Journal of Statistical Software

May 2015, Volume 65, Issue 7. *http://www.jstatsoft.org/*

DTR: An R Package for Estimation and Comparison of Survival Outcomes of Dynamic Treatment Regimes

Xinyu Tang
University of Arkansas
for Medical Sciences

Maria Melguizo
University of Arkansas
for Medical Sciences

Abstract

Sequentially randomized designs, more recently known as sequential multiple assignment randomized trial (SMART) designs, are widely used in biomedical research, particularly in clinical trials, to assess and compare the effects of various treatment sequences. In such designs, patients are initially randomized to one of the first-stage therapies. Then patients meeting some criteria (e.g., no relapse of disease) participate in the second-stage randomization to one of the second-stage therapies. The advantage of such a design is that it allows the investigator to study various treatment sequences where the patients' second-stage therapies can be adjusted based on their responses to the first-stage therapies. In the past few years, substantial improvement has been made in the statistical methods for analyzing the data from SMARTs. Much of the proposed statistical approaches focus on estimating and comparing the survival outcomes of treatment sequences embedded in the SMART designs. In this article, we introduce the R package **DTR**, which provides a set of functions that can be used to estimate and compare the effects of different treatment sequences on survival outcomes using the newly proposed statistical approaches. The proposed package is also illustrated using simulated data from SMARTs.

Keywords: adaptive treatment strategy, dynamic treatment regime, inverse-probability weighting, survival analysis, sequential multiple assignment randomized trial (SMART) design, sequentially randomized design, treatment sequence.

1. Introduction

Sequentially randomized designs, more recently known as sequential multiple assignment randomized trial (SMART) designs, are widely used in biomedical research, particularly in clinical

trials, to assess and compare the effects of various treatment sequences (Robins 1986, 1987; Lavori and Dawson 2000; Murphy 2005; Bembom and van der Laan 2007; Chakraborty and Murphy 2013; Kidwell 2014). In such designs, patients are initially randomized to one of the first-stage therapies. Then patients meeting some criteria (e.g., no relapse of disease) participate in the second-stage randomization to one of the second-stage therapies. The second-stage therapy could be a rescue therapy if patients show primary resistance to the first-stage therapy, or a maintenance therapy if favorable responses are observed. There are different types of SMART designs with respect to the participants in the second-stage randomization: (i) SMART designs in which only the responders to one of the first-stage therapies participate in the second-stage randomization to one of the maintenance therapies; (ii) SMART designs in which only the non-responders to one of the first-stage therapies participate in the second-stage randomization to one of the rescue therapies; (iii) SMART designs in which the responders to all the first-stage therapies participate in the second-stage randomization; (iv) SMART designs in which the non-responders to all the first-stage therapies participate in the second-stage randomization; and (v) SMART designs in which all the responders and non-responders participate in the second-stage randomization (Lei, Nahum-Shani, Lynch, Oslin, and Murphy 2012; Chakraborty and Moodie 2013). The advantage of such a design is that it allows the investigator to study various treatment sequences where the patients' second-stage therapies can be adjusted based on their responses to the first-stage therapies. Here is an example of a SMART design investigating the effect of a combination of myeloablative chemotherapy, total-body irradiation and transplantation of autologous bone marrow purged of cancer cells (ABMT) to a standard chemotherapy in treating children with high risk neuroblastoma (Matthay *et al.* 1999, 2009). All the children with high risk neuroblastoma were treated with the same initial regimen of chemotherapy, and those without disease progression participated in the first-stage randomization to either ABMT or three cycles of intensive chemotherapy. All the children who completed cytotoxic therapy without disease progression then participated in the second-stage randomization to either treatment with 13-cis-retinoic acid (cis-RA) for six months or no further therapy.

In the past few years, substantial improvement has been made regarding the statistical methods for analyzing data from SMART designs in terms of either continuous outcomes (Dawson and Lavori 2010, 2012), binary outcomes (Buyze and Goetghebeur 2013) or survival outcomes (Lunceford, Davidian, and Tsiatis 2002; Guo and Tsiatis 2005). This article focuses on the proposed statistical approaches for estimating and comparing the survival endpoints of treatment sequences from these trials. Those treatment sequences are usually referred to as dynamic treatment regimes (DTRs), adaptive treatment strategies, or treatment policies. A DTR is defined as a sequence of the first-stage therapy, an intermediate response to the first-stage therapy, and the second-stage therapy. The construction of DTRs is usually from SMART designs or longitudinal observational studies. However, this article only discusses the DTRs constructed from SMART designs. In the high risk neuroblastoma study described above, four DTRs can be constructed: (i) treat with ABMT followed by cis-RA if no disease progression; (ii) treat with ABMT followed by no further therapy if no disease progression; (iii) treat with chemotherapy followed by cis-RA if no disease progression; and (iv) treat with chemotherapy followed by no further therapy if no disease progression. Evaluating the effect of a sequence of therapies is more efficient and more clinically meaningful than looking at the first- and second-stage therapies separately (Chakraborty and Murphy 2013; Kidwell 2014). Lunceford *et al.* (2002) introduced the survival and mean restricted survival estimators for

treatment regimes in a two-stage randomization design using two forms of inverse-probability weighting for second-stage randomization and censoring respectively. Guo and Tsiatis (2005) proposed a weighted risk set estimator (WRSE) for the cumulative hazard function. Tang and Wahed (2011) proposed a generalized Cox model for comparing any combination of treatment regimes after adjustment for auxiliary variables. Kidwell and Wahed (2013) proposed weighted log-rank test statistics to compare survival distributions of DTRs. Tang and Wahed (2015) introduced the cumulative hazard ratio (CHR) estimator between two DTRs, and a testing procedure to compare the effects of DTRs based on the CHR estimator.

Although we have seen considerable advancement in the development of statistical methods for estimating and comparing the effects of DTRs on survival outcomes, no statistical package has been developed to implement those newly-proposed statistical approaches. Investigators and statisticians hesitate to implement the newly-proposed statistical approaches, mainly because (i) the statistical formulae for estimating and comparing the survival distributions of DTRs are theoretically complex, and (ii) the newly-proposed statistical methodology has not been included in any widely-used statistical software. Therefore, we developed the R package **DTR** (Tang and Melguizo 2015) to implement the newly-proposed statistical approaches for estimating and comparing the survival outcomes of different DTRs. The current version of the package and documentation implemented the SMART designs in which the responders to all the first-stage therapies participate in the second-stage randomization. However, it can also be applied to SMART designs where the non-responders to all the first-stage therapies participate in the second-stage randomization by switching the responders with non-responders. In Section 2 we briefly introduce the notation and the formulae for calculating the survival quantities based on each statistical method included in this version of the package. In Section 3 the functionality of each function is described and illustrated using simulated data from SMARTs. A more complete overview of the functionality is given in the reference manual of the **DTR** package which is available from the Comprehensive R Archive Network (CRAN) at http://CRAN.R-project.org/package=DTR. Some concluding remarks are provided in Section 4.

2. Statistical methods for dynamic treatment regimes

2.1. Notation

In the SMART designs used for the **DTR** package, patients are initially randomized to J first-stage therapies (A_1, A_2, ..., A_J), then patients who responded to the first-stage therapies are then randomized to K second-stage therapies (B_1, B_2, ..., B_K). The treatment regime A_jB_k, where $j = 1, 2, \ldots, J; k = 1, 2, \ldots, K$, is defined as "treat with A_j, followed by B_k if responds to A_j." The primary research interest is to estimate and compare the effects of different DTRs in terms of survival quantities (e.g., survival function, cumulative hazard function). The set of observed data from this design can be described as

$$\{X_{ji}, R_i, R_i T_i^R, R_i Z_{ki}, U_i, \Delta_i, V_i\},$$

for $j = 1, 2, \ldots, J; k = 1, 2, \ldots, K$; and $i = 1, 2, \ldots, n$; where

X_{ji}: the indicator for the jth first-stage therapy, $X_{ji} = 1$ if the ith patient was assigned to first-stage therapy A_j, and $X_{ji} = 0$ if otherwise;

R_i: the indicator for response, $R_i = 1$ if the ith patient responded, and $R_i = 0$ if otherwise;

T_i^R (optional): the time to response for the ith patient;

Z_{ki}: the indicator for kth second-stage therapy, $Z_{ki} = 1$ if the ith patient was assigned to second-stage therapy B_k, and $R_i = 1$ and $Z_{ki} = 0$ if otherwise; (note that $R_i = 0$ and $Z_{ki} = 0$ if the ith patient did not respond, and thus did not participate in the second-stage randomization);

U_i: the observed event/censoring time for the ith patient;

Δ_i: the censoring indicator, $\Delta_i = 0$ if the ith patient was censored, and $\Delta_i = 1$ if the ith patient's event was observed; and

V_i (optional): the covariate vector for the ith patient.

When the response status is considered as a time-varying process, we define the time-varying response variable to be $R_i(t) = R_i I(T_i^R < t)$.

Based on the analytical framework of inverse-probability weighting (Lunceford *et al.* 2002; Guo and Tsiatis 2005; Wahed and Tsiatis 2004, 2006; Tang and Wahed 2015; Kidwell and Wahed 2013), the weight functions used by various statistical approaches for treatment regime A_jB_k, $j = 1, 2, \ldots, J$ and $k = 1, 2, \ldots, K$ are defined in Table 1. Because patients randomized to A_1, A_2, ..., and A_J are independent samples, some statistical methods (Lunceford *et al.* 2002; Guo and Tsiatis 2005) estimate the survival quantities for treatment regime A_jB_k based on A_j data only. Based on the counting process notation described in Fleming and Harrington (1991), the risk, event and censoring indicators used in this article are outlined in Table 2. The survival function, cumulative hazard function, and hazard function for treatment regime A_jB_k are denoted as $S_{jk}(t)$, $\Lambda_{jk}(t)$ and $\lambda_{jk}(t)$ respectively for $j = 1, 2, \ldots, J$ and $k = 1, 2, \ldots, K$.

2.2. Inverse-probability weighting estimator

The survival estimator for treatment regime A_jB_k using the two forms of inverse-probability weighting (Lunceford *et al.* 2002) is

$$\hat{S}_{jk}(t) = 1 - \left\{ \sum_{i=1}^{n} \frac{\Delta_i W'_{jki}}{\hat{K}(U_i)} \right\}^{-1} \sum_{i=1}^{n} \frac{\Delta_i W'_{jki}}{\hat{K}(U_i)} I(U_i \leq t), \tag{1}$$

$j = 1, 2, \ldots, J$ and $k = 1, 2, \ldots, K$ where $\hat{K}(t)$ is the Kaplan-Meier estimate of the censoring curve, and its variance estimator is written as

$$\begin{aligned} \widehat{\mathsf{VAR}}\{\hat{S}_{jk}(t)\} \quad = \quad & \frac{1}{n} \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{\Delta_i W'^2_{jki}}{\hat{K}(U_i)} \left\{ I(U_i \leq t) - 1 + \hat{S}_{jk}(t) \right\}^2 \right. \\ & \left. + \int_0^L \frac{dN^c(u)}{\hat{K}(u)Y(u)} \hat{\mathsf{E}} \left\{ L_{jki}(t,u) \right\}^2 \right\}, \end{aligned} \tag{2}$$

where L is defined as the restricted lifetime, which is smaller than the maximum follow-up time of the SMART, and

$$\hat{\mathsf{E}} \left\{ L_{jki}(t,u) \right\}^2 = \frac{1}{n} \sum_{i=1}^{n} \frac{\Delta_i}{\hat{K}(U_i)} \left[W'_{jki} \left\{ I(U_i \leq t) - 1 + \hat{S}_{jk}(t) \right\} - \hat{G}_{jk}(t,u) \right]^2 I(U_i \geq u),$$

Definition	Estimation based on A_j data?	Time-varying response?	Usage
$W'_{jki} = (1 - R_i) + R_i Z_{ki}/\pi_{jk}$	Yes	No	Inverse-probability weighting estimator
$W'_{jki}(t) = [1 - R_i(t)] + R_i(t) Z_{ki}/\pi_{jk}$	Yes	Yes	Weighted risk set estimator
$W_{jki} = X_{ji}\{(1 - R_i) + R_i Z_{ki}/\pi_{jk}\}/\pi_j$	No	No	Cumulative hazard ratio estimator (see $N_{jki}(t)$ and $Y_{jki}(t)$ in Table 2)
$W_{jki}(t) = X_{ji}\{[1 - R_i(t)] + R_i(t) Z_{ki}/\pi_{jk}\}/\pi_j$	No	Yes	Weighted log-rank test (see $N^*_{jki}(t)$ and $Y^*_{jki}(t)$ in Table 2)

Table 1: Various weight functions for treatment regime $A_j B_k$ for $j = 1, 2, \ldots, J$ and $k = 1, 2, \ldots, K$, where $\pi_j = \mathsf{P}(X_{ji} = 1)$ and $\pi_{jk} = \mathsf{P}(Z_{ki} = 1 | X_{ji} = 1, R_i = 1)$.

At-risk process at time t	
Patient i	Number of patients
$Y_i(t) = I(U_i \geq t)$	$Y(t) = \sum_{i=1}^{n} Y_i(t)$
$Y_{ji}(t) = I(U_i \geq t, X_{ji} = 1)$	$Y_j(t) = \sum_{i=1}^{n} Y_{ji}(t)$
$Y_{ji}^{NR}(t) = \{1 - R_i(t)\}Y_{ji}(t)$	$Y_j^{NR}(t) = \sum_{i=1}^{n}\{1 - R_i(t)\}Y_{ji}(t)$
$Y_{jki}(t) = W_{jki}Y_i(t)$	$Y_{jk}(t) = \sum_{i=1}^{n} W_{jki}Y_i(t)$
$Y_{jki}^{*}(t) = W_{jki}(t)Y_i(t)$	$Y_{jk}^{*}(t) = \sum_{i=1}^{n} W_{jki}(t)Y_i(t)$
Event process at time t	
Patient i	Number of patients
$N_i(t) = I(U_i \geq t, \Delta_i = 1)$	$N(t) = \sum_{i=1}^{n} N_i(t)$
$N_{ji}(t) = I(U_i \geq t, \Delta_i = 1, X_{ji} = 1)$	$N_j(t) = \sum_{i=1}^{n} N_{ji}(t)$
$N_{ji}^{NR}(t) = \{1 - R_i(t)\}N_{ji}(t)$	$N_j^{NR}(t) = \sum_{i=1}^{n}\{1 - R_i(t)\}N_{ji}(t)$
$N_{jki}(t) = W_{jki}N_i(t)$	$N_{jk}(t) = \sum_{i=1}^{n} W_{jki}N_i(t)$
$N_{jki}^{*}(t) = W_{jki}(t)N_i(t)$	$N_{jk}^{*}(t) = \sum_{i=1}^{n} W_{jki}(t)N_i(t)$
Censoring process at time t	
Patient i	Number of patients
$N_i^c(t) = I(U_i \geq t, \Delta_i = 0)$	$N^c(t) = \sum_{i=1}^{n} N_i^c(t)$

Table 2: Counting process notations.

where

$$\hat{G}_{jk}(t,u) = \{n\hat{S}(u)\}^{-1}\sum_{i=1}^{n}\frac{\Delta_i W'_{jki}}{\hat{K}(U_i)}\left\{I(U_i \leq t) - 1 + \hat{S}_{jk}(t)\right\}I(U_i \geq u).$$

The detailed steps for calculating the variance estimator are outlined in Appendix A. The inverse-probability weighting estimator is mostly used when the primary research question of a SMART is to estimate the survival probabilities of DTRs over time, and the time to response is not observed or it is known that the time to response does not affect the overall survival outcome.

2.3. Weighted risk set estimator

The weighted risk set estimator (WRSE) of the survival function for treatment regime A_jB_k (Guo and Tsiatis 2005) is

$$\hat{S}_{jk}(t) = \exp\left\{-\sum_{i=1}^{n}\int_0^t \frac{W'_{jki}(u)dN_i(u)}{\sum_{p=1}^{n} W'_{jkp}(u)Y_p(u)}\right\}, \tag{3}$$

and its variance estimator is written as

$$\widehat{\mathrm{VAR}}\{\hat{S}_{jk}(t)\} = \frac{1}{n}\left\{\hat{S}_{jk}(t)\right\}^2\hat{\sigma}^2,$$

where

$$\hat{\sigma}^2 = n \sum_{i=1}^{n} \left(\int_0^t \frac{W'_{jki}(u) dN_i(u)}{\sum_{p=1}^{n} W'_{jkp}(u) Y_p(u)} - \int_0^t \frac{W'_{jki}(u) Y_i(u) \sum_{p=1}^{n} W'_{jkp}(u) dN_p(u)}{\left[\sum_{p=1}^{n} W'_{jkp}(u) Y_p(u) \right]^2} \right)^2. \tag{4}$$

The detailed steps for calculating $\hat{\sigma}^2$ are outlined in Appendix B. The weighted risk set estimator is mostly used when the primary research question of a SMART is to estimate the survival probabilities of DTRs over time, and it is suspected that the time to response would impact the overall survival outcome.

2.4. Cumulative hazard ratio estimator

Based on stratified proportional hazards with treatment regimes as strata (Tang and Wahed 2015), the hazard function for treatment regime $A_j B_k$, $j = 1, \ldots, J$ and $k = 1, \ldots, K$ can be written as

$$\lambda_{jk}(t) = \lambda_{jk0}(t) \exp\left\{\beta^\top V\right\}, j = 1, \ldots, J \text{ and } k = 1, \ldots, K,$$

where $\lambda_{jk0}(t)$ is the baseline hazard function for treatment regime $A_j B_k$, and β is a vector of coefficients corresponding to baseline covariates V. The coefficient estimate $\hat{\beta}$ can be obtained by solving the pseudo-score equation

$$U(\beta) = \sum_{i=1}^{n} \sum_{j=1}^{J} \sum_{k=1}^{K} \int_0^L \left\{ V_i - \overline{V}_{jk}(t, \beta) \right\} dN_{jki}(t) = 0,$$

where

$$\overline{V}_{jk}(t, \beta) = \frac{\sum_{p=1}^{n} V_p Y_{jkp}(t) \exp\left\{\beta^\top V_p\right\}}{\sum_{p=1}^{n} Y_{jkp}(t) \exp\left\{\beta^\top V_p\right\}}.$$

The cumulative baseline hazard for treatment regime $A_j B_k$ can be obtained as

$$\hat{\Lambda}_{jk0}(t, \hat{\beta}) = \sum_{i=1}^{n} \int_0^t \frac{dN_{jki}(s)}{\sum_{p=1}^{n} Y_{jkp}(s) \exp\{\hat{\beta}^\top V_p\}}.$$

Based on the estimated cumulative baseline hazards, the estimated cumulative hazard ratio (CHR) for comparing treatment regimes $A_j B_k$ to $A_{j'} B_{k'}$ can be obtained as

$$\hat{\theta}_{jkj'k'}(t) = \frac{\hat{\Lambda}_{jk0}(t)}{\hat{\Lambda}_{j'k'0}(t)}. \tag{5}$$

The cumulative hazard ratio estimator $\hat{\theta}_{jkj'k'}(t)$ follows a Gaussian process with mean $\theta_{jkj'k'}(t)$ and variance function $\sigma^2_{jkj'k'}(t)$. The detailed steps for computing the variance estimator $\hat{\sigma}^2_{jkj'k'}(t)$ for $\hat{\theta}_{jkj'k'}(t)$ are described in Section 3 and Appendix B of Tang and Wahed (2015). The cumulative hazard ratio estimator is mostly used when the primary research interest is to

compare the survival curves of different DTRs after adjusting for other potential risk factors, and it is suspected that the proportional hazards assumption is violated.

2.5. Wald-type test

Wald-type tests can be used for comparing the survival quantities of different treatment regimes based on either the inverse-probability weighting estimator (Lunceford *et al.* 2002), WRSE (Guo and Tsiatis 2005) or the logarithm of the CHR estimator (Tang and Wahed 2015). For example, let us denote the estimates for the survival functions to be

$$\hat{S}(t) = \{\hat{S}_{11}(t), \hat{S}_{12}(t), \ldots, \hat{S}_{J(K-1)}(t), \hat{S}_{JK}(t)\},$$

and their estimated variance covariance matrix to be $\hat{\Sigma}$. For testing the null hypothesis $H_0 : DS(t) = 0$, where D is a $d \times (JK)$ matrix, the test statistic can be written as $\{D\hat{S}(t)\}^\top \{D\hat{\Sigma}D^\top\}^{-1}\{D\hat{S}(t)\}$, and it follows a chi-square distribution with d degrees of freedom. The Wald-type test is mostly used when the primary research interest is to compare the survival rates of different DTRs at certain times (e.g., 3-year survival, and 5-year survival).

2.6. Weighted log-rank test

Feng and Wahed (2008) introduced a supremum weighted log-rank test statistic for testing for the survival differences between two DTRs. However, the supremum weighted log-rank tests can only be applied for comparing two separate-path DTRs. Kidwell and Wahed (2013) extended Feng and Wahed (2008)'s work and proposed the weighted log-rank tests for comparing the survival distributions of shared-path DTRs. The current version of the package incorporated the weighted log-rank tests proposed by Kidwell and Wahed (2013). The standardized weighted log-rank test statistic for comparing treatment regime A_jB_k and $A_jB_{k'}$ can be written as

$$T^W_{jkjk'}(t) = n^{-1/2} Z^W_{jkjk'}(t)/\hat{\sigma}(t), \tag{6}$$

where

$$Z^W_{jkjk'}(t) = \int_0^t \frac{Y^*_{jk}(s)Y^*_{jk'}(s)}{Y^*_{jk}(s) + Y^*_{jk'}(s)} \left\{ \frac{dN^*_{jk}(s)}{Y^*_{jk}(s)} - \frac{dN^*_{jk'}(s)}{Y^*_{jk'}(s)} \right\}, \tag{7}$$

and

$$\begin{aligned} \hat{\sigma}^2(t) \quad = \quad & \frac{1}{n}\int_0^t \frac{{Y^*_{jk'}}^2(s)\sum_{p=1}^n W^2_{jkp}(s)Y_{jp}(s) + {Y^*_{jk}}^2(s)\sum_{p=1}^n W^2_{jk'p}(s)Y_{jp}(s)}{\left\{Y^*_{jk}(s) + Y^*_{jk'}(s)\right\}^2} \left\{\frac{dN_j(s)}{Y_j(s)}\right\} \\ & -\frac{2}{n}\int_0^t \frac{Y^*_{jk}(s)Y^*_{jk'}(s)}{\left\{Y^*_{jk}(s) + Y^*_{jk'}(s)\right\}^2} \left\{\pi_j^{-2} Y_j^{NR}(s)\frac{dN_j(s)}{Y_j(s)}\right\}. \end{aligned} \tag{8}$$

The detailed steps for calculating $Z^W_{jkjk'}(t)$ and $\hat{\sigma}^2(t)$ are outlined in Appendix C. The weighted log-rank test is mostly used when the primary research interest is to compare the survival curves of different DTRs, and there is no other risk factors to be considered.

2.7. Generalized Cox model

For comparing the effects of different DTRs after adjustment for auxiliary variables, Tang and Wahed (2011) proposed to use the following version of the Cox model based on two-stage randomization designs where $J = K = 2$:

$$\begin{aligned}\lambda(t) &= \lambda_0(t)\exp\Big\{\beta_1^{(1)}X_1 + \beta^{(2)}R(t) + \beta_1^{(3)}X_1R(t) + \beta_1^{(4)}Z_1R(t) + \beta_{11}^{(5)}X_1Z_1R(t) \\ &\quad +\gamma^\top V\Big\}, \end{aligned} \tag{9}$$

where $\lambda(t)$ is the general hazard function at time t; $\lambda_0(t)$ denotes the baseline hazard function (when all the covariates are equal to 0); $R(t)$ denotes the time-varying measurement of response and consent as defined before; and β is the vector of coefficients denoted as $\beta = \left[\beta_1^{(1)}, \beta^{(2)}, \beta_1^{(3)}, \beta_1^{(4)}, \beta_{11}^{(5)}, \gamma\right]^\top$. Under model (9), comparisons of treatment regimes in terms of their hazards can be interpreted based on the coefficient vector β. For example, for comparing all four treatment regimes in a simple two-stage randomization design, the null hypothesis $H_0 : \lambda_{A_1B_1}(t) = \lambda_{A_1B_2}(t) = \lambda_{A_2B_1}(t) = \lambda_{A_2B_2}(t)$ can be interpreted as $H_0 : \beta_1^{(1)} = \beta_1^{(3)} = \beta_1^{(4)} = \beta_{11}^{(5)} = 0$. The generalized Cox model is mostly used when the primary research interest is to compare the survival curves of different DTRs after adjusting for other potential risk factors under the assumption of proportional hazards across regimes.

3. The DTR package

The **DTR** package for the R environment for statistical computing and graphics (R Core Team 2015) can be downloaded from CRAN at `http://CRAN.R-project.org/package=DTR`. This package is designed for the estimation and comparison of survival distributions of DTRs from SMARTs in which the responders to all the first-stage therapies participate in the second-stage randomization. In SMART designs, there could be more than two therapies available at each stage. For simplicity, and to maintain the similarity to the most common SMARTs, a two-stage randomization design allowing for two treatment options at each stage is used in the current version of the package. The basic structure of a SMART design used for this package is depicted in Figure 1. In detail, patients are initially randomized to either A_1 or A_2 as first-stage therapy, then patients who responded to the first-stage therapies are randomized to either B_1 or B_2 as second-stage therapy. Four DTRs are embedded in this design:

A_1B_1: treat with A_1, followed by B_1 if responds to A_1;

A_1B_2: treat with A_1, followed by B_2 if responds to A_1;

A_2B_1: treat with A_2, followed by B_1 if responds to A_2; and

A_2B_2: treat with A_2, followed by B_2 if responds to A_2.

The current version of the **DTR** package includes the functions for simulating data from various two-stage randomization designs, estimating the survival quantities (e.g., survival function, restricted mean survival, cumulative hazard ratio) of treatment regimes, plotting the survival estimates over time, and comparing the survival distributions of different regimes. The statistical approaches proposed in Lunceford *et al.* (2002); Guo and Tsiatis (2005); Tang

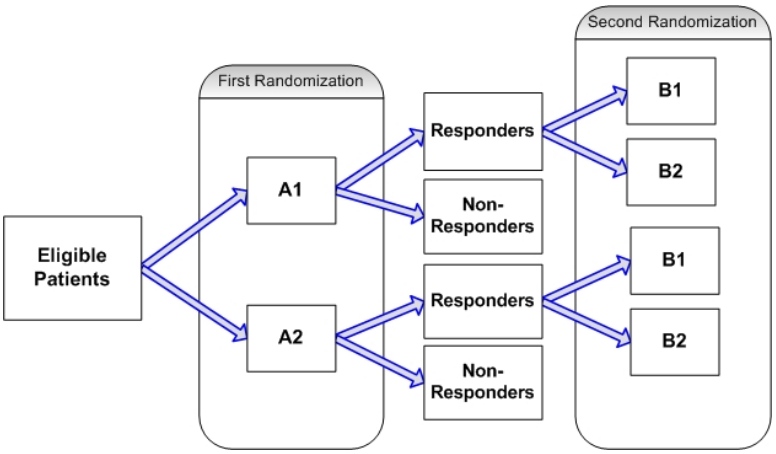

Figure 1: The basic structure of a SMART design used for the **DTR** package.

and Wahed (2011, 2015); and Kidwell and Wahed (2013) are implemented in the current version of the **DTR** package. A brief introduction of each statistical method is given in Sections 2.2–2.7. For the convenience of the reader, we also summarize the functionality of the **DTR** package in Table 3.

3.1. Output/input data

The data from a two-stage randomization design shown in Figure 1 is stored in an R data frame consisting of the following arguments:

`X`: First-stage indicator, `X` $= 0$ if assigned to A_1, and `X` $= 1$ if assigned to A_2.

`TR`: Time to response.

`R`: Response status, `R` $= 1$ for responders, and `R` $= 0$ for non-responders.

`Z`: Second-stage indicator, `R` $= 1$ and `Z` $= 0$ if assigned to B_1 (note that `R` $= 0$ and `Z` $= 0$ for non-responders), and `Z` $= 1$ if assigned to B_2.

`U`: Observed time, `U` is event time if `delta` $= 1$, and `U` is censoring time if `delta` $= 0$.

`delta`: Censoring indicator, `delta` $= 1$ for observed events, and `delta` $= 0$ for censored.

`V`: Covariate(s) to be adjusted. `V` may include one column for one covariate, or more than one column for multiple covariates.

Because (i) only the data from the A_j arm was described in the simulation studies of Lunceford *et al.* (2002) and Guo and Tsiatis (2005), (ii) the approaches described in Guo and Tsiatis (2005) and Kidwell and Wahed (2013) treated response status as a time-varying process, (iii) Tang and Wahed (2011) and Tang and Wahed (2015) took into account the information from auxiliary covariates, different sets of the above arguments are generated by different data simulation functions, and required as the input data for different estimation and comparison functions. For clarification, we summarize the output/input data for each function in Table 4.

Function	Arguments	Description
Data simulation functions		
`simLDTdata`	`(n, ...)`	Data simulation
`simWRSEdata`	`(n, ...)`	Data simulation
`simCHRdata`	`(n, ...)`	Data simulation
`simLRdata`	`(n, ...)`	Data simulation
`simPHdata`	`(n, ...)`	Data simulation
Estimation functions		
`LDTestimate`	`(data, L)`	Survival estimates
`WRSEestimate`	`(data)`	Survival estimates
`CHRestimate`	`(data, covar)`	Cumulative hazard ratio estimates
S3 `summary` methods		
for class '`DTR`'	`(object, ...)`	Summarize '`DTR`' object
for class '`CHR`'	`(object, ...)`	Summarize '`CHR`' object
S3 `print` methods		
for class '`DTR`'	`(x, ...)`	Print '`DTR`' object
for class '`summary.DTR`'	`(x, ...)`	Print '`summary.DTR`' object
for class '`CHR`'	`(x, ...)`	Print '`CHR`' object
for class '`summary.CHR`'	`(x, ...)`	Print '`summary.CHR`' object
S3 `plot` methods		
for class '`DTR`'	`(x, ...)`	Survival plot
for class '`CHR`'	`(x, ...)`	Cumulative hazard ratio plot
Comparison functions		
`contrast_wald`	`(est, t)`	Wald-type test based on survival estimates
`contrast_chr`	`(est, t)`	Wald-type test based on cumulative hazard ratio estimates
`contrast_logrank`	`(data)`	Weighted log-rank test
`PHfit`	`(data, covar)`	Fit generalized Cox model
`contrast_ph`	`(fit)`	Wald-type test based on generalized Cox model

Table 3: Summary of the functionality of the **DTR** package.

3.2. Inverse-probability weighting estimator functions

The function `LDTestimate()` computes the inverse-probability weighting survival estimates and their estimated standard errors for DTRs at observed event times based on Equations 1 and 2. The technical details are briefly described in Section 2.2 (see more details in Lunceford *et al.* 2002). The examples in this section will be illustrated using a simulated dataset called `LDTdata` which is included in the **DTR** package. The `LDTdata` dataset was simulated as

Function	Input argument	Output/input data
`simLDTdata`	—	{`R`, `Z`, `U`, `delta`}
`simWRSEdata`	—	{`TR`, `R`, `Z`, `U`, `delta`}
`simCHRdata`	—	{`X`, `R`, `Z`, `U`, `delta`, `V`}
`simLRdata`	—	{`X`, `TR`, `R`, `Z`, `U`, `delta`}
`simPHdata`	—	{`X`, `TR`, `R`, `Z`, `U`, `delta`, `V`}
`LDTestimate`	`data`	{`X`, `R`, `Z`, `U`, `delta`}
`WRSEestimate`	`data`	{`X`, `TR`, `R`, `Z`, `U`, `delta`}
`CHRestimate`	`data`	{`X`, `R`, `Z`, `U`, `delta`, `V`}
`contrast_logrank`	`data`	{`X`, `TR`, `R`, `Z`, `U`, `delta`}
`PHfit`	`data`	{`X`, `TR`, `R`, `Z`, `U`, `delta`, `V`}

Table 4: Summary of output/input data for each function.

described in the simulation study of Lunceford *et al.* (2002) using the `simLDTdata()` function as follows.

After installing package **DTR**, load the package:

```
R> library("DTR")
```

Simulate the `LDTdata` dataset using the `simLDTdata()` function:

```
R> set.seed(123)
R> data.A1 <- simLDTdata(n = 100, max.c = 2.5, pi.r = 0.5, pi.z = 0.5,
+    lambda = 1.33, alpha = 6.67, beta1 = 0.29, beta2 = -0.67, L = 1.5)
R> data.A2 <- simLDTdata(n = 100, max.c = 2.5, pi.r = 0.5, pi.z = 0.5,
+    lambda = 1.33, alpha = 6.67, beta1 = 0.29, beta2 = -0.67, L = 1.5)
R> LDTdata <- cbind(X = c(rep(0, 100), rep(1, 100)),
+    rbind(data.A1, data.A2))
```

The same data set can also be loaded into the workspace using the command below:

```
R> data("LDTdata", package = "DTR")
```

The `LDTdata` dataset is a data frame with 200 rows corresponding to patients and 5 columns corresponding to variables. There were 200 patients equally assigned to either A_1 or A_2 at the first stage. Afterwards, 98 (49%) patients responded to either A_1 or A_2, and they were then equally assigned to either B_1 or B_2 at the second stage. The total follow-up time was 2.5 years, and the survival time was restricted to 1.5 years.

```
R> dim(LDTdata)
```

```
[1] 200   5
```

```
R> head(LDTdata)
```

```
  X R Z          U delta
1 0 1 0 0.71894380     0
2 0 0 0 0.54286779     1
3 0 0 0 1.02244230     0
4 0 1 1 0.04301965     1
5 0 0 0 0.04672240     1
6 0 1 1 0.11389125     0
```

The function `LDTestimate()` creates an object of class '`DTR`' that contains a list of components: DTRs, the number of observation for each regime, the number of events for each regime, observed event times, the number of patients at risk at each event time, the number of events at each event time, the survival estimates for each regime, and their variance/covariance estimates.

```
R> est <- LDTestimate(data = LDTdata)
R> est

Call: LDTestimate(data = LDTdata)

  DTR records events    median     LCL95     UCL95
 A1B1      75     54 0.8363004 0.5499846 1.0112419
 A1B2      75     51 0.5245500 0.4290100 0.7144589
 A2B1      77     58 0.5495142 0.4500212 0.8327760
 A2B2      75     55 0.5773552 0.3939629 1.1364390
```

The median survival times for treatment regimes A_1B_1, A_1B_2, A_2B_1, and A_2B_2 were estimated to be 0.84 (95% confidence interval (CI): 0.55–1.01), 0.52 (95% CI: 0.43–0.71), 0.55 (95% CI: 0.45–0.83), and 0.58 (95% CI: 0.39–1.14) years respectively. We can call the `summary()` function of the '`DTR`' object to print a full list of the survival estimates.

```
R> summary(est)

Call: LDTestimate(data = LDTdata)

        time n.risk n.event    SURV11    SURV12    SURV21
 0.003283737    199       1 0.9909211 0.9886175 1.0000000
 0.004230144    198       1 0.9818422 0.9772351 1.0000000
 0.004784251    197       1 0.9818422 0.9772351 0.9899313
 0.006796506    196       1 0.9818422 0.9772351 0.9798627
 0.017014431    195       1 0.9818422 0.9772351 0.9697940
 0.018175468    194       1 0.9727633 0.9658526 0.9697940
     SURV22       SE11       SE12        SE21        SE22
 1.0000000 0.01008932 0.01008263 0.000000000 0.000000000
 1.0000000 0.01425190 0.01423295 0.000000000 0.000000000
 0.9900736 0.01425190 0.01423295 0.009984863 0.009991568
 0.9801472 0.01425190 0.01423295 0.014099290 0.014118278
 0.9702207 0.01425190 0.01423295 0.017241736 0.017276660
 0.9702207 0.01743465 0.01739978 0.017241736 0.017276660
```

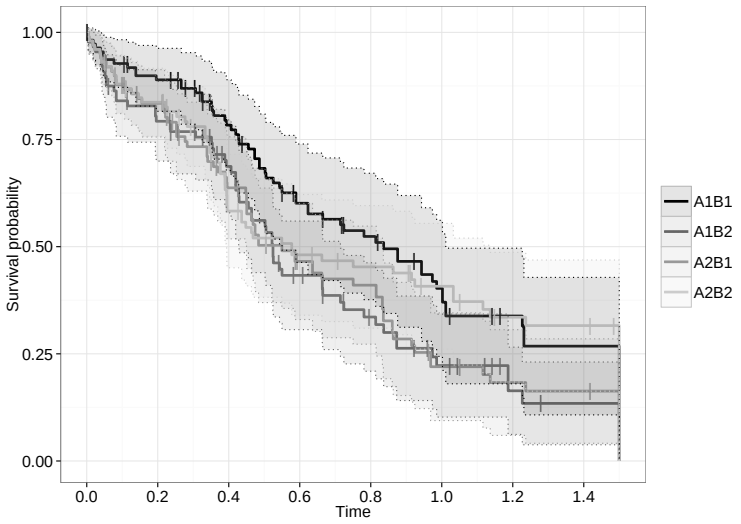

Figure 2: The inverse-probability weighting estimates under four treatment regimes based on the simulated data. The | represents censoring.

Only the first few rows of the `summary()` output are shown above. The `plot()` function can be used to plot the survival estimates and their corresponding 95% confidence bands of four DTRs, displayed in Figure 2. Comparisons of survival estimates at a specific time point can be carried out among DTRs by calling the function `contrast_wald()`. Both overall and pairwise comparisons are performed.

```
R> contrast_wald(est, t = 1)
```

```
              H0 (t=1) test statistic df          p
1 A1B1=A1B2=A2B1=A2B2   7.9059901218  3 0.04799509
2           A1B1=A1B2   3.7484056950  1 0.05285791
3           A1B1=A2B1   3.1334772144  1 0.07669998
4           A1B1=A2B2   0.0019567425  1 0.96471702
5           A1B2=A2B1   0.0008756533  1 0.97639289
6           A1B2=A2B2   3.6607866126  1 0.05570733
7           A2B1=A2B2   4.1219916252  1 0.04232922
```

There was a significant difference in the 1-year survival among four treatment regimes (overall $p = 0.048$), and patients following regime A_2B_2 appeared to have a better 1-year survival compared to patients following regime A_2B_1 (A_2B_1 vs. A_2B_2: $p = 0.04$; Figure 2).

3.3. Weighted risk set estimator functions

The function `WRSEestimate()` calculates the weighted risk set estimates of the survival functions and their estimated standard errors based on Equations 3 and 4. The technical details are briefly described in Section 2.3 (see more details in Guo and Tsiatis 2005). The examples in this section will be illustrated using a simulated dataset called `WRSEdata` which is included in the **DTR** package. The `WRSEdata` dataset was simulated as described in the simulation study of Guo and Tsiatis (2005) using the `simWRSEdata()` function.

Load the WRSEdata dataset into the workspace:

```
R> data("WRSEdata", package = "DTR")
```

The WRSEdata dataset is a data frame with 200 rows corresponding to patients and 6 columns corresponding to variables. In this dataset, 200 eligible patients were equally assigned to first-stage therapy A_1 or A_2. All the patients were followed from the time of randomization until they responded/did not respond to the first-stage therapy, and the time to response varied across patients. 98 (49%) of the patients responded to the first-stage therapy, and were further assigned to one of the second-stage therapies B_1 or B_2 at a $1:1$ ratio. The maximum follow-up time was 3.5 years.

```
R> dim(WRSEdata)

[1] 200   6

R> head(WRSEdata)

  X        TR R Z         U delta
1 0 415.28788 1 0 367.38028     0
2 0   0.00000 0 0 131.76759     1
3 0   0.00000 0 0 286.84001     1
4 0  30.20926 1 1 211.68924     1
5 0   0.00000 0 0  11.34070     1
6 0 599.66131 1 1  58.19843     0
```

The function WRSEestimate() creates an object of the same class 'DTR' as described above.

```
R> est <- WRSEestimate(data = WRSEdata)
R> est

Call: WRSEestimate(data = WRSEdata)

  DTR records events   median    LCL95    UCL95
 A1B1      75     60 261.4049 182.9489 386.9716
 A1B2      75     57 262.6695 202.9910 416.0386
 A2B1      77     60 209.5315 133.3808 312.1586
 A2B2      75     57 275.8422 160.7573 441.5636
```

The estimated median survival times were 261 (95% CI: 183–387), 263 (95% CI: 203–416), 210 (95% CI: 133–312), and 276 (95% CI: 161–442) days for regimes A_1B_1, A_1B_2, A_2B_1, and A_2B_2 respectively. The functions summary() and plot() can be used to display and plot the survival estimates at observed event times respectively. The plot of the survival estimates and their 95% confidence bands across study period is shown in Figure 3. Comparisons of survival estimates can be carried out between different DTRs by testing the equality of the weighted risk set estimates at a specific time point by calling the function contrast_wald().

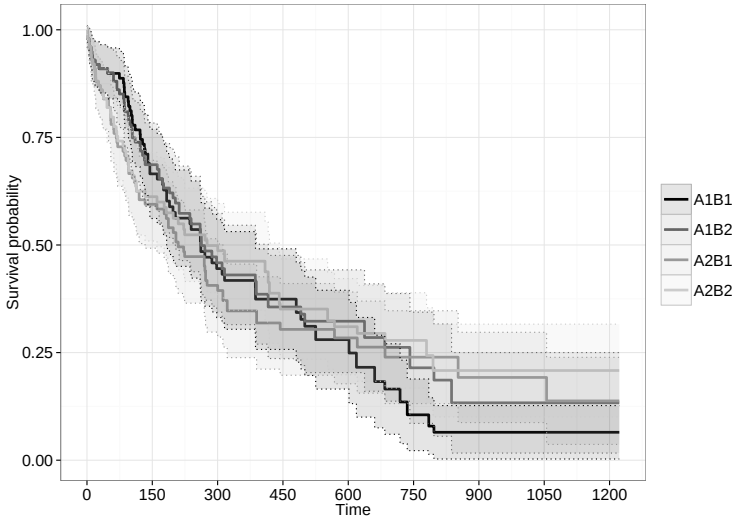

Figure 3: The weighted risk set estimates under four treatment regimes based on the simulated data.

```
R> contrast_wald(est, t = 500)
```

```
           H0 (t=500) test statistic df         p
1 A1B1=A1B2=A2B1=A2B2      0.78786426  3 0.8523675
2           A1B1=A1B2      0.04674220  1 0.8288323
3           A1B1=A2B1      0.08414960  1 0.7717508
4           A1B1=A2B2      0.08279477  1 0.7735452
5           A1B2=A2B1      0.18895723  1 0.6637859
6           A1B2=A2B2      0.01882311  1 0.8908748
7           A2B1=A2B2      0.72328550  1 0.3950683
```

No significant difference was found in the 500-day survival among the four DTRs (overall $p = 0.85$).

3.4. Cumulative hazard ratio estimator functions

Based on the statistical methods proposed in Tang and Wahed (2015), we developed the `CHRestimate()` function for calculating the CHR estimates between pairwise DTRs across times. A brief introduction of the statistical method can be found in Section 2.4. The examples in this section will be illustrated using a simulated dataset called `CHRdata` which is included in the **DTR** package. The `CHRdata` dataset was simulated as described in the simulation study of Tang and Wahed (2015) using the `simCHRdata()` function.

Load the `CHRdata` dataset into the workspace:

```
R> data("CHRdata", package = "DTR")
```

The `CHRdata` dataset is a data frame with 200 rows corresponding to patients and 7 columns corresponding to variables. In this dataset, a total of 200 eligible patients participated in the first-stage randomization, and were equally assigned to either A_1 or A_2 as first-stage

therapy. Among the 200 patients, 121 (60.5%) responded to the first-stage therapies they were assigned to, and then participated in the second-stage randomization. At the second-stage randomization 121 patients were equally assigned to either B_1 or B_2 as second-stage therapy. Approximately 25% of the patients were censored during the 5-year study period. The investigator suspected that gender (column `V1`, 1 for males and 0 for females) may have an impact on the survival outcomes, and thus would like to adjust for the gender effect when assessing the effects of the four DTRs in terms of survival.

```
R> dim(CHRdata)

[1] 200   7

R> head(CHRdata)

  X R Z          U delta V1 V2
1 0 1 0 4.07642397     1  0  0
2 1 1 1 1.98054567     1  1  1
3 0 1 0 2.31292696     1  0  1
4 0 1 1 3.94075459     0  1  1
5 0 1 1 3.48862215     0  1  0
6 1 0 0 0.08659921     1  0  1
```

The function `CHRestimate()` creates an object of class '`CHR`' that contains a list of components: coefficient estimate(s) for covariate(s), comparisons between DTRs, the 75th percentile of the observed times, observed event times, the number of patients at risk at each event time, the number of events at each event time, the CHR estimates, their variance/covariance estimates, the log CHR estimates, and their variance/covariance estimates.

```
R> est <- CHRestimate(data = CHRdata, covar = "V1")
R> est$coefficients

          [,1]
[1,] 0.3691321
```

The gender coefficient was estimated to be 0.37. The hazards of death among males were estimated to be 1.45 times the hazards among females.

```
R> est

Call: CHRestimate(data = CHRdata, covar = "V1")

    comparison time75P       CHR     LCL95     UCL95     LOGCHR
 A1B2 vs. A1B1    2.95 0.8675783 0.4858320 1.2493246 -0.1420495
 A2B1 vs. A1B1    2.95 1.9842268 1.0744243 2.8940293  0.6852293
 A2B2 vs. A1B1    2.95 0.5147304 0.2563395 0.7731213 -0.6641120
 A2B1 vs. A1B2    2.95 2.2870867 1.2401984 3.3339751  0.8272788
 A2B2 vs. A1B2    2.95 0.5932956 0.2956825 0.8909088 -0.5220625
```

```
A2B2 vs. A2B1    2.95 0.2594111 0.1430457 0.3757765 -1.3493413
  LOGLCL95    LOGUCL95
-0.5820631  0.29796412
 0.2267120  1.14374671
-1.1661046 -0.16211933
 0.3695400  1.28501764
-1.0236896 -0.02043535
-1.7979166 -0.90076600
```

For example, the estimated CHR for comparing the treatment regime A_2B_1 to regime A_1B_1 was 0.87 (95% CI: 0.49–1.25) at 2.95 years. The `summary()` function prints the CHR estimates (default `log.CHR = FALSE`) or the log CHR estimates (`log.CHR = TRUE`) at observed event times.

```
R> summary(est, log.CHR = TRUE)

Call: CHRestimate(data = CHRdata, covar = "V1")

       time n.risk n.event     CHR1211     CHR2111     CHR2211
 0.01370667    200       1          NA          NA          NA
 0.02271074    199       1          NA          NA          NA
 0.02364251    198       1          NA          NA          NA
 0.03483425    197       1          NA          NA          NA
 0.03876242    196       1  1.08209451  1.15707144  1.04708406
 0.06430915    195       1  0.66809770  0.45782539  0.34783801
      CHR2112      CHR2212    CHR2221    SE1211    SE2111
          NA           NA -0.1087876        NA        NA
          NA           NA -0.1094524        NA        NA
  0.05929750 -0.050154913 -0.1094524        NA        NA
  0.46987657  0.359889189 -0.1099874        NA        NA
  0.07497693 -0.035010453 -0.1099874 0.9617040 1.1516823
 -0.21027231 -0.320259695 -0.1099874 0.6396431 0.9123984
     SE2211    SE2112    SE2212    SE2221
         NA        NA        NA 0.1607660
         NA        NA        NA 0.1617506
         NA 1.2217712 1.2199936 0.1617506
         NA 1.1521291 1.1496529 0.1625413
 1.1535057 0.9463734 0.9448615 0.1625413
 0.9147200 0.8462893 0.8454683 0.1625413
```

Only the first few rows of the `summary()` output are shown above. The pairwise log CHR estimates between any two different DTRs and their 95% confidence bands can be plotted by calling the `plot()` function with `log.CHR = TRUE` and `confidence.interval = TRUE` (Figure 4). The comparisons of different treatment regimes are carried out by performing the Wald-type tests based on the log CHR estimates by calling the `contrast_chr()` function.

```
R> contrast_chr(est, t = 3)
```

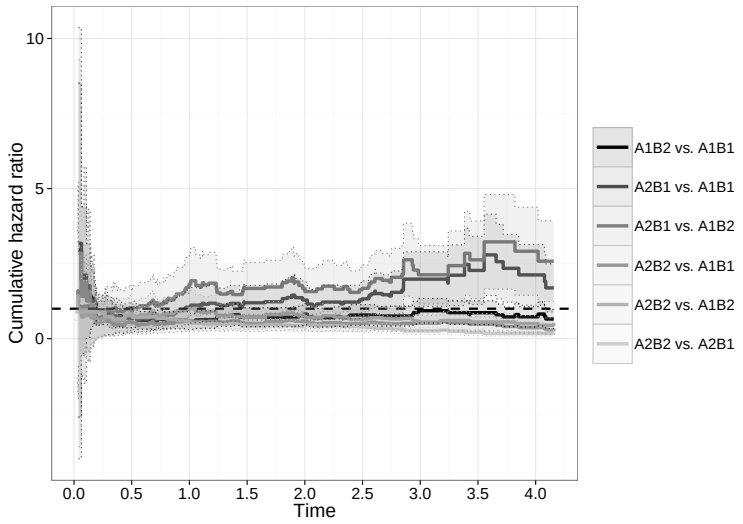

Figure 4: The cumulative hazard ratio estimates for comparing different treatment regimes based on the simulated data.

```
             H0 (t=3) test statistic df            p
1 A1B1=A1B2=A2B1=A2B2     35.9076643  3 7.832658e-08
2           A1B1=A1B2      0.1009254  1 7.507219e-01
3           A1B1=A2B1      8.5797065  1 3.399302e-03
4           A1B1=A2B2      6.7235720  1 9.514688e-03
5           A1B2=A2B1     10.6019870  1 1.129662e-03
6           A1B2=A2B2      5.4159672  1 1.995340e-02
7           A2B1=A2B2     34.7604241  1 3.728768e-09
```

As shown in Figure 4, there was no significant difference in the 3-year survival between treatment regimes A_1B_2 and A_1B_1 ($p = 0.75$). Other than that, all other treatment regimes were significantly different from each other at 3 years (overall $p < 0.001$).

3.5. The weighted log-rank test function

The `contrast_logrank()` function was developed with regard to the statistical methods proposed in Kidwell and Wahed (2013). We briefly introduced the methods in Section 2.6. The examples in this section will be illustrated using a simulated dataset called `LRdata` which is included in the **DTR** package. The `LRdata` dataset was simulated as described in the simulation study of Kidwell and Wahed (2013) using the `simLRdata()` function.

Load the `LRdata` dataset into the workspace:

```
R> data("LRdata", package = "DTR")
```

The `LRdata` dataset is a data frame with 100 rows corresponding to patients and 6 columns corresponding to variables. In this dataset, 100 eligible patients were equally assigned to one of the first-stage therapies A_1 or A_2. They were followed until they responded to or did not respond to the first-stage therapies. 61 (61%) responders were then equally assigned to one

of the second-stage therapies B_1 or B_2. The maximum follow-up time was 12 years, and 24% of the patients were lost to follow-up during the study.

```
R> dim(LRdata)

[1] 100   6

R> head(LRdata)

  X        TR R Z         U delta
1 1 0.3564477 1 1 3.3707436     1
2 0 0.0000000 0 0 1.9288219     1
3 0 0.0000000 0 0 0.2135722     1
4 1 1.2715099 1 0 7.1945839     1
5 0 2.6705747 1 1 2.7690037     1
6 1 0.0000000 0 0 0.2741043     1
```

The `contrast_logrank()` function compares the survival distributions of different DTRs using the weighted log-rank tests (Kidwell and Wahed 2013).

```
R> contrast_logrank(data = LRdata)

                   H0 (standardized) test statistic df      p
1 A1B1=A1B2=A2B1=A2B2                        1.2950  2 0.5234
2           A1B1=A1B2                       -0.5799  1 0.5620
3           A1B1=A2B1                       -0.3268  1 0.7438
4           A1B1=A2B2                        0.5964  1 0.5509
5           A1B2=A2B1                        0.3759  1 0.7070
6           A1B2=A2B2                        0.8272  1 0.4081
7           A2B1=A2B2                        1.1500  1 0.2501
```

We did not find any significant difference in the survival distributions of four DTRs (overall $p = 0.52$).

3.6. Generalized Cox model functions

The `PHfit()` and `contrast_ph()` functions were developed with regard to the statistical methods proposed in Tang and Wahed (2011). We briefly introduced the methods in Section 2.7. The examples in this section will be illustrated using a simulated dataset called `PHdata` which is included in the **DTR** package. The `PHdata` dataset was simulated as described in the simulation study of Tang and Wahed (2011) using the `simPHdata()` function.

Load the `PHdata` dataset into the workspace:

```
R> data("PHdata", package = "DTR")
```

The `PHdata` dataset is a data frame with 400 rows corresponding to patients and 7 columns corresponding to variables. In this dataset, 400 patients participated in the first-stage randomization to either A_1 or A_2, and then were followed until their response status was observed

at different times. Among 400 patients, 227 (57%) patients responded, and then participated in the second-stage randomization to either B_1 or B_2. The censoring rate was 53% throughout the study with a maximum follow-up time of 14 years.

```
R> dim(PHdata)

[1] 400   7

R> head(PHdata)

  X        TR R Z         U delta         V
1 1 0.0000000 0 0 0.1904523     1 0.7197622
2 1 0.0000000 0 0 1.8259748     1 0.8849113
3 0 0.5323748 1 1 1.6978088     0 1.7793542
4 1 0.0000000 0 0 0.6579115     0 1.0352542
5 0 0.0000000 0 0 2.0384931     1 1.0646439
6 1 0.0000000 0 0 4.6701414     1 1.8575325
```

The generalized Cox model can be fitted by calling the `PHfit()` function. The function returns an object of class '`coxph`' that is introduced in the **survival** package (Therneau and Grambsch 2000; Therneau 2015). The Cox proportional hazard model output can be printed by calling the `summary()` function of the '`coxph`' object.

```
R> fit <- PHfit(data = PHdata, covar = "V")
R> summary(fit)

Call:
"coxph(Surv(U, delta) ~ X + R + XR + RZ + XRZ + V)"

  n= 573, number of events= 187

       coef exp(coef) se(coef)      z Pr(>|z|)
X   -0.4613    0.6305   0.1869 -2.468  0.01357 *
R   -0.7918    0.4530   0.3266 -2.425  0.01532 *
XR   0.6769    1.9679   0.4263  1.588  0.11229
RZ   1.2743    3.5761   0.3557  3.583  0.00034 ***
XRZ -1.4178    0.2423   0.5105 -2.777  0.00548 **
V   -0.2520    0.7772   0.1539 -1.638  0.10152
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

    exp(coef) exp(-coef) lower .95 upper .95
X      0.6305     1.5862   0.43709    0.9094
R      0.4530     2.2074   0.23886    0.8592
XR     1.9679     0.5082   0.85335    4.5379
RZ     3.5761     0.2796   1.78102    7.1803
XRZ    0.2423     4.1279   0.08907    0.6589
```

```
V      0.7772     1.2866   0.57487    1.0509

Concordance= 0.601  (se = 0.023 )
Rsquare= 0.048   (max possible= 0.968 )
Likelihood ratio test= 28.2  on 6 df,   p=8.627e-05
Wald test            = 29.91  on 6 df,   p=4.082e-05
Score (logrank) test = 31.72  on 6 df,   p=1.845e-05
```

We can perform the comparisons of the survival distributions among different treatment regimes by calling the function `contrast_ph()` on the '`coxph`' object returned by the `PHfit()` function.

```
R> contrast_ph(fit)
```

```
                   H0 test statistic df            p
1 A1B1=A1B2=A2B1=A2B2     26.7316696  4 2.252281e-05
2               A1=A2     18.9904445  3 2.746455e-04
3               B1=B2     12.9969214  2 1.505755e-03
4           A1B1=A1B2     12.8367145  1 3.398839e-04
5           A1B1=A2B1      6.4112311  2 4.053394e-02
6           A1B1=A2B2      6.1272860  2 4.671719e-02
7           A1B2=A2B1     16.5567266  2 2.539525e-04
8           A1B2=A2B2     18.6510113  2 8.912188e-05
9           A2B1=A2B2      0.1527972  1 6.958765e-01
```

There was a significant difference in the survival distribution among the four DTRs (overall $p < 0.001$). Patients assigned to A_1 at the first stage had significantly different survival distribution compared to those assigned to A_2. Among patients who were assigned to A_1 at the first stage, whether they were assigned to B_1 or B_2 significantly affected their survival outcomes. However, if patients were assigned to A_2 at the first stage, there was no significant difference in the survival irrespective of their assignment at the second stage.

4. Discussion

In this article, we illustrated the **DTR** package designed for estimating and comparing the effects of treatment regimes from SMARTs in the context of survival endpoints. The package implements the statistical estimating and testing procedures proposed in Lunceford *et al.* (2002); Guo and Tsiatis (2005); Tang and Wahed (2011); Tang and Wahed (2015); and Kidwell and Wahed (2013). Each statistical method was briefly introduced, and the functionality of the **DTR** package was demonstrated using the simulated data from various SMARTs as described in aforementioned papers. The simulated datasets were generated for illustration purpose, and may not reflect the dataset in a real-life situation. For example, in the `WRSEdata` and `LRdata` datasets, some response times were observed after the censoring times, which may not be realistic in SMARTs. The package allows users to perform the generation of artificial data from SMARTs, point estimation of the survival quantities, estimates of their standard errors, overall and pairwise comparisons of the effects of different DTRs, and a visualization of the

survival curves for DTRs over the study period. We hope that this package can be useful for researchers in several areas where SMART designs with survival outcomes are applicable. Although it is still common to separately evaluate the first- and second-stage therapies when analyzing the survival data from SMARTs, we hope that the analyses could be substantially improved by using the functions provided by this package, as there has been a growing interest in studying the effects of different treatment sequences in biomedical research.

Acknowledgments

This research is in part supported by the Medical Research Endowment Fund G1-41175-01, University of Arkansas for Medical Sciences. We thank Dr. Abdus S. Wahed from University of Pittsburgh, the handling Editor of the Journal of Statistical Software, and the referees for their insightful comments. The authors thank Suzanne J. Speaker for assistance in preparation of the manuscript. The authors also thank Jeffrey D. Newmiller for kindly providing the functions for plotting the 95% confidence bands of the survival functions.

References

Bembom O, van der Laan MJ (2007). "Statistical Methods for Analyzing Sequentially Randomized Trials." *Journal of the National Cancer Institute*, **99**(21), 1577–1582.

Buyze J, Goetghebeur E (2013). "Evaluating Dynamic Treatment Strategies: Does It Have to be More Costly?" *Pharmaceutical Statistics*, **12**(1), 35–42.

Chakraborty B, Moodie EEM (2013). *Statistical Methods for Dynamic Treatment Regimes*. Springer-Verlag, New York.

Chakraborty B, Murphy SA (2013). "Dynamic Treatment Regimes." *Annual Review of Statistics and Its Application*, **1**, 447–464.

Dawson R, Lavori P (2010). "Sample Size Calculations for Evaluating Treatment Policies in Multi-Stage Designs." *Clinical Trials*, **7**(6), 643–652.

Dawson R, Lavori P (2012). "Efficient Design and Inference for Multistage Randomized Trials of Individualized Treatment Policies." *Biostatistics*, **13**(1), 142–152.

Feng W, Wahed AS (2008). "Supremum Weighted Log-Rank Test and Sample Size for Comparing Two-Stage Adaptive Treatment Strategies." *Biometrika*, **95**(3), 695–707.

Fleming TR, Harrington DP (1991). *Counting Processes and Survival Analysis*. John Wiley & Sons, New York.

Guo X, Tsiatis AA (2005). "A Weighted Risk Estimator for Survival Distributions in Two-Stage Randomization Designs with Censored Survival Data." *The International Journal of Biostatistics*, **1**(1), 1–15.

Kidwell KM (2014). "SMART Designs in Cancer Research: Past, Present, and Future." *Clinical Trials*, **11**(4), 445–456.

Kidwell KM, Wahed AS (2013). "Weighted Log-Rank Statistic to Compare Shared-Path Adaptive Treatment Strategies." *Biostatistics*, **14**(2), 299–312.

Lavori P, Dawson R (2000). "A Design for Testing Clinical Strategies: Biased Adaptive Within-Subject Randomization." *Journal of the Royal Statistical Society A*, **163**(1), 29–38.

Lei H, Nahum-Shani I, Lynch K, Oslin D, Murphy SA (2012). "A SMART Design for Building Individualized Treatment Sequences." *The Annual Review of Psychology*, **8**, 21–48.

Lunceford JK, Davidian M, Tsiatis AA (2002). "Estimation of Survival Distributions of Treatment Policies in Two-Stage Randomization Designs in Clinical Trials." *Biometrics*, **58**(1), 48–57.

Matthay KK, Reynolds CP, Seeger RC, Shimada H, Adkins ES, Haas-Kogan D, *et al.* (2009). "Long-Term Results for Children with High-Risk Neuroblastoma Treated on a Randomized Trial of Myeloablative Therapy Followed by 13-Cis-Retinoic Acid: A Children's Oncology Group Study." *Journal of Clinical Oncology*, **27**(7), 1007–1013.

Matthay KK, Villablanca JG, Seeger RC, Stram DO, Harris RE, Ramsay NK, *et al.* (1999). "Treatment of High-Risk Neuroblastoma with Intensive Chemotherapy, Radiotherapy, Autologous Bone Marrow Transplantation, and 13-Cis-Retinoic Acid." *The New England Journal of Medicine*, **341**(16), 1165–1173.

Murphy SA (2005). "An Experimental Design for the Development of Adaptive Treatment Strategies." *Statistics in Medicine*, **24**(10), 1455–1481.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Robins J (1986). "A New Approach to Causal Inference in Mortality Studies with Sustained Exposure Periods-Application to Control of the Healthy Survivor Effect." *Mathematical Modelling*, **7**(9–12), 1393–1512.

Robins JM (1987). "Addendum to "A New Approach to Causal Inference in Mortality Studies with Sustained Exposure Periods-Application to Control of the Healthy Survivor Effect"." *Computers & Mathematics with Applications*, **14**(9–12), 923–945.

Tang X, Melguizo M (2015). ***DTR**: Estimation and Comparison of Dynamic Treatment Regimes*. R package version 1.6, URL http://CRAN.R-project.org/package=DTR.

Tang X, Wahed AS (2011). "Comparison of Treatment Regimes with Adjustment for Auxiliary Variables." *Journal of Applied Statistics*, **38**(12), 2925–2938.

Tang X, Wahed AS (2015). "Cumulative Hazard Ratio Estimation for Treatment Regimes in Sequentially Randomized Clinical Trials." *Statistics in Biosciences*. doi:10.1007/s12561-013-9089-6.

Therneau TM (2015). ***survival**: A Package for Survival Analysis in S*. R package version 2.38, URL http://CRAN.R-project.org/package=survival.

Therneau TM, Grambsch PM (2000). *Modeling Survival Data: Extending the Cox Model*. Springer-Verlag, New York.

Wahed AS, Tsiatis AA (2004). "Optimal Estimator for the Survival Distribution and Related Quantities for Treatment Policies in Two-Stage Randomization Designs in Clinical Trials." *Biometrics*, **60**(1), 124–133.

Wahed AS, Tsiatis AA (2006). "Semiparametric Efficient Estimation of Survival Distributions in Two-Stage Randomization Designs in Clinical Trials with Censored Data." *Biometrika*, **93**(1), 163–177.

A. Computation of Equation 2

The detailed steps for computing $\widehat{\mathsf{VAR}}\{\hat{S}_{jk}(t)\}$ using Equation 2 are as follows:

Step 1: Calculate the survival estimate regardless of treatment regimes:

$$\hat{S}(u) = 1 - \left\{\sum_{i=1}^{n} \frac{\Delta_i}{\hat{K}(U_i)}\right\}^{-1} \sum_{i=1}^{n} \frac{\Delta_i}{\hat{K}(U_i)} I(U_i \leq u).$$

Step 2: Calculate:

$$\hat{G}_{jk}(t,u) = \{n\hat{S}(u)\}^{-1} \sum_{i=1}^{n} \frac{\Delta_i W'_{jki}}{\hat{K}(U_i)} \left\{I(U_i \leq t) - 1 + \hat{S}_{jk}(t)\right\} I(U_i \geq u).$$

Step 3: Calculate:

$$\hat{\mathsf{E}}\left\{L_{jki}(t,u)\right\}^2 = \frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i}{\hat{K}(U_i)} \left[W'_{jki}\left\{I(U_i \leq t) - 1 + \hat{S}_{jk}(t)\right\} - \hat{G}_{jk}(t,u)\right]^2 I(U_i \geq u).$$

Step 4: Calculate:

$$\begin{aligned}
\widehat{\mathsf{VAR}}\{\hat{S}_{jk}(t)\} &= \frac{1}{n}\left\{\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i W'^2_{jki}}{\hat{K}(U_i)} \left\{I(U_i \leq t) - 1 + \hat{S}_{jk}(t)\right\}^2 \right. \\
&\quad \left. + \int_0^L \frac{dN^c(u)}{\hat{K}(u)Y(u)} \hat{\mathsf{E}}\left\{L_{jki}(t,u)\right\}^2 \right\} \\
&= \frac{1}{n}\left\{\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i W'^2_{jki}}{\hat{K}(U_i)} \left\{I(U_i \leq t) - 1 + \hat{S}_{jk}(t)\right\}^2 \right. \\
&\quad \left. + \sum_{p=1}^{n} \int_0^L \frac{\hat{\mathsf{E}}\left\{L_{jki}(t,u)\right\}^2}{\hat{K}(u)Y(u)} dN^c_p(u) \right\} \\
&= \frac{1}{n}\left\{\frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i W'^2_{jki}}{\hat{K}(U_i)} \left\{I(U_i \leq t) - 1 + \hat{S}_{jk}(t)\right\}^2 \right. \\
&\quad \left. + \sum_{p=1}^{n} \frac{\hat{\mathsf{E}}\left\{L_{jki}(t,U_p)\right\}^2}{\hat{K}(U_p)Y(U_p)} (1-\Delta_p) I(U_p \leq L) \right\}.
\end{aligned}$$

B. Computation of Equation 4

The detailed steps for computing $\hat{\sigma}^2$ using Equation 4 are as follows:

$$\begin{aligned}
\hat{\sigma}^2 &= n\sum_{i=1}^{n}\left(\int_0^t \frac{W'_{jki}(u)dN_i(u)}{\sum_{p=1}^{n} W'_{jkp}(u)Y_p(u)} - \int_0^t \frac{W'_{jki}(u)Y_i(u)\sum_{p=1}^{n} W'_{jkp}(u)dN_p(u)}{\left[\sum_{p=1}^{n} W'_{jkp}(u)Y_p(u)\right]^2}\right)^2 \\
&= n\sum_{i=1}^{n}\left(\int_0^t \frac{W'_{jki}(u)dN_i(u)}{\sum_{p=1}^{n} W'_{jkp}(u)Y_p(u)} - \sum_{p=1}^{n}\int_0^t \frac{W'_{jki}(u)Y_i(u)W'_{jkp}(u)dN_p(u)}{\left[\sum_{q=1}^{n} W'_{jkq}(u)Y_q(u)\right]^2}\right)^2 \\
&= n\sum_{i=1}^{n}\left(\frac{W'_{jki}(U_i)\Delta_i I(U_i \le t)}{\sum_{p=1}^{n} W'_{jkp}(U_i)Y_p(U_i)} - \sum_{p=1}^{n}\frac{W'_{jki}(U_p)Y_i(U_p)W'_{jkp}(U_p)\Delta_p I(U_p \le t)}{\left[\sum_{q=1}^{n} W'_{jkq}(U_p)Y_q(U_p)\right]^2}\right)^2.
\end{aligned}$$

C. Computation of Equations 7 and 8

The detailed steps for computing $Z^W_{jkjk'}(t)$ using Equation 7 are as follows:

$$\begin{aligned}
Z^W_{jkjk'}(t) &= \int_0^t \frac{Y^*_{jk}(s)Y^*_{jk'}(s)}{Y^*_{jk}(s)+Y^*_{jk'}(s)}\left\{\frac{dN^*_{jk}(s)}{Y^*_{jk}(s)} - \frac{dN^*_{jk'}(s)}{Y^*_{jk'}(s)}\right\} \\
&= \int_0^t \frac{Y^*_{jk'}(s)dN^*_{jk}(s)}{Y^*_{jk}(s)+Y^*_{jk'}(s)} - \int_0^t \frac{Y^*_{jk}(s)dN^*_{jk'}(s)}{Y^*_{jk}(s)+Y^*_{jk'}(s)} \\
&= \sum_{i=1}^{n}\int_0^t \frac{Y^*_{jk'}(s)W_{jki}(s)dN_{jki}(s)}{Y^*_{jk}(s)+Y^*_{jk'}(s)} - \sum_{i=1}^{n}\int_0^t \frac{Y^*_{jk}(s)W_{jk'i}(s)dN_{jk'i}(s)}{Y^*_{jk}(s)+Y^*_{jk'}(s)} \\
&= \sum_{i=1}^{n}\frac{Y^*_{jk'}(U_i)W_{jki}(U_i)\Delta_i I(U_i \le t)}{Y^*_{jk}(U_i)+Y^*_{jk'}(U_i)} - \sum_{i=1}^{n}\frac{Y^*_{jk}(U_i)W_{jk'i}(U_i)\Delta_i I(U_i \le t)}{Y^*_{jk}(U_i)+Y^*_{jk'}(U_i)},
\end{aligned}$$

where $Y^*_{jk}(U_i) = \sum_{p=1}^{n} W_{jkp}(t)Y_p(U_i)$ and $Y^*_{jk'}(U_i) = \sum_{p=1}^{n} W_{jk'p}(t)Y_p(U_i)$.

The detailed steps for computing $\hat{\sigma}^2(t)$ using Equation 8 are as follows:

$$\begin{aligned}
\hat{\sigma}^2(t) &= \frac{1}{n}\int_0^t \frac{{Y^*_{jk'}}^2(s)\sum_{p=1}^n W^2_{jkp}(s)Y_{jp}(s) + {Y^*_{jk}}^2(s)\sum_{p=1}^n W^2_{jk'p}(s)Y_{jp}(s)}{\left\{Y^*_{jk}(s)+Y^*_{jk'}(s)\right\}^2}\left\{\frac{dN_j(s)}{Y_j(s)}\right\} \\
&\quad -\frac{2}{n}\int_0^t \frac{Y^*_{jk}(s)Y^*_{jk'}(s)}{\left\{Y^*_{jk}(s)+Y^*_{jk'}(s)\right\}^2}\left\{\pi_j^{-2}Y_j^{NR}(s)\frac{dN_j(s)}{Y_j(s)}\right\} \\
&= \frac{1}{n}\sum_{i=1}^n\int_0^t \frac{{Y^*_{jk'}}^2(s)\sum_{p=1}^n W^2_{jkp}(s)Y_{jp}(s) + {Y^*_{jk}}^2(s)\sum_{p=1}^n W^2_{jk'p}(s)Y_{jp}(s)}{\left\{Y^*_{jk}(s)+Y^*_{jk'}(s)\right\}^2 Y_j(s)}dN_{ji}(s) \\
&\quad -\frac{2}{n}\sum_{i=1}^n\int_0^t \frac{Y^*_{jk}(s)Y^*_{jk'}(s)Y_j^{NR}(s)}{\pi_j^2\left\{Y^*_{jk}(s)+Y^*_{jk'}(s)\right\}^2 Y_j(s)}dN_{ji}(s) \\
&= \frac{1}{n}\sum_{i=1}^n \frac{{Y^*_{jk'}}^2(U_i)\left[\sum_{p=1}^n W^2_{jkp}(U_i)Y_{jp}(U_i)\right]X_{ji}\Delta_i I(U_i\le t)}{\left\{Y^*_{jk}(U_i)+Y^*_{jk'}(U_i)\right\}^2 Y_j(U_i)} \\
&\quad +\frac{1}{n}\sum_{i=1}^n \frac{{Y^*_{jk}}^2(s)\left[\sum_{p=1}^n W^2_{jk'p}(U_i)Y_{jp}(U_i)\right]X_{ji}\Delta_i I(U_i\le t)}{\left\{Y^*_{jk}(U_i)+Y^*_{jk'}(U_i)\right\}^2 Y_j(U_i)} \\
&\quad -\frac{2}{n}\sum_{i=1}^n \frac{Y^*_{jk}(U_i)Y^*_{jk'}(U_i)Y_j^{NR}(U_i)X_{ji}\Delta_i I(U_i\le t)}{\pi_j^2\left\{Y^*_{jk}(U_i)+Y^*_{jk'}(U_i)\right\}^2 Y_j(U_i)},
\end{aligned}$$

where $Y_{jp}(U_i) = X_{jp}Y_p(U_i)$, $Y_j(U_i) = \sum_{p=1}^n X_{jp}Y_p(U_i)$, and $Y^*_{jk}(U_i)$ and $Y^*_{jk'}(U_i)$ are defined above.

Affiliation:

Xinyu Tang
Biostatistics Program
Department of Pediatrics
College of Medicine
University of Arkansas for Medical Sciences
1 Children's Way, Slot 512-43
Little Rock, AR, United States of America
E-mail: xtang@uams.edu
URL: http://www.arpediatrics.org/index.php?option=com_uams&view=faculty&sap=00035081

Journal of Statistical Software http://www.jstatsoft.org/
published by the American Statistical Association http://www.amstat.org/
Volume 65, Issue 7 *Submitted:* 2013-12-01
May 2015 *Accepted:* 2014-09-23