A New Multistate Transition Model For Effect Estimation In Randomized Trials With Treatment Switching And A Cured Subgroup

Haixia Hu (huhx_vero@163.com)
Air Force Medical University
Ling Wang
Air Force Medical University
Chen Li
Air Force Medical University
Wei Ge
Air Force Medical University
Jielai Xia
Air Force Medical University

Research Article

Keywords: Multistate model, treatment switching, cured subgroup, semi-competing risks, effect estimation

Posted Date: January 3rd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1183057/v1

Title: A New Multistate Transition Model for Effect Estimation in Randomized Trials with Treatment Switching and a Cured Subgroup

Running title: Multistate Transition Model for Treatment Switching

Authors: Haixia Hu[1], Ling Wang[1], Chen Li[1], Wei Ge[1], and Jielai Xia[1,2]

1 Department of Health Statistics, Faculty of Preventive Medicine, Air Force Medical University, Xi'an, Shaanxi, China.

2 Correspondence to: Jielai Xia. Department of Health Statistics, Faculty of Preventive Medicine, Air Force Medical University, No.169 Changle West Road, Xi'an 710032, Shaanxi, China. Email: jielaixia@yahoo.com.

Abstract

Background: Many methods, including multistate models, have been proposed in the literature to estimate the treatment effect on overall survival in randomized trials with treatment switching permit after the disease progression. Nevertheless, the cured fraction of patients has not been considered. The cured would never experience the progressive disease, but they may suffer death with a hazard comparable to that of people without the disease. With the mix of the cured subgroup, existing methods yield highly biased effect estimation and fail to reflect the truth in uncured patients.

Methods: In this paper, we propose a new multistate transition model to incorporate the cure, progression, treatment switching, and death states during trials. In the proposed model, the probability of cure and the death hazard of the cured are modeled separately. For the not cured patients, the semi-competing risks model is used with the treatment effect evaluated via transitional hazards between states. The particle swarm optimization algorithm is adopted to estimate the model parameters.

Results: Extensive simulation studies have been conducted to evaluate the performance of the proposed multistate model and compare it with existing treatment switching adjustment methods. Results show that in all scenarios, the treatment effect estimation of the proposed model is more accurate than that of existing treatment switching adjustment methods. Besides, the application to diffuse large B-cell lymphoma data has also illustrated the superiority of the proposed model.

Conclusions: The superiority and robustness of the proposed multistate transition model qualify it to estimate the treatment effect in trials with the treatment switching permit after

progression and a cured subgroup.

Keywords: Multistate model; treatment switching; cured subgroup; semi-competing risks; effect estimation

Background

In randomized controlled trials (RCT), patients are randomly assigned to either experimental or control groups and followed up until the occurrence of the event of interest or the termination of the trial. In studies comparing the effectiveness of different treatments, progression-free survival (PFS) is often adopted as a surrogate endpoint of overall survival (OS) due to shorter follow-up requirements and well prediction for OS. In some trials with PFS as the primary endpoint, patients in the control group are permitted to switch onto the experimental group after progression if the experimental efficacy is uncovered by existing or external data. The treatment switching allowance is increasingly common, especially in cancer clinical trials [1]. It is driven by ethical and practice concerns [2, 3]. It is ethically recommended to offer patients early access to the new treatment to obtain potential benefits. Besides, the treatment switching permit has the advantage of boosting trial recruitment without impacting on the analysis of short-term endpoint (i.e., PFS), since the balance between groups is not damaged before treatment switching.

However, for studies on the long-term treatment effect (i.e., OS), which is critical for health technology assessment (HTA) decision-making, the treatment switching contributes to considerable confounding [4]. Many statistical methods have been proposed to adjust for treatment switching in the literature. Robins and Tsiatis [5] proposed the rank preserving structural failure time model (RPSFTM) to estimate the treatment effect in trials with treatment switching. Via accelerated failure time (AFT) models, the counterfactual latent failure time of switchers is calculated and the grid search is adopted to obtain the accelerated factor (AF) estimate based on the randomization. In the process of counterfactual failure time calculation,

White et al. [6] explained the necessity of re-censoring in breaking the dependence of censoring time and treatment. Further, Branson and Whitehead [7] proposed an iterative parameter estimation (IPE) algorithm to adjust for treatment crossover. Similarly, re-censoring is also needed in the IPE algorithm. The difference is that the likelihood-based analysis is used in IPE instead of the rank-test approach in RPSFTM, which greatly accelerates the effect estimation computation. The IPE method has been extended to the prognostic assessment-based treatment switching [8]. Both the RPSFTM and IPE methods are randomization-based and have the key assumption of common treatment effect (CTE). The CTE assumption limits their application. In cases that the CTE assumption is not satisfied, i.e., the treatment effect in switchers is different from the treatment effect in patients initially randomized to the experimental treatment, a two-stage estimation (TSE) method was proposed by Latimer et al [9]. The treatment effect in switchers, which is defined as the switching effect in this paper, is estimated by comparing the survival of switchers and no-switchers in control group after the progressive disease (PD). The time of PD is regarded as the "second baseline". To obtain the unbiased estimate of the switching effect, the "no unmeasured confounders" assumption is required at the second baseline. The inverse probability of censoring weights (IPCW) method [10], an observational-based method to estimate treatment effect in studies with treatment switching, is also constrained by the "no unmeasured confounders" assumption. Extensive comparisons of these treatment switching adjustment methods have been carried out in the literature [2, 9, 11-15]. A major flaw of these methods is that the treatment effect change with patient status is neglected. Some patients may experience PD before death while some do not. For patients who experienced PD, the treatment effects on OS before and after PD are probably different, which

has not been accounted for in these methods.

Multistate models have been proposed in the literature to describe the different trajectories to death and the corresponding treatment effect. Specifically, Zeng et al [16] proposed to model the observed PD and death times via a semi-competing risks model. The logistic model was used to model the progression status (i.e., progression or not before death). Semi-parametric hazard models are used in the transitions between randomization to death, randomization to progression, and progression to death. Zhang et al [17] extended Zeng's method with the gamma frailty model to account for dependence between PD and death times and a Bayesian procedure was adopted to estimate the treatment effect. Huang et al [18] adopted copula models to establish the joint distribution of PD and death times subsequently. Chen et al [19] extended Zeng's method to accommodate two-way time-varying switching. With design-based treatment switching in the context of recurrent events data, Chen [20] proposed a semiparametric frailty modeling approach to estimate time-varying effects.

Though gradually refined and improved, the semi-competing risks model has not considered the subgroup for whom the disease is cured. Those patients are insusceptible to PD. It is worth noting that the cured or not is different from the patient disparity in trajectories to death mentioned before. In the semi-competing risks model, some patients may not experience PD due to the censoring of death. However, the cure means that patients would never experience PD even without the censoring of death. With the development of medicine, many diseases, such as early-stage cancers [21], could be clinically curable and patients would never experience PD and would have a life expectancy similar to that of people without the disease. In the analysis of acute graft-versus-host disease (aGVHD), Lee et al [22] took the fraction of

patients insusceptible to the aGVHD into consideration and proposed a novel multistate model incorporating semi-competing risks model and the cured proportion. Treatment switching has not been considered in Lee's study nevertheless. Zhang et al [23] have also considered the cured fraction to the disease and they adopted a cure rate model for the PD time. But the emphasis of their study was the joint modeling of the longitudinal biomarkers and the survival data, the treatment effect assessment accounting for switching has not been comprehensively considered.

In this paper, we propose a novel multistate transition model which includes the cured fraction of patients, the semi-competing risks between progression and death, and the treatment switching. In the proposed model, logistic models are adopted for the patient heterogeneity (i.e., cured or not) and the treatment switching choice, semi-competing risks model with patient-specific shared frailty is used for the progression and death times, as well as the correlation between them. The treatment effect on cure rate, progression, and death hazard is quantitatively measured using the coefficient of the covariate "group" in each sub-model. The proposed model evaluates the treatment effects on the cured and uncured subgroups separately. The impact of the cured subgroup on the treatment effect estimation in uncured patients is eliminated successfully.

Methods

Multistate transition model

In this subsection, we introduce the proposed multistate transition model in detail. Five states are incorporated in the model, including cured, not cured, progression, treatment switching, and death. The transitions between these five states are shown in Figure 1. The cured fraction (bold in Figure 1) is innovatively considered in the proposed model in comparison with

Zeng's semi-competing risks model [16].

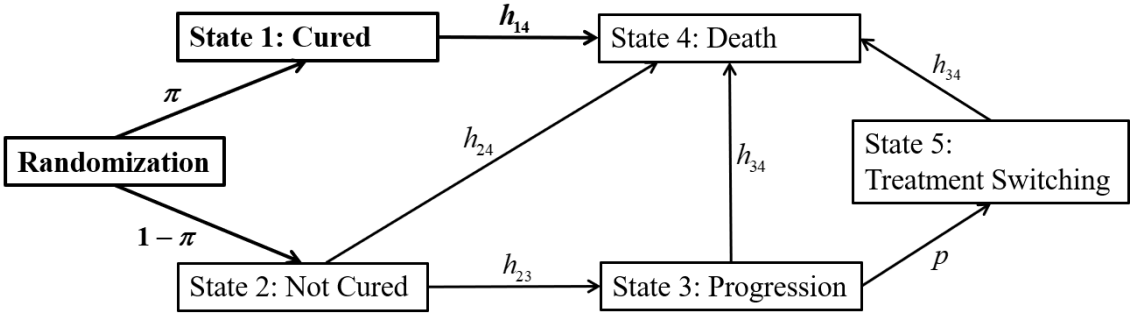


Figure 1. Proposed multistate transition model in which the cured fraction (bold) is innovatively considered.

In RCTs comparing the effectiveness of experimental and control treatments, a proportion of patients might be cured and they would never experience the PD. Let *s* be the cure indicator with *s*=1 representing the cured patient and *s*=0 representing the not cured patient. The probability of being cured (denoted by π) is expressed as

$$\pi = P(s=1) = \frac{1}{1+\exp\left[-(a_0 + a_1 trt)\right]}, \quad (1)$$

where trt is the treatment indicator with $trt=1$ representing the experimental treatment and $trt=0$ representing the control treatment. With the specified cure rate of experimental and control treatments (denoted by $\pi_{\exp}$ and π_{ctr}, respectively), the intercept (a_0) and coefficient (a_1) in equation (1) could be calculated easily. Assume that patients randomized to control treatment are allowed to switch onto the experimental treatment after PD. The treatment switching is voluntary and may be influenced by the time to PD, which is denoted by t_{PD}. For example, patients who do well and have longer progression time are advised to stay on the control treatment rather than switching [24]. Similarly, a logistic model is used to establish the probability of treatment switching (denoted by *p*), i.e.,

$$p = P\left(V = 1 \mid s = 0, trt = 0, \delta_{PD} = 1\right) = \frac{1}{1 + \exp\left[-\left(b_0 + b_1 Q_{t_{PD}}\right)\right]}, \quad (2)$$

where V is the treatment switching indicator with $V = 1$ representing switching and $V = 0$ otherwise, δ_{PD} is the indicator for PD occurrence with $\delta_{PD} = 1$ representing that the time to PD is observed and $\delta_{PD} = 0$ otherwise. $Q_{t_{PD}}$ is an ordinal categorical variable converted by t_{PD}, with $Q_{t_{PD}} = 1, 2, 3, 4$ representing that the t_{PD} is at <25%, 25%-50%, 50%-75%, >70% quantiles of the progression time, respectively. The ordinal categorical variable $Q_{t_{PD}}$ rather than the original continuous variable t_{PD} is used to describe the impact of the progression time on the treatment switching probability for ease of clinical interpretation. Based on similar studies [16, 22] in the literature, we express the transition hazards between states 1-4 as follows.

$$h_{14}\left(t_D \mid s = 1\right) = \gamma \lambda_{14}\left(t_D\right), 0 < t_D \quad (3)$$

$$h_{23}\left(t_{PD} \mid trt, s = 0\right) = \gamma \lambda_{23}\left(t_{PD}\right) \exp\left(\beta_{23} trt\right), 0 < t_{PD} \quad (4)$$

$$h_{24}\left(t_D \mid trt, s = 0\right) = \gamma \lambda_{24}\left(t_D\right) \exp\left(\beta_{24} trt\right), 0 < t_D \quad (5)$$

$$h_{34}\left(t_{PPD} \mid t_{PD}, trt, s = 0\right) = \gamma \lambda_{34}\left(t_{PPD}\right) \exp\left(\beta_{34,1} trt + \beta_{34,2} V\left(1 - trt\right)\right), 0 < t_{PD} < t_D \quad (6)$$

In equations (3)-(6), t_{PD}, t_D, t_{PPD} represent the time to PD, time to death without progression, and time to post-progression death, respectively, γ is the patient-specific shared frailty following a gamma distribution with mean one and variance θ, i.e., $Ga\left(\theta^{-1}, \theta^{-1}\right)$ [25]. The correlation between the progression time and the death time, as well as the patient heterogeneity, is reflected by the shared frailty. In equation (3), only the baseline hazard is included and the treatment has no effect on the death hazard of the cured subgroup, as we assume that the cured have similar death hazards as people who have never suffered the disease. In equation (6), the coefficient $\beta_{34,1}$ and $\beta_{34,2}$ represent the experimental treatment effect and the treatment switching effect, respectively, on the hazard of post-progression death. Under the assumption

of CTE, we have $\beta_{34,1} = \beta_{34,2} = \beta_{24}$. For treatments that have a smaller effect on progressed patients, $\beta_{34,1}$ is closer to zero than β_{24}. In cases that patients switched onto the experimental treatment after progression benefit less compared to patients randomized to the experimental treatment at first, $\beta_{34,2}$ is closer to zero than $\beta_{34,1}$.

Maximum likelihood estimation (MLE) of the treatment effect

In equations (1)-(6), the cure rate, treatment switching probability, and the transition hazards between states are explicitly modeled. Based on the observed data, the MLE method is used to estimate the parameters in the model.

Suppose a trial with N patients, the observed data for each patient include $\left(trt_i, t_{PDi}, \delta_{PDi}, t_{Di}, \delta_{Di}\right), i = 1, 2, ..., N$, where trt_i is the treatment indicator, t_{PDi} and t_{Di} are the PD and death times, respectively, δ_{PDi} and δ_{Di} are the censoring indicator for PD and death, respectively, for the *i*-th patient. The PD time could be censored by both death and trial termination (denoted by τ) while the death time could only be censored by τ. For patients who have experienced PD, we have $t_{Di} = t_{PDi} + t_{PPDi}$, where t_{PPDi} is the post-progression death time. Besides, for patients who have experienced PD in the control arm, the treatment switching indicator V illustrates whether he/she switches onto the experimental arm. During the follow-up of the trial, six scenarios may be observed regarding progression, treatment switching and death status (i.e., δ_{PD}, V, δ_D), as shown in Table 1. The likelihood of each status could be calculated accordingly.

Table 1. Observed status indicator and likelihood contribution corresponding to six possible scenarios during the trial.

Status	Description	Observed data $(\delta_{PD}, V, \delta_D)$	Likelihood
A	Observed PD, switching, and death	(1, 1, 1)	f_1
B	Observed PD, switching, and censored death	(1, 1, 0)	f_2
C	Observed PD and death without switching	(1, 0, 1)	f_3
D	Observed PD and censored death without switching	(1, 0, 0)	f_4
E	Observed death without PD	(0, -, 1)	f_5
F	Censored PD and death	(0, -, 0)	f_6

According to the trial setting, patients in the control arm are allowed to switch onto the experimental arm only after the occurrence of PD, the switching indicator is not applicable in statuses E and F since the PD is not observed. Besides, patients with observed PD are not cured without a doubt, i.e., patients in statuses A-D are not cured ($s=0$), while patients without observed PD could belong to either the cured or the not cured subgroup. Therefore, the likelihood functions $f_5 \sim f_6$ should be marginalized with respect to the distribution of *s*. Besides, the likelihood functions $f_1 \sim f_6$ are the marginal distributions with respect to the distribution of shared frailty (γ), i.e., $f_k(t) = \int f_k(t|\gamma) d\gamma, k = 1,2,...,6$. The corresponding conditional distributions (i.e., $f_k(t|\gamma), k = 1,2,...,6$) are shown as follows,

$$f_1(t_{PD}, t_{PPD}|\gamma) = P(s=0)\left[S_{24}(t_{PD}|\gamma) S_{23}(t_{PD}|\gamma) h_{23}(t_{PD}|\gamma)\right] P(V=1)^{I(trt=0)} \left[S_{34}(t_{PPD}|\gamma) h_{34}(t_{PPD}|\gamma)\right]$$

$$f_2(t_{PD}, t_{PPD}|\gamma) = P(s=0)\left[S_{24}(t_{PD}|\gamma) S_{23}(t_{PD}|\gamma) h_{23}(t_{PD}|\gamma)\right] P(V=1)^{I(trt=0)} S_{34}(t_{PPD}|\gamma)$$

$$f_3(t_{PD}, t_{PPD}|\gamma) = P(s=0)\left[S_{24}(t_{PD}|\gamma) S_{23}(t_{PD}|\gamma) h_{23}(t_{PD}|\gamma)\right] P(V=0)^{I(trt=0)} \left[S_{34}(t_{PPD}|\gamma) h_{34}(t_{PPD}|\gamma)\right]$$

$$f_4(t_{PD}, t_{PPD}|\gamma) = P(s=0)\left[S_{24}(t_{PD}|\gamma) S_{23}(t_{PD}|\gamma) h_{23}(t_{PD}|\gamma)\right] P(V=0)^{I(trt=0)} S_{34}(t_{PPD}|\gamma)$$

$$f_5(t_D|\gamma) = P(s=0)\left[S_{23}(t_D|\gamma) S_{24}(t_D|\gamma) h_{24}(t_D|\gamma)\right] + P(s=1)\left[S_{14}(t_D|\gamma) h_{14}(t_D|\gamma)\right]$$

189 $f_6(t_D|\gamma) = P(s=0)\left[S_{23}(t_D|\gamma)S_{24}(t_D|\gamma)\right] + P(s=1)S_{14}(t_D|\gamma)$

190 where $I(trt=0)$ is the control treatment indicator, $S_{24}(t|\gamma) = \exp\left[-\int_0^t h_{24}(u|\gamma)du\right]$,

191 $S_{23}(t|\gamma) = \exp\left[-\int_0^t h_{23}(u|\gamma)du\right]$, $S_{34}(t|\gamma) = \exp\left[-\int_0^t h_{34}(u|\gamma)du\right]$, and

192 $S_{14}(t|\gamma) = \exp\left[-\int_0^t h_{14}(u|\gamma)du\right]$. Assume the transitional baseline hazards are constant, i.e.,

193 $\lambda_{14}(t) = \lambda_{14}$, $\lambda_{23}(t) = \lambda_{23}$, $\lambda_{24}(t) = \lambda_{24}$, and $\lambda_{34}(t) = \lambda_{34}$, with marginal distributions of the

194 conditional distributions above, the observed data likelihood for

195 $\Phi = \left(a_0, a_1, b_0, b_1, \lambda_{14}, \lambda_{23}, \beta_{23}, \lambda_{24}, \beta_{24}, \lambda_{34}, \beta_{34,1}, \beta_{34,2}, \theta\right)$ is

196 $$L(\Phi) = \prod_{i=1}^{N}\left\{\left[\begin{array}{l}\left(f_1^{\delta_{PD}\delta_D V} f_2^{\delta_{PD}(1-\delta_D)V} f_3^{\delta_{PD}\delta_D(1-V)} f_4^{\delta_{PD}(1-\delta_D)(1-V)}\right)^{I(trt=0)} \\ \cdot\left(f_1^{\delta_{PD}\delta_D} f_2^{\delta_{PD}(1-\delta_D)}\right)^{I(trt=1)}\end{array}\right] f_5^{(1-\delta_{PD})\delta_D} f_6^{(1-\delta_{PD})(1-\delta_D)}\right\}, \quad (7)$$

197 where $I(trt=1)$ is the experimental treatment indicator. To maximize $L(\Phi)$ in equation (7),

198 let

199 $$\begin{aligned} Z(\Phi) &= \log\left[L(\Phi)\right] \\ &= \sum_{i=1}^{N}\left\{\begin{array}{l} I(trt=0)\left[\begin{array}{l}\left(\delta_{PD}\delta_D V\right)\log\left(f_1\right) + \left(\delta_{PD}(1-\delta_D)V\right)\log\left(f_2\right) \\ +\left(\delta_{PD}\delta_D(1-V)\right)\log\left(f_3\right) + \left(\delta_{PD}(1-\delta_D)(1-V)\right)\log\left(f_4\right)\end{array}\right] \\ +I(trt=1)\left[\left(\delta_{PD}\delta_D\right)\log\left(f_1\right) + \left(\delta_{PD}(1-\delta_D)\right)\log\left(f_2\right)\right] \\ +\left((1-\delta_{PD})\delta_D\right)\log\left(f_5\right) \\ +\left((1-\delta_{PD})(1-\delta_D)\right)\log\left(f_6\right)\end{array}\right\}. \end{aligned} \quad (8)$$

200 Then, the goal is to find Φ that maximizes $Z(\Phi)$ in equation (8).

201 ***Particle swarm optimization (PSO) algorithm***

202 There are many unknown parameters in Φ to estimate, a simple and easily implemented

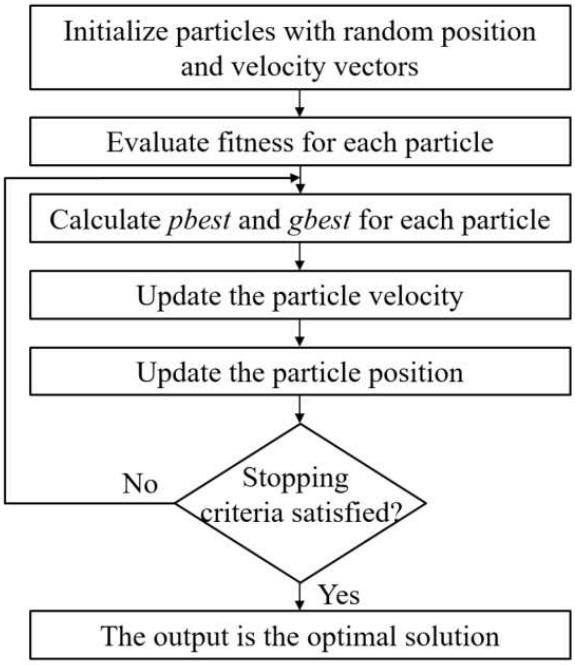
203 method is adopted. PSO algorithm is a population-based stochastic optimization technique

204 developed by Kennedy and Eberhart [26], inspired by the social behavior of bird flocking or

205 fish schooling. It has been successfully applied in the area of medicine, such as medical images

206 analysis [27], information clustering [28], disease diagnosis [29], and so on. The flow chart of

207 the PSO algorithm is shown in Figure 2 [30].



208

209 **Figure 2.** Flow chart of the PSO algorithm [30].

210 There are five steps in the PSO algorithm.

211 1) Initialize *M* solutions, which are named "particles" in PSO, for Φ. There are a position
212 vector and a velocity vector for each particle. The position vector (denoted by $\boldsymbol{\varphi}$)
213 represents the candidate solution to maximize $Z(\Phi)$ in equation (8), or similarly, to
214 minimize $-Z(\Phi)$, while the velocity vector (denoted by $\mathbf{v}$) represents the direction and
215 amplitude of the position change in the next iteration to search for a better solution. The
216 number of dimensions of the position and velocity vectors is equal to the number of
217 elements in Φ.

218 2) Calculate the value of the target function to evaluate the fitness of each particle. In the
219 context of this study, calculate $Z(\boldsymbol{\varphi}_m), m = 1, 2, ..., M$ in equation (8).

220 3) Calculate "*pbest*" and "*gbest*", where "*pbest*" is the best position of each particle in history

iterations while "*gbest*" is the best position among all particles in history iterations. The larger the value of $Z(\boldsymbol{\varphi})$, the better the position of particles.

4) Update the particle velocity and position. For the *m*-th particle, the velocity and position in the (*j*+1)-th iteration are

$$\mathbf{v}_{m(j+1)} = \omega \times \mathbf{v}_{m(j)} + c_1 \times u_1 \times \left(pbest_{m(j)} - \boldsymbol{\varphi}_{m(j)} \right) + c_2 \times u_2 \times \left(gbest_{(j)} - \boldsymbol{\varphi}_{m(j)} \right)$$

and $\boldsymbol{\varphi}_{m(j+1)} = \boldsymbol{\varphi}_{m(j)} + \mathbf{v}_{m(j+1)}, m = 1, 2, ..., M$,

where ω is the inertia weight, we set it 0.5 in this paper, more elaborated setting of ω are referenced to [31]. c_1 and c_2 are accelerating factors and we set them to be the commonly used value of two, i.e., $c_1 = c_2 = 2$. u_1 and u_2 are random numbers from the uniform distribution $U(0,1)$. To keep the best solution search within an appropriate space, the velocity is clamped to a maximum velocity, we set the maximum to be 0.1 in this study. Similarly, the search space is also limited for each parameter, which is discussed later.

5) Repeat steps 2-4 until the maximum iteration is reached or no better fitness is attained. The latter is manifested as that the increase of the target function $Z(\Phi)$ value for the "*gbest*" particle is smaller than a pre-specified threshold, which is 1e-8 in this paper.

The initial particles in step 1 of the PSO algorithm are set as follows. Firstly, we adopt logistic models and AFT models to fit the observed data to get the crude estimate of each parameter. Specifically, patients with observed PD are classified as not cured and patients without observed PD are classified as cured with a probability of 0.5. Then the estimates of a_0 and a_1 are obtained by fitting the cure status with a logistic model. Based on the patients who have observed PD in the control treatment arm, the logistic model is used to fit the treatment switching status, then the estimates of b_0 and b_1 are obtained. For parameters

λ_{14}, λ_{23}, β_{23}, λ_{24}, β_{24}, λ_{34}, $\beta_{34,1}$ and $\beta_{34,2}$, AFT models are used to fit the death time of classified cured patients, observed PD time of classified not cured patients, death time of classified not cured patients without observed PD, and post-progression death time of classified not cured patients with observed PD, respectively. The initial value of θ is set at 0.5. Secondly, define the solution search space for each parameter. Specifically, for the parameter θ, the search space is limited within $(0.1, 2.5)$, for parameters λ_{14}, λ_{23}, λ_{24} and λ_{34}, the search space is limited within the scope of crude point estimate multiplied by 0.01 to 100, while the search space for other parameters is limited within the scope of crude point estimate plus or minus one. Thirdly, generate the random initial position of each particle within the search space of each parameter. The elements in the initial velocity of each particle are generated randomly within $(-0.1, 0.1)$.

Results

Simulation study

Study design

We simulate independent datasets based on an RCT in which patients are randomized 1:1 to experimental and control groups and a certain proportion of patients could be cured in both groups. For cured patients, the progression would never happen, the death hazard is small and independent of the treatment. Uncured patients in the control group are permitted to switch onto the experimental group after PD and the switching probability is dependent on the time to PD. The transition hazards between states are known. For the basic scenario, the true values of the parameters are as follows.

$$\pi_{\exp}=30\%,\ \pi_{ctr}=15\%,\ p=30\%,\ b_1=\log(3),$$

$$\lambda_{14}=0.0003,\ \lambda_{23}=0.02,\ \beta_{23}=\log(0.4),\ \lambda_{24}=0.005,\ \beta_{24}=\log(0.4),$$

$$\lambda_{34}=0.03,\ \beta_{34,1}=\log(0.4),\ \beta_{34,2}=\log(0.4),\ \theta=1.$$

That is, the cure rates in experimental and control groups are 30% and 15%, respectively, the values of a_0 and a_1 are calculated accordingly. With the switching proportion of 30% and the value of b_1, the value of b_0 could be computed via iterations. Under the setting of $\lambda_{23}, \beta_{23}, \lambda_{24}$ and β_{24}, about 20% uncured patients would die without progression. Besides, in this scenario, $\beta_{23}=\beta_{24}=\beta_{34,1}=\beta_{34,2}=\log(0.4)$. That is, the experimental therapy has the same treatment effect on prolonging the PFS and OS, on prolonging OS in patients with or without PD, as well as in switchers.

Control methods for switching adjustment

To provide context on the performance of the proposed multistate transition model, we present the estimation results of several control methods. The control methods are grouped into simple treatment switching adjustment methods and multistate model methods. For the former, the treatment effect is assumed to be a constant for all patients. Only one parameter is to be estimated for these methods. These methods include:

1) intention to treat (ITT) analysis which ignores the treatment switching;
2) per-protocol analysis which censors the switchers at the PD time (PPcen);
3) per-protocol analysis which excludes the switchers from the analysis dataset (PPexc);
4) RPSFTM method with the treatment effect estimated via grid searching (GE) [5];
5) RPSFTM method with the treatment effect estimated via iteration parameter estimation (IPE) [7].

Re-censoring is considered in GE and IPE methods. Besides, the two-stage estimation (TSE)

method, which estimates the switching effect and treatment effect separately in two stages [9], is also classified into the category of simple treatment switching adjustment methods in this paper. The multistate model methods highlight the treatment effect heterogeneity across patient subcategories and patient statuses. In this paper, we consider the proposed multistate transition model without considering the shared frailty and Zeng's semi-competing risks model. For comparability, the PSO algorithm is adopted for the parameter estimation of Zeng's method instead of the expectation-maximization (EM) algorithm adopted in their research [16].

Investigated scenarios

To evaluate the robustness of the proposed multistate transition model, as well as the performance sensitivity of control methods concerning different trial settings, the simulation study covers several scenarios. Only one key variable changes across scenarios to show the influence of the variable on the performance of the proposed and control methods. Specifically, based on the basic scenario described in the "Study design" subsection, the changing variables are listed as follows across scenarios.

1) Cure rate: $\pi_{\exp} = \pi_{ctr} = 15\%$ with an equal cure rate between groups; $\pi_{\exp} = 30\%, \pi_{ctr} = 15\%$ with the cure rate ratio being two; and $\pi_{\exp} = 45\%, \pi_{ctr} = 15\%$ with the cure rate ratio being three.
2) Switching proportion: 10% (low), 30% (moderate), and 50% (high).
3) Treatment effect (TE) assumption:
   A. CTE satisfied: the treatment effect on OS is the same for switchers and patients initially randomized to experimental treatment, no matter whether the PD happens or not, i.e., $\beta_{24} = \beta_{34,1} = \beta_{34,2} = \log(0.4)$.

When CTE is not satisfied, two scenarios are considered as follows.

B. The treatment effects on OS are different for patients with or without PD and we assume the protective effect of experimental treatment is larger for patients without PD, i.e., $\beta_{24} = \log(0.4), \beta_{34,1} = \beta_{34,2} = \log(0.6)$.

C. The treatment effects for switchers and patients randomized to the experimental group at first are different and we assume the switchers benefit less, i.e., $\beta_{24} = \beta_{34,1} = \log(0.4), \beta_{34,2} = \log(0.6)$.

4) Sample size: 500, 1000, and 2000.

Performance measures

The main purpose of this paper is to estimate the experimental treatment effect on OS in RCTs with treatment switching permission after progression in control group and with a proportion of patients cured in both groups. In multistate models, including the proposed multistate transitional models, the treatment effects on transitions towards different states are described by different parameters (β_{23}, β_{24}, and $\beta_{34,1}$). We evaluate the performance of the proposed model and other multistate methods via the accuracy of the estimates of β_{23}, β_{24}, and $\beta_{34,1}$, as well as other parameters. The estimation accuracy is measured by the percentage bias (PB(%)), mean squared error (MSE), and empirical standard error (SE) of the estimate. Specifically, PB(%) is calculated as $\left(\bar{\hat{\beta}} - \beta\right) \Big/ |\beta| \times 100$, where β refers to the true value of parameters and $\bar{\hat{\beta}}$ is the mean of the parameter estimates. The percentage bias is more preferable to the bias because of the different scales of the parameters to estimate. MSE is calculated as $\sum\left(\hat{\beta} - \beta\right)^2 \Big/ N_{\text{itr}}$, where N_{itr} is the number of the simulation replicates.

Simple treatment switching adjustment methods (i.e., ITT, PPcen, PPexc, GE, IPE, and

TSE) ignore the treatment effect heterogeneity for different patient subcategories or the effect change with the PD status of patients. The treatment effect estimates of these methods reflect the average effect in all patients across all statuses. We evaluate them from two aspects. On the one hand, the treatment effect on death without progression (β_{24}) is served as the true value, since patients who have not progressed are recruited in the trial to evaluate the treatment effect on OS. On the other hand, under each scenario, we simulate data for 1 000 000 patients without incorporating treatment switching and estimate the average hazard ratio ($ave.\text{HR}$) on OS. The logarithm of the average HR, i.e., $\log\left(ave.\text{HR}\right)$, is served as the true value to evaluate the performance of the simple treatment switching adjustment methods. For GE, IPE, and TSE methods, the death time data corrected by the estimated AF are used to estimate the HR via a proportional hazard model [6]. It is worth noting that the average treatment effect estimate, i.e., $\log\left(ave.\text{HR}\right)$, is prone to error because it is obtained via simulation instead of the calculation, but the error is likely to be extremely minimal given the large number of patients simulated [32].

Simulation results

The simulation results are displayed in two parts. In the first part, the multistate models are compared (i.e., proposed multistate transition model with or without shared frailty versus Zeng's semi-competing risks model). In the second part, the performances of simple treatment switching adjustment methods are investigated. In either part, the estimation performances concerning different scenarios are presented. 100 datasets are simulated in each scenario.

Under the basic scenario, the estimation performances of the proposed multistate transition model and Zeng's semi-competing risks model are shown in Table 2. As shown in the third and

fourth columns, the proposed multistate transition model could provide an almost unbiased estimation of the parameters via the PSO algorithm, especially for the baseline hazards of transitions between states (i.e., $\lambda_{14}, \lambda_{23}, \lambda_{24}$, and λ_{34}). For parameters except λ_{14}, the estimation bias is less than 4% of the true value. From the results of the proposed multistate transition model without considering shared frailty (middle part in Table 2), we see that the estimation bias is greatly increased. It is observed that the baseline hazards (i.e., $\lambda_{14}, \lambda_{23}, \lambda_{24}$, and λ_{34}) are underestimated uniformly. The treatment effects are also underestimated from the perspective of HR (i.e., $\exp(\beta_{23}), \exp(\beta_{24})$, and $\exp(\beta_{34,1})$), while the treatment effect for switchers (i.e., $\exp(\beta_{34,2})$) is overestimated. The probabilities of cure and treatment switching are also overestimated. As seen in the last four columns in Table 2, the semi-competing risks model without considering the cured fraction and shared frailty leads to large estimation biases, particularly in the effect estimate on the death without progression (i.e., λ_{24}). The direction of the estimation bias for the semi-competing risks model is the same as that for the multistate transition model without considering shared frailty. Therefore, failing to consider the cured fraction and the shared frailty results in large estimation biases.

Table 2. Parameter estimation performances of the proposed multistate transition model with or without shared frailty and Zeng's semi-competing risks model under the basic scenario, N=2000.

Parameter	True value	Proposed multistate transition model				Proposed multistate transition model$^{\gamma-}$				Semi-competing risks model			
		Est	PB (%)	SE	MSE	Est	PB (%)	SE	MSE	Est	PB(%)	SE	MSE
a_0	-1.730	-1.770	**-2.338**	0.201	0.041	-1.174	**32.147**	0.069	0.314				
a_1	0.890	0.917	**3.065**	0.190	0.036	0.933	**4.848**	0.104	0.013				
b_0	-3.870	-3.917	**-1.224**	0.299	0.091	-3.012	**22.169**	0.301	0.826				
b_1	1.099	1.112	**1.223**	0.095	0.009	2.118	**92.770**	0.102	1.049				
λ_{14}	0.0003	0.0003	**-7.285**	0.001†	0.001†	0.001†	**-20.342**	0.001†	0.001†				
λ_{23}	0.020	0.020	**-0.913**	0.002	0.001†	0.013	**-35.298**	0.001	0.001†	0.016	**-21.880**	0.001	0.001†
β_{23}	-0.916	-0.913	**0.371**	0.117	0.014	-0.536	**41.490**	0.075	0.150	-0.534	**41.773**	0.073	0.152
λ_{24}	0.005	0.005	**1.722**	0.001†	0.001†	0.003	**-36.624**	0.001†	0.001†	0.002	**-62.394**	0.001†	0.001†
β_{24}	-0.916	-0.917	**-0.118**	0.146	0.021	-0.566	**38.236**	0.129	0.139	-0.786	**14.181**	0.112	0.029
λ_{34}	0.030	0.030	**-0.877**	0.002	0.001†	0.019	**-36.511**	0.001	0.001†	0.019	**-36.790**	0.001	0.001†
$\beta_{34,1}$	-0.916	-0.896	**2.216**	0.114	0.013	-0.654	**28.644**	0.098	0.078	-0.678	**26.016**	0.107	0.068
$\beta_{34,2}$	-0.916	-0.908	**0.950**	0.134	0.018	-1.084	**-18.354**	0.128	0.045	-1.127	**-23.001**	0.121	0.059
θ	1.000	0.991	**-0.922**	0.125	0.016								

Est: parameter estimate; PB(%): percentage bias; SE: empirical standard error; MSE: mean squared error. $^{\gamma-}$: multistate transition model without shared frailty. †: values less than 0.001.

Results of the comparison among the three multistate model methods under other scenarios are shown in Tables S1-S8 of the Appendix 1 in Additional file. On the whole, the performances of the three methods in scenarios of different cure rates, switching proportions, TE assumptions, and sample sizes are similar to that in the basic scenario. The bias of the parameter estimate based on the proposed multistate transition model is consistently small across different scenarios except for the case of a small sample size. When N=500 (see Table S8), the estimation bias cannot be ignored. The sample size of 500 might be too small to estimate 13 parameters at once. Besides, for the parameter λ_{14}, the estimation in some scenarios is not satisfactory due to the impact of extreme estimates. The possible reason is that the true value of λ_{14} (i.e., 0.0003) is too small to be easily affected by inaccurate estimates. In all scenarios, the parameter estimation biases are larger when the shared frailty is not considered in the proposed multistate transition model. For Zeng's semi-competing risks model, the parameter estimation bias is large in all scenarios. With the change of cure rates, switching proportions, and TE assumptions, the estimation biases of certain parameters change significantly. Specifically, when the cure rate ratio increases, the parameter β_{24} is greatly overestimated (see Tables S1-S2 and Table 2). Because the increased cured patients in the experimental group greatly decrease the overall death hazard, which amplifies the treatment effect of the experimental therapy. With the increase of the switching proportion, the estimation biases of the parameters λ_{34} and $\beta_{34,1}$ decrease while the estimation bias of the parameter $\beta_{34,2}$ increases (see Tables S3-S4 and Table 2). Under different TE assumptions, the estimation biases of parameters $\beta_{34,1}$ and $\beta_{34,2}$ change. Multiple factors may contribute to the result and it is hard to figure out. With the decrease of the sample size, the SE and MSE of the parameter estimate increase, which is true

for all three methods (see Tables S7-S8 and Table 2).

Figure 3 shows PB(%) of the parameter estimate with the proposed multistate transition model and simple treatment switching adjustment methods (i.e., ITT, PPcen, PPexc, GE, IPE, and TSE) under scenarios of different cure rates (Figure 3 A1-A2), switching proportions (Figure 3 B1-B2), TE assumptions (Figure 3 C1-C2), and sample sizes (Figure 3 D1-D2). The logarithm of HR for death without progression (i.e., β_{24}) is served as the true value as shown in the left column in Figure 3. Besides, for simple treatment switching adjustment methods, the estimation performance is also evaluated by comparing the estimate with the logarithm of average HR for death across all patients, as shown in the right column in Figure 3.

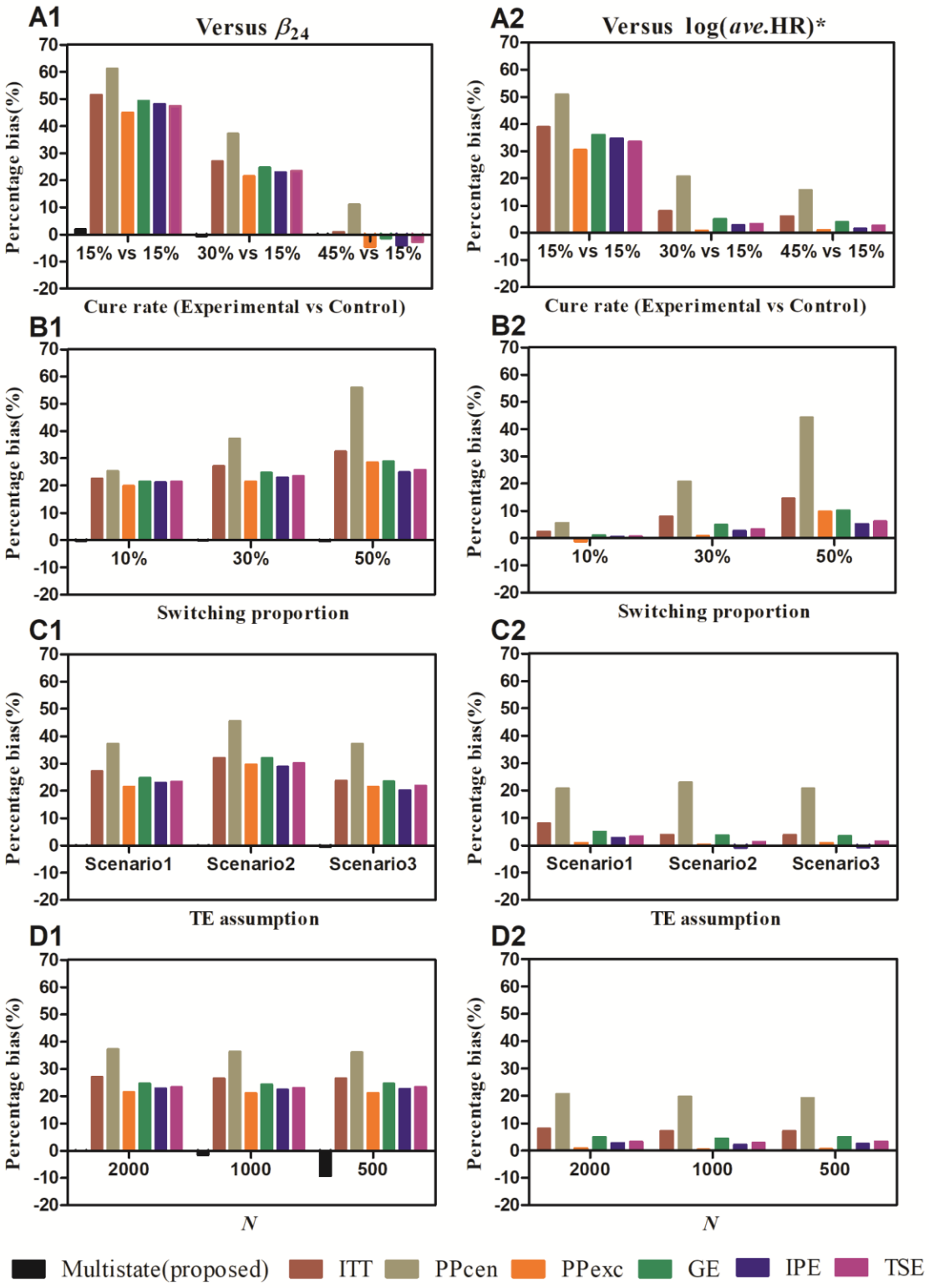


Figure 3. PB(%) of the estimate using the proposed model and contrast methods under different scenarios.

*: Only for control methods (i.e., homogeneity assumption-based methods). TE: treatment effect. In C1-C2,

scenario 1 is the case that the CTE assumption is satisfied with $\beta_{24} = \beta_{34,1} = \beta_{34,2} = \log(0.4)$; scenario 2 is the case that the CTE assumption is not satisfied with $\beta_{24} = \log(0.4), \beta_{34,1} = \beta_{34,2} = \log(0.6)$; scenario 3 is the case that the CTE assumption is not satisfied with $\beta_{24} = \beta_{34,1} = \log(0.4), \beta_{34,2} = \log(0.6)$. *N*: sample size.

As shown in Figure 3, for the estimation of β_{24} (left column), the bias of the proposed multistate transition model is much smaller than that of simple treatment switching adjustment methods (i.e., ITT, PPcen, PPexc, GE, IPE, and TSE). Besides, the performance of the proposed multistate transition model is robust with scenarios of different cure rates, switching proportions, and TE assumptions. In cases of small sample sizes, the estimation bias increases but is still acceptable (less than 10%). For simple treatment switching adjustment methods, the estimation bias of β_{24} is much larger. Besides, the bias is positive in most scenarios. That is, the protective treatment effect of the experimental therapy is underestimated. On the one hand, the treatment switching after PD in the control group narrows the death hazard gap between groups. On the other hand, the presence of the cured patients dilutes the treatment effect of the experimental therapy, since the experimental therapy has no effect on the death hazard of cured patients. In comparison with the left column, the estimation biases of simple treatment switching adjustment methods are much smaller in the right column in Figure 3. Because these methods measure the average treatment effect in essence, which is consistent with the true value in the right column. Nevertheless, the performances of ITT and PPcen methods are not satisfactory in the right column. When estimating the average treatment effect with simple treatment switching adjustment methods, the cure rates and switching proportion have large impacts on the estimation performance (see Figure 3 A2 and B2). When the cure rate in the experimental group increases, the estimation bias for the average treatment effect decreases. Because the average

death hazard of patients in the experimental group is decreased by the increasing proportion of the cured, the estimated treatment effect is increasing naturally. The estimation bias increases with the switching proportion, especially for ITT, PPcen, and PPexc methods.

The details of the estimation bias variations with the cure rate, switching proportion, TE assumption, and sample size for simple treatment switching adjustment methods and proposed multistate transition model are as follows.

The simulations in Figure 3 A1 are carried out under the scenario of N=2000, CTE assumption satisfied, and switching proportion of 30%. It is observed that the estimation bias of the proposed multistate transition model is consistently small (less than 2%) in different cure rate scenarios. For simple treatment switching adjustment methods, the estimation bias of β_{24} changes greatly with cure rates. In the scenario of 15% vs 15% (i.e., the cure rates are equal in the experimental and control groups), the estimation biases of β_{24} based on simple treatment switching adjustment methods are extremely large. Because the death hazard of cured patients is independent of the received therapy, the equal proportions of the cured in experimental and control groups dilutes the treatment effect of the experimental therapy. The treatment effect is greatly underestimated. In scenarios of 30% vs 15% and 45% vs 15%, the cure rate in the experimental group increases, the treatment effect underestimation is gradually eased and the estimation bias of β_{24} is decreased. Under the scenario of 45% vs 15%, influenced by the large cure rate in the experimental group, the treatment effect is overestimated for PPexc, GE, IPE, and TSE methods.

As shown in Figure 3 B1, with the increase of the switching proportion, the estimation bias of all methods increases. The bias of the proposed multistate transition model is much

smaller than that of the simple treatment switching adjustment methods. With a sample size of 2000, the bias of the former is less than 1% while the biases of the latter methods are larger than 20%. The estimation biases of ITT and PPcen methods are more sensitive to the switching proportion.

In Figure 3 C1, the CTE assumption is satisfied in scenario 1 (i.e., the treatment effect on OS is a constant before and after progression across all uncured patients) while the CTE assumption is not satisfied in the other two scenarios. It is observed that the proposed multistate transition model performs well in all scenarios. The estimation bias of simple treatment switching adjustment methods increases in scenario 2. Because in scenario 2, the treatment effect in progressed patients and switchers is smaller, the average treatment effect decreases. Then the treatment effect underestimation is exacerbated. In comparison with scenario 1, the treatment effect in switchers is smaller in scenario 3, i.e., the switchers benefit less in comparison with the patients who have been randomized to the experimental group from the beginning. The estimation biases of ITT, GE, IPE, and TSE methods decrease in scenario 3. Because the decreased switching effect attenuates the impact of treatment switching on OS. Besides, we find that the estimation biases of PPcen and PPexc methods are unchanged in scenario 3 compared to that in scenario 1. That is because the observation after treatment switching is excluded in analysis, the switching effect does not impact the estimation of these two methods.

Figure 3 D1 shows the estimation bias versus the sample size. It is observed that the proposed multistate transition model is more sensitive to the sample size. Because there are much more parameters to estimate in the multistate transition model. The estimation bias

increases with the decrease of sample size. By contrast, the performance of simple treatment switching adjustment methods is less affected by the sample size. Nevertheless, the estimation bias of the proposed multistate transition model is smaller than that of simple treatment switching adjustment methods in cases of small sample sizes.

Case study

Diffuse large B-cell lymphoma (DLBCL) is a curable lymphoma [33]. The standard treatment for DLBCL consists of cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP) chemotherapy with anti-CD20 antibody rituximab (R) [34]. It has been proved that about 45% of DLBCL patients would be cured with R-CHOP treatments [35]. However, patients with treatment failure after R-CHOP often have a poor outcome [36]. Therefore, it is reasonable to allow patients who progressed with R-CHOP treatment to switch onto the experimental arm in RCT testing the effectiveness of a new treatment for DLBCL.

Assume a new treatment for DLBCL with the cure rate of 50% and the HR of 0.6 in comparison with R-CHOP treatment, we generated simulation datasets of the RCT comparing the new treatment and R-CHOP. Patients who are progressed in R-CHOP arm are allowed to switch onto the new treatment arm in the simulated RCT. The parameter setting of the simulation data is based on Coiffer's research [35]. More details are elaborated in Appendix 2 of the Additional file. With the switching proportion of 10%, 30%, and 50%, the sample size of 2000, 1000, and 500, the bias of the HR estimates of the new treatment obtained by the proposed multistate transitional model, Zeng's semi-competing risks model, ITT, PPcen, PPexc, GE, IPE, and TSE methods are shown in Figure 4. It shows that based on one dataset, the estimation bias of the proposed model is much smaller in comparison with the other methods. With the increase

of the switching proportion, the estimation bias of contrast methods increases, while the estimation bias of the proposed model is robust. Under the scenario of *N*=1000, the estimation bias of the proposed model, as well as that of contrast methods, is larger, it could be attributed to the sample error of the simulated dataset.

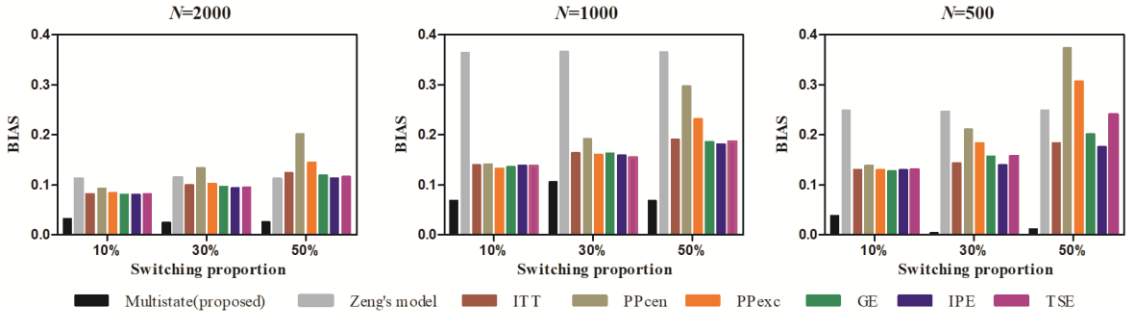


Figure 4. HR estimation bias of the proposed model and contrast methods based on the DLBCL dataset.

Discussion

To estimate the treatment effect of experimental therapy in RCTs with treatment switching permit in the presence of a cured subgroup, we propose a new multistate transition model in which the disease cure, progression, treatment switching, and death are accommodated. The proposed model hopes to separately quantify the treatment effect on the cure rate, progression hazard, and death hazard with or without progression. There are three trajectories to death in the model, including death for the cured to whom the progression would never happen, death without progression for the uncured, and death after progression for the uncured. The experimental effects on different trajectories are estimated separately in the model. Simulation studies under various scenarios show the good estimation performance and robustness of the proposed model. The results of comparisons with other methods illustrate the superiority of the

proposed model.

As a multistate model, Zeng's semi-competing risks model is included in the simulation as a control. The main difference between the proposed method and Zeng's method is that the cured fraction and patient-specific frailty are not considered in Zeng's method. Simulation results show that failing to incorporate the cured fraction in the model leads to large estimation biases. The effect estimation based on the proposed model without considering shared frailty is also biased. Therefore, when there is a cured subgroup, the proposed model is preferable. The shared frailty connects the progression and death times and describes the patient heterogeneity at the same time, which makes the model more scientific, rigorous, and rational.

Different from the proposed multistate transition model, some existing methods, including ITT, PPcen, PPexc, GE, IPE, and TSE methods, ignore the multiple trajectories to death in patients. These methods assume that the treatment effect keeps the same in all patients, and the possible change of the treatment effect after progression is also neglected. We call them simple treatment switching adjustment methods in this paper. As shown in the simulation study, the treatment effect estimates of these methods are more close to the effect averaged over all patients and all disease stages. When aiming to estimate the treatment effect on the hazard of one specific transition, for example, the death hazard in patients without progression, simple treatment switching adjustment methods bring about large biases. The bias varies with the cure rate, switching proportion, treatment effect, and switching effect. Therefore, these methods are incapable of providing specific treatment effect estimations on transition hazards between states. Instead, they provide average treatment effect estimation, which could not meet the needs of clinical practice in some cases. Besides, among these simple treatment switching adjustment

methods, the PPcen method produces the largest bias when estimating the average treatment effect. The GE, IPE, and TSE methods perform better in adjusting the treatment switching and estimating the treatment effect on OS. Many studies have been conducted in the literature to compare these methods, details are not elaborated here since it is not the core of this paper. Readers interested are referred to [2, 9, 11-14].

The multistate transition model proposed in this paper accommodates the possible cured fraction of patients innovatively, which makes the model more applicable. The shared frailty modeled by gamma distribution accounts for the individual heterogeneity of patients and dependence of progression and death times, which improves the generality of the proposed model. The limitation of the study is that the possible dropout or loss to follow-up and covariates that might influence the progression and death hazard have not been considered in the model, the extension on these aspects is under consideration.

Conclusions

The superiority and robustness of the proposed multistate transition model qualify it to estimate the treatment effect in trials with the treatment switching permit after progression and a cured subgroup.

List of abbreviations

AF: accelerated factor; AFT: accelerated failure time; aGVHD: acute graft-versus-host disease; CTE: common treatment effect; DLBCL: diffuse large B-cell lymphoma; EM: expectation-maximization; GE: grid estimation; HR: hazard ratio; HTA: health technology assessment; IPCW: inverse probability censoring weights; IPE: iterative parameter estimation; ITT: intention-to-treat; MSE: mean squared error; OS: overall survival; PB: percentage bias; PD:

progressive disease; PFS: progression-free survival; PP: per-protocol; PSO: particle swarm optimization; R-CHOP: rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone; RCT: randomized controlled trial; RPSFTM: rank preserving structural failure time model; SE: standard error; TE: treatment effect.

Declarations

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of Data and Materials: Parameter settings for the simulated DLBCL dataset used in the case study are publicly available in Additional file.

Competing interests: The authors declare that they have no competing interests.

Funding: This work was supported by the National Nature Science Foundation of China (grant no. 81773553, 81973141, 81803328).

Authors' contributions: HH and JX proposed the conception and designed the work. HH and LW performed the data analysis and drafted the manuscript. CL and WG participated in the results interpretation and manuscript revision. All the authors reviewed and approved the final manuscript.

Additional files

Additional file.pdf contains the simulation results of the proposed multistate transitional model and Zeng's semi-competing risks model under other scenarios (Appendix 1) and the parameter setting consideration for the simulated DLBCL dataset used in the case study (Appendix 2).

Acknowledgement

The authors would like to thank Dr. Xueqian Wang in Tsinghua University for his helpful

suggestions which significantly improved the quality of the paper.

References

1. Yeh J, Gupta S, Patel SJ, Kota V and Guddati AK. Trends in the crossover of patients in phase III oncology clinical trials in the USA. eCancerMedicalScience. 2020;14:1142.
2. Latimer NR, Abrams KR, Lambert PC, Crowther MJ, Wailoo AJ, Morden JP, et al. Adjusting survival time estimates to sccount for treatment switching in randomized controlled trials – an economic evaluation context. Med Decis Making. 2014;34:387-402.
3. Motzer RJ, Escudier B, Oudard S, Hutson TE, Porta C, Bracarda S, et al. Efficacy of everolimus in advanced renal cell carcinoma: a double-blind, randomised, placebo-controlled phase III trial. the Lancet. 2008;372:449-56.
4. Latimer NR. Treatment switching in oncology trials and the acceptability of adjustment methods. Expert Rev Pharm Out. 2015;15:561-4.
5. Robins JM and Tsiatis AA. Correcting for non-compliance in randomized trials using rank preserving structural failure time models. Commun Stat-Thero M. 1991;20:2609-31.
6. White IR, Babiker AG, Walker S and Darbyshire JH. Randomization-based methods for correcting for treatment changes: examples from the Concorde trial. Stat Med. 1999;18:2617-34.
7. Branson M and Whitehead J. Estimating a treatment effect in survival studies in which patients switch treatment. Stat Med. 2002;21:2449-63.
8. Shao J, Chang M and Chow S. Statistical inference for cancer trials with treatment switching. Stat Med. 2005;24:1783-90.
9. Latimer NR, Abrams KR, Lambert PC, Crowther MJ, Wailoo AJ, Morden JP, et al.

595 Adjusting for treatment switching in randomised controlled trials − A simulation study and
596 a simplified two-stage method. Stat Methods Med Res. 2014;26:724-51.
597 10. Robins JM and Finkelstein DM. Correcting for noncompliance and dependent censoring in
598 an aids clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests.
599 Biometrics. 2000;56:779-88.
600 11. Morden JP, Lambert PC, Latimer N, Abrams KR and Wailoo AJ. Assessing methods for
601 dealing with treatment switching in randomised controlled trials: a simulation study. BMC
602 Med Res Methodol. 2011;11:4.
603 12. White I. Methods for handling treatment switching: rank-preserving structural nested
604 failure time models, inverse-probability-of-censoring weighting, and marginal structural
605 models. Report, MRC Biostatistics Unit, UK, February 2012.
606 13. Watkins C, Huang X, Latimer N, Tang Y and Wright EJ. Adjusting overall survival for
607 treatment switches: commonly used methods and practical application. Pharm Stat.
608 2013;12:348-57.
609 14. Latimer NR, Abrams KR, Lambert PC, Morden JP and Crowther MJ. Assessing methods
610 for dealing with treatment switching in clinical trials: A follow-up simulation study. Stat
611 Methods Med Res. 2016;27:765-84.
612 15. Montané L, Ferranti D, Chabaud S, Pujade-Lauraine E, Pérol D and Paoletti X. Assessment
613 of post-switch endpoints in randomized clinical trials. Revue d'Épidémiologie et de Santé
614 Publique. 2018;66:S130.
615 16. Zeng D, Chen Q, Chen MH and Ibrahim JG. Estimating treatment effects with treatment
616 switching via semicompeting risks models: an application to a colorectal cancer study.

617 Biometrika. 2012;99:167-84.
618 17. Zhang Y, Chen M, Ibrahim JG, Zeng D, Chen Q, Pan Z, et al. Bayesian gamma frailty
619 models for survival data with semi-competing risks and treatment switching. Lifetime Data
620 Anal. 2014;20:76-105.
621 18. Huang CH, Chen YH, Wang JL and Wang M. Semiparametric copula-based analysis for
622 treatment effects in the presence of treatment switching. Stat Med. 2020;39:2936-48.
623 19. Chen Q, Zhang F, Chen M and Cong XJ. Estimation of treatment effects and model
624 diagnostics with two-way time-varying treatment switching: an application to a head and
625 neck study. Lifetime Data Anal. 2020;26:685-707.
626 20. Chen Q, Zeng D, Ibrahim JG, Akacha M and Schmidli H. Estimating time-varying effects
627 for overdispersed recurrent events data with treatment switching. Biometrika.
628 2013;100:339-54.
629 21. Jia X, Sima CS, Brennan MF and Panageas KS. Cure models for the analysis of time-to-
630 event data in cancer studies. J Surg Oncol. 2013;108:342-7.
631 22. Lee C, Lee SJ and Haneuse S. Time-to-event analysis when the event is defined on a finite
632 time interval. Stat Methods Med Res. 2020;29:1573-91.
633 23. Zhang F, Chen M, Cong XJ and Chen Q. Assessing importance of biomarkers: A Bayesian
634 joint modelling approach of longitudinal and survival data with semi-competing risks. Stat
635 Model. 2021;21:30-55.
636 24. Henshall C, Latimer NR, Sansom L and Ward RL. Treatment switching in cancer trials:
637 issues and proposals. Int J Technol Assess. 2016;32:167-74.
638 25. Xu J, Kalbfleisch JD and Tai B. Statistical analysis of illness-death processes and

semicompeting risks data. Biometrics. 2010;66:716-25.

26. Kennedy J and Eberhart R. Particle Swarm Optimization. Proceedings of ICNN'95-international conference on neural networks. IEEE. 1995;4:1942-8.

27. Abdel-Basset M, Fakhry AE, El-henawy I, Qiu T and Sangaiah AK. Feature and Intensity Based Medical Image Registration Using Particle Swarm Optimization. J Med Syst. 2017;41:197.

28. Patibandla RSML, Rao BT, Krishna PS and Maddumala VR. Medical data clustering using particle swarm optimization method. Journal of Critical Reviews. 2020;7:363-7.

29. Beheshti Z, Shamsuddin SMH, Beheshti E and Yuhaniz SS. Enhancement of artificial neural network learning using centripetal accelerated particle swarm optimization for medical diseases diagnosis. Soft Comput. 2014;18:2253-70.

30. Amrane Y and Kouba NEY. A Multiobjective optimal VAR dispatch using FACTS devices considering voltage stability and contingency analysis. In: Deo R, Samui P and Roy SS, editors. Predictive Modelling for Energy Management and Power Systems Engineering. Amsterdam: Elsevier; 2021. pp. 1-26.

31. Bansal J, Singh P, Saraswat M, Verma A, Jadon SS and Abraham A. Inertia Weight Strategies in Particle Swarm Optimization. In: Third World Congress on Nature and Biologically Inspired Computing, Salamanca, Spain, 19-21 Oct 2011, pp. 633-640. New York: IEEE.

32. Latimer NR, White IR, Abrams KR and Siebert U. Causal inference for long-term survival in randomised trials with treatment switching: Should re-censoring be applied when estimating counterfactual survival times? Stat Methods Med Res. 2018;28:2475-93.

33. Friedberg JW and Fisher RI. Diffuse Large B-Cell Lymphoma. Hematol Oncol Clin N. 2008;22:941-52.
34. Goda JS, Lewis SC, Laskar S, Kannan S, Khanna N, Jain H, et al. A phase III randomized controlled trial of radiation dose optimization in non-Hodgkin lymphoma-diffuse large B-cell lymphoma (DOBL study): Study protocol and design. Cancer reports. 2019;2:e1161.
35. Coiffier B, Lepage E, Briere J, Herbrecht R, Tilly H, Bouabdallah R, et al. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. N Engl J Med. 2002;346:235-42.
36. Sehn LH, Salles G. Diffuse Large B-Cell Lymphoma. N Engl J Med. 2021;384:842-58.

Figure Legends

Figure 1. Proposed multistate transition model in which the cured fraction (bold) is innovatively considered.

Figure 2. Flow chart of the PSO algorithm.

Figure 3. PB(%) of the estimate using the proposed model and contrast methods under different scenarios.

Figure 4. HR estimation bias of the proposed model and contrast methods based on the DLBCL dataset.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile.pdf