved in the weights (in terms of numbers of parameters fitted), the better the precision of estimation of β and γ, suggesting using separate models for each treatment. However, finite sample degradation of performance can occur when numerous parameters in these models are estimated, suggesting use of more parsimonious models. This trade-off should be evaluated by the analyst. In Sections 5 and 6 of the main paper, we fitted models of both types.

Programs such as `proc phreg` treat the estimated weights as fixed, so do not take into account estimation of parameters in the models defining the weights in step (2) above. As noted by Robins *and others* (2000), this should lead to conservative confidence intervals and hypothesis tests based on "robust" sandwich errors calculated by the software. Faithful application of the sandwich theory to obtain standard errors accounting for this estimation is quite unwieldy. Alternatively, a nonparametric bootstrap could be used. In all simulations we have conducted, including those in Section 6 of the main paper, the effect of using the output standard errors "as-is," so ignoring the effect of estimating the weights, has been negligible.

As noted at the end of Section 3 of the main paper, the methods are also applicable to a binary endpoint. We consider the simple case, as in SYNERGY with $t_{\max} = 30$ days, where occurrence of the event if prior to $t_{\max}$ is observed without censoring for virtually all subjects. The observed endpoint is thus $R_i = I(U_i \leq t_{\max})$, where U_i is now the uncensored

event time if before $t_{\max}$ and otherwise is known to be $> t_{\max}$. Although some subjects may discontinue treatment for optional reasons prior to $t_{\max}$, so deviating from the regimes of interest, we may conceptualize potential outcomes and a logistic regression model analogous to (3.1) in which the (unconditional) effect of the two regimes is formally defined. To estimate consistently the log-odds ratio β, say, corresponding to this effect, a weighted version of the maximum likelihood score equations for logistic regression based on (R_i, Z_i), $i = 1, \ldots, n$, may be formulated, where the summand for subject i is multiplied by an estimate of $I(O_i \geq t_{\max})/K\{t_{\max}, Z, Q(\cdot), S\}$; here, stabilizing the weights has no effect. This may be implemented by using, e.g., SAS `proc genmod` (SAS Institute, 2006) with a `weight` statement. Again, inferences based on the output (default robust) standard errors may be conservative.

In the event the outcome is censored prior to $t_{\max}$ for some subjects, one can estimate the desired odds ratio by calculating the ratio of the estimated probabilities of the outcome occurring by $t_{\max}$ derived from fitting treatment-specific survival curves using inverse weighting as in Cole and Hernán (2003).

E. SAS code

The following code implements the inverse probability risk set weighted methods for a time-to-event outcome in a scenario similar to that of the simulations in Section 6 of the main paper but that also allows subjects who optionally discontinue assigned treatment at time zero. Code is given for two situations: the case where $w(u, Z, X) \equiv 1$, and the case where stabilized weights are used. The former is straightforward, and is shown first. Implementation with stabilized weights is considerably more complicated, owing to the fact that, while the denominator $K\{u, Z, Q(\cdot), S\}$ ceases to change after time S, the numerator will not; see below.

This code demonstrates implementation in the case where $w(u, Z, X) \equiv 1$ (no stabilizing of weights).

```
/*  In the dataset, every row corresponds to a distinct subject  */


/*  First model the probability of optional discontinuation at time 0.
    pH is the probability of optional discontinuation at 0 */

proc logistic data=dataset descending;
   model H=Z X1 X2;
   output out=initstop(keep= i pH) p=pH;
run;

data dataset; merge dataset initstop; by i;
run;

/*  create data in counting process input format for modeling
    time-to-optional-discontinuation if O>0.  For simplicity, time is
    rounded up to integers and we split time into intervals of length 1

    Lambda2 is the indicator for observing optional discontinuation in
    counting process input format.

    Lambda is as defined in the implementation section.

    Vh is a time-dependent covariate. */

data modelstop; set dataset;
   if H=0; keep H i d1 d2 S Lambda Lambda2 X1 X2 Vh D Z;
   do d2=1 to S;
      d1=d2-1;
      Vh=0;
      if d1>=D then Vh=1;
      Lambda2=0;
      if d2=S then Lambda2=Lambda;
      output;
   end;
run;

/*  model time-to-optional-discontinuation for creating weight -
    separately build model for each treatment group  */

proc phreg data=modelstop; where Z=0;
   id i;
   model (d1,d2)*Lambda2(0)= X1  X2 Vh;
   output out=stopprob0(keep=stopprob i d1 d2) survival=stopprob;
run;

proc phreg data=modelstop; where Z=1;
   id i;
   model (d1,d2)*Lambda2(0)= X1 X2 Vh;
   output out=stopprob1(keep=stopprob i d1 d2) survival=stopprob;
```

```
run;

data stopprob; set stopprob0 stopprob1;
run;

proc sort data=stopprob; by i d1 d2;
run;

/*  create data in counting process input format for modeling the event time.

    Delta2 is the indicator for observing event in counting process input
    format.

    Delta is as defined in the main paper */

data modeldth; set dataset;
   if H=0;
   keep  pH H i d1 d2 S Delta Delta2 U Z Lambda ;
   do d2=1 to S;
      d1=d2-1;
      Delta2=0;
      if d2=U then Delta2=Delta;
      output;
   end;
   if S<U then do;
      d1=S;
      d2=U;
      Delta2=Delta;
      output;
   end;
run;

/*  here we take w(u,Z,X) = 1.  If instead you want to take
    w(u,Z,X) to depend on (Z,X), must also fit another proportional
    hazards model with only these covariates  */

data modeldth; merge stopprob modeldth; by i d1 d2;
   stopprob2=lag1(stopprob);
   if stopprob=. then stopprob=stopprob2;
   invwt=1/(stopprob*(1-pH));
   if Lambda=1 and d2>S then invwt=0;
run;

/*  Solve (4.4) and get sandwich variance  */

proc phreg data=modeldth COVSANDWICH(aggregate);
   id i;
   model (d1,d2)*Delta2(0)=Z;
   weight invwt;
run;
```

This code demonstrates implementation in the case where $w(u, Z, X)$ is taken to depend on u; in this case, $w(u, Z)$ was defined by fitting a proportional hazards model as in Section

D. As noted above, this is considerably more complicated than the foregoing code due to the fact that the numerator $w(u, Z)$ continues to change with time while the denominator does not. This introduces additional complexity due to the need to convert the data set into the appropriate counting process format in order to solve equation (4.4) in the main paper using `proc phreg`. In fitting the proportional hazards model for $w(u, Z)$ in this case, because $w(u, Z)$ changes with time for $u > S$, one must further split the time interval from S to u into subintervals. The user should be aware that adaptation of this code to his or her situation will require some effort.

```
/*  In the dataset, every row corresponds to a distinct subject  */

/*  First model the probability of optional discontinuation at time 0 given Z
    and X. pH is the probability of optional discontinuation at 0 */

proc logistic data=dataset descending;
   model H=Z X1 X2;
   output out=initstop(keep= i pH) p=pH;
run;

/*  Model the probability of optional discontinuation at time 0 given Z for
    creating the stabilized weights */

proc logistic data=dataset descending;
   model H=Z;
   output out=numinitstop(keep= i numpH) p=numpH;
run;

data dataset; merge dataset  initstop numinitstop; by i; run;

/* Kaplan-Meier estimator for treatment specific survival functions
   for creating the stabilized weights */

proc phreg data=dataset;  where Z=0;
   model S*Lambda(0)=;
   baseline out=numstopprob0 survival=numstopprob0;
run;

proc phreg data=dataset;  where Z=1;
   model S*Lambda(0)=;
   baseline out=numstopprob1 survival=numstopprob1;
run;

data numstopprob0; set numstopprob0; d2=S; drop S;run;
data numstopprob1; set numstopprob1; d2=S; drop S;run;

/*  Create data in counting process input format for modeling
    time-to-optional-discontinuation if O>0.  For simplicity, time is
```

```
    rounded up to integers and we split time into intervals of length 1

    Lambda2 is the indicator for observing optional discontinuation in
    counting process input format.

    Lambda is as defined in the implementation section.

    Vh is a time-dependent covariate. */

data modelstop; set dataset;
   if H=0; keep H i d1 d2 S Lambda Lambda2 X1 X2 Vh D Z;
   do d2=1 to S;
      d1=d2-1;
      Vh=0;
      if d1>=D then Vh=1;
      Lambda2=0;
      if d2=S then Lambda2=Lambda;
      output;
   end;
run;

/*  Model time-to-optional-discontinuation for creating weight -
    separately build model for each treatment group  */

proc phreg data=modelstop; where Z=0;
   id i;
   model (d1,d2)*Lambda2(0)= X1  X2 Vh;
   output out=stopprob0(keep=stopprob i d1 d2) survival=stopprob;
run;

proc phreg data=modelstop; where Z=1;
   id i;
   model (d1,d2)*Lambda2(0)= X1 X2 Vh;
   output out=stopprob1(keep=stopprob i d1 d2) survival=stopprob;
run;

data stopprob; set stopprob0 stopprob1;
run;

proc sort data=stopprob; by i d1 d2;
run;

/*  Create data in counting process input format for modeling the event time.
    Delta2 is the indicator for observing event in counting process input
    format.

    Delta is as defined in the main paper */

data modeldth; set dataset;
   if H=0; keep  pH numpH H i d1 d2 S Delta2 Delta U Z Lambda ;
   do d2=1 to U;
       d1=d2-1;
       Delta2=0;
       if d2=U then Delta2=Delta;
       output;
```

```
   end;
run;

/* Calculate the Kaplan-Meier estimator at each time point */

proc sort data=dataset out=dataset2; by descending U; run;
data dataset2; set dataset2; if _N_=1; keep U; run;

data modeldth2; set dataset2; keep U d2;
   do d2=1 to U;
      output;
   end;
run;

data modeldth2; merge modeldth2 numstopprob0 numstopprob1; by d2; run;

data modeldth2; set modeldth2; retain numstopprob0a numstopprob1a;
   if numstopprob0^=. then numstopprob0a=numstopprob0;
   else  numstopprob0a=numstopprob0a;
   if numstopprob1^=. then numstopprob1a=numstopprob1;
  else  numstopprob1a=numstopprob1a;
   if d2=0 then delete;
run;

/* As the denominator of the weight, K{uVS, Q (uVS)}, for a subject will
   be constant in time if u>S, we need to find K{S, Q(S)} */

data laststopprob; set stopprob; by i d1 d2; if last.i;
   stopprob2=stopprob; drop stopprob;
run;

data modeldth; merge modeldth laststopprob(keep=i stopprob2); by i;run;

data modeldth; merge  stopprob  modeldth; by i d1 d2;
   if stopprob=. and d2>S then stopprob=stopprob2;
run;

proc sort data=modeldth; by d2; run;

data modeldth; merge modeldth modeldth2(keep=d2 numstopprob0a numstopprob1a);
by d2;run;

proc sort data=modeldth; by i d1 d2;run;

/* Construct the stabilized weight */

data modeldth; set modeldth;
   invwt=1/(stopprob*(1-pH));
   if Z=0 then
   stbinvwt=numstopprob0a*(1-numpH)*invwt;
   else stbinvwt=numstopprob1a*(1-numpH)*invwt;
   if Lambda=1 and d2>S then do; invwt=0;  stbinvwt=0; end;
run;

/*  Solve (4.4) and get sandwich variance  */
```

```
proc phreg data=modeldth COVSANDWICH(aggregate);
   id i;
   model (d1,d2)*Delta2(0)=Z;
   weight stbinvwt;
run;
```

Additional references

Cole, S.R. and Herán, M.A. (2003). Adjusted survival curves with inverse probability weights. *Computer Methods and Programs in Biomedicine* **75**, 45–49.

Kalbfleisch, J.D. and Prentice, R.L. (2002). *The Statistical Analysis of Failure Time Data*. Hoboken, New Jersey: John Wiley and Sons.