

# Inverse Probability Weighted Estimating Equations for Randomized Trials in Transfusion Medicine

RICHARD J. COOK

*Department of Statistics and Actuarial Science,  
University of Waterloo, Waterloo, ON, N2L 3G1, Canada  
E-mail: rjcook@uwaterloo.ca*

KER-AI LEE

*Department of Statistics and Actuarial Science,  
University of Waterloo, Waterloo, ON, N2L 3G1, Canada*

MEAGHAN CUERDEN

*Department of Statistics and Actuarial Science,  
University of Waterloo, Waterloo, ON, N2L 3G1, Canada*

CECILIA A. COTTON

*Department of Statistics and Actuarial Science,  
University of Waterloo, Waterloo, ON, N2L 3G1, Canada*

## Summary

Thrombocytopenia is a condition characterized by extremely low platelet counts, which puts patients at elevated risk of morbidity and mortality because of bleeding. Trials in transfusion medicine are routinely designed to assess the effect of experimental platelet products on patients platelet counts. In such trials, patients may receive multiple platelet transfusions over a predefined period of treatment, and a response is available from each such administration. The resulting data comprised multiple responses per patient, and although it is natural to want to use this data in testing for treatment effects, naive analyses of the multiple responses can yield biased estimates of the probability of response and associated treatment effects. These biases arise because only subsets of the patients randomized contribute response data on the second and subsequent administrations of therapy and the balance between treatment groups with respect to potential confounding factors is lost. We discuss the design and analysis issues involved in this setting and make recommendations for the design of future platelet transfusion trials.

*Keywords:* causation, confounder, longitudinal study, marginal model, randomized trial

**This is the peer reviewed version of the following article: Cook, R. J., Lee, K.-A., Cuerden, M. and Cotton, C. A. (2013), Inverse probability weighted estimating equations for randomized trials in transfusion medicine. *Statist. Med.*, 32: 4380–4399. doi:10.1002/sim.5827, which has been published in final form at <http://dx.doi.org/10.1002/sim.5827>. This article may be used for non-commercial purposes in accordance With Wiley Terms and Conditions for Self-Archiving: <http://olabout.wiley.com/WileyCDA/Section/id-820227.html#terms>.**

## 1 INTRODUCTION

### 1.1 PLATELET TRANSFUSION TRIALS IN THROMBOCYTOPENIA

Thrombocytopenia is a chronic medical condition characterized by abnormally low platelet counts in which patients are at significant risk of morbidity and mortality due to bleeding [1]. The causes of thrombocytopenia are multifaceted and include immunological abnormalities [2], treatment with anti-coagulants in orthopedic surgery patients [3], infection [4], and exposure to certain medications such as methotrexate [5]. It also arises in patients with cancer or hematologic disorders as a consequence to chemotherapy [6]. Treatment often includes prednisone, infusions of intravenous immunoglobulin, and platelet transfusion [7]. Platelet transfusion, however, can lead to antibody responses in recipients, and as a result, there has been considerable research directed at the development of pathogen-inactivated platelet products to reduce the risks of these reactions [8–13]. This article is concerned with methodological challenges arising in the design and analysis of randomized trials directed at comparing the effectiveness of standard and pathogen inactivated platelets.

While the ultimate goal of platelet transfusion is to reduce the risk of bleeding, hemorrhagic events are fairly rare and there are challenges in how best to summarize bleeding outcome data and make associated treatment comparisons [14–16]. As a result, trials typically examine the effect of a platelet transfusion in terms of the change in the recipients platelet count before and after the transfusion, called the count increment. When this difference is standardized with respect to body mass and the dose of platelets transfused, the resulting measure is called the corrected count increment (CCI), and this is perhaps the most common response adopted in randomized trials of platelet products [17]. The post-transfusion platelet counts are measured at different times in different studies, but in European studies, blood samples are often taken 1 h after the transfusion, in which case, the associated CCI reflects the short term effect of the transfusion. The CCI can be used as a continuous response or dichotomized; 1 h CCI values above  $7500 (\times 10^6 m^2/\ell)$  represent “successful” transfusions.

The timing and need for transfusions vary considerably between patients because of differences in the severity of the underlying condition, the schedule and intensity of chemotherapy, and other known and unknown factors. Protocols often dictate that data on the responses to a pre-specified number of transfusions are to be analysed [9]. If all patients are followed until they experience the pre-specified number of transfusions, then this is a reasonable strategy. Under suitable assumptions, there are obvious efficiency gains from using data beyond the first transfusion, and data over the broader course of treatment may be viewed as more reflective of the overall burden of disease and benefit of transfusion therapy. In most settings however, there is a fixed period of treatment (e.g., 4 weeks) and the number of transfusions required over this time period will be random and may be far lower than the prespecified number. This is further complicated by the fact that the number of transfusions administered in a fixed window of time is potentially related to the platelet product to which patients have been assigned, the response to the intervention, and observed and unobserved covariates. In this case, standard use of data beyond the first administration is problematic because the balance in the distribution of potential confounding variables between treatment arms is lost for second and subsequent transfusions. In essence, marginal treatment comparisons based on the second administration of platelet therapy are based only on those patients requiring a second administration and thus constitutes an improper subgroup comparison [18].

In this article, we discuss particular models for the transfusion process, and study the biases resulting from standard longitudinal analyses. We discuss ways to analyse the data from such a study, which permit causal inference using a simple marginal model, investigate the performance of the methods through simulation studies, and illustrate their use through application to data from a motivating trial. We organized the remainder of this paper as follows. In Section 2, we define notation and formulate models for the need for transfusions and the response to transfusions. In Section 3, we highlight the effect of confounding in the second and subsequent transfusions by deriving limiting

values of parameters obtained from naive analyses. In Section 4, we conduct simulation studies to point out how marginal models can be fitted using weights to adjust for the effect of confounders arising, despite the initial randomization, from the finite period of observation. We give an application to a recent platelet transfusion trial for illustration. We make general remarks in Section 5.

## 1.2 THE MIRASOL TRIAL OF PATHOGEN-INACTIVATED PLATELETS

Cazenave et al. [11] report on a recent multicentre trial of 118 hematology/oncology patients with chemotherapy-induced thrombocytopenia who were randomized to receive either pathogen-inactivated platelets (Mirasol) or standard platelets as required over a 28 day treatment period. The primary outcome in this trial was the CCI with the posttransfusion platelet count measured 1 hour after transfusion. The protocol stipulated that CCI responses from a maximum of eight transfusions occurring over the 28 day treatment period would be used for each patient in the assessment of the experimental and reference products. The primary analysis is based on comparing the probability of successful response between treatment arms under the implicit assumption that this probability does not change over time within treatment groups.

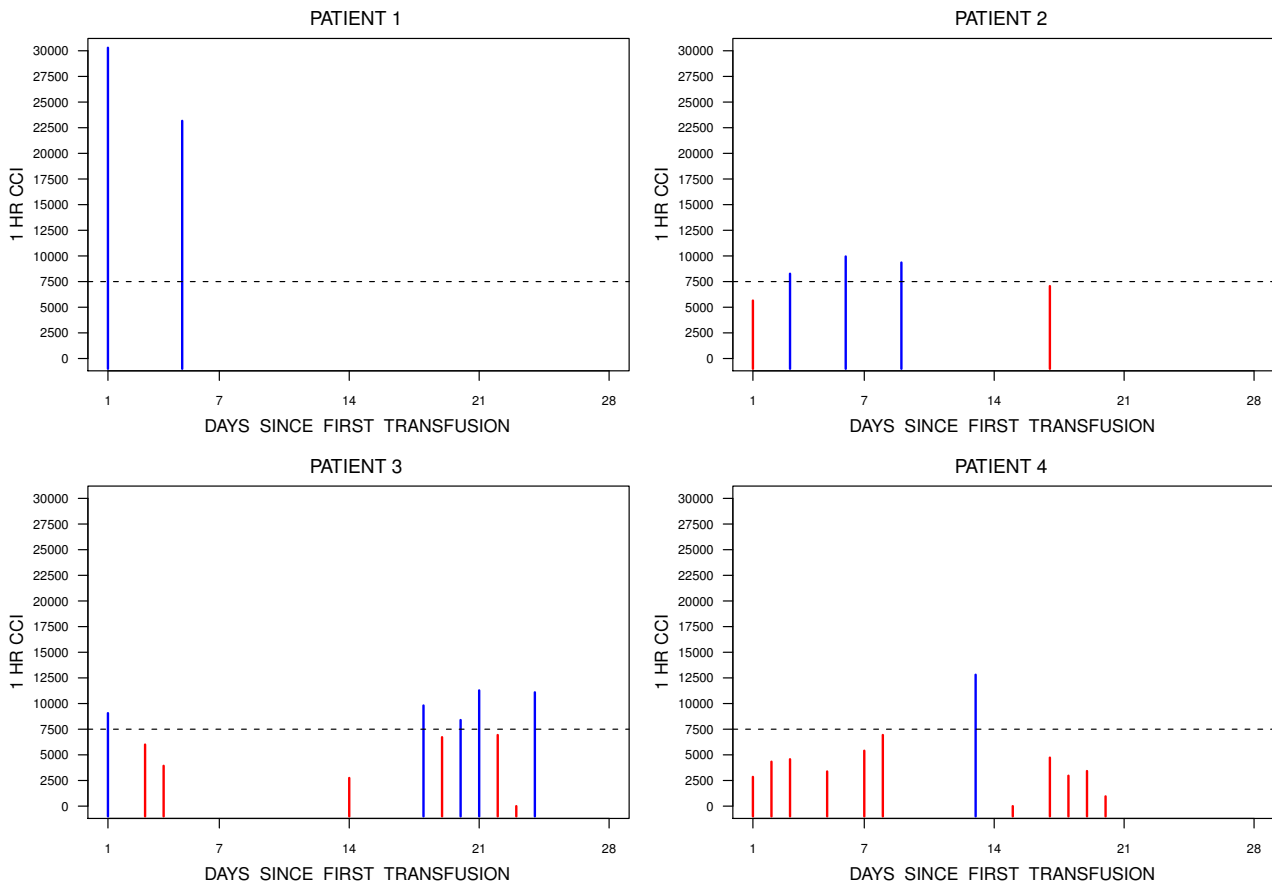


Figure 1: Sample timeline plots of platelet transfusions and associated 1 h corrected count increments for four individuals from the Mirasol study [11] ; the horizontal line at 7500 is the threshold for defining successful responses to transfusion.

Cazenave et al. [19] give the rationale for restricting attention to eight transfusions, and this is now common practice in such studies. In examining data from an earlier platelet transfusion trial, when considering all transfusions received, there was considerable heterogeneity in the mean CCI between patients such that there was a tendency for the mean CCI response per transfusion to be higher among

patients who had a smaller number of transfusions. Restricting the number of transfusions used in the analysis to a maximum of eight was put forward as one method for reducing the heterogeneity. It is important, however, to consider the source of this heterogeneity. Poor response (low CCI) to a transfusion means a patient's platelet count remained low and they are therefore at relatively higher risk of needing subsequent transfusions. In contrast, if a patient has a large CCI from a particular transfusion, there will likely be a longer time to the next transfusion. This could explain the association between the mean CCI per patient and the number of transfusions they require in any given treatment period.

Figure 1 displays data for a sample of four patients from the Mirasol study. The vertical axis is the CCI, and the horizontal axis indicates the number of days since randomization. The height of the vertical lines reflect the 1 h CCI, and the horizontal dashed line at 7500 is the threshold defining successful (CCI > 7500) or unsuccessful (CCI ≤ 7500) transfusions. Patient 1 required only two transfusions during the 28 day treatment period, and both led to large increases in the platelet counts satisfying the definition of a successful response. Patient 4, on the other hand, required twelve transfusions, and only one of these led to a successful response. Patients 2 and 3 had intermediate numbers of transfusions, which were of mixed success.

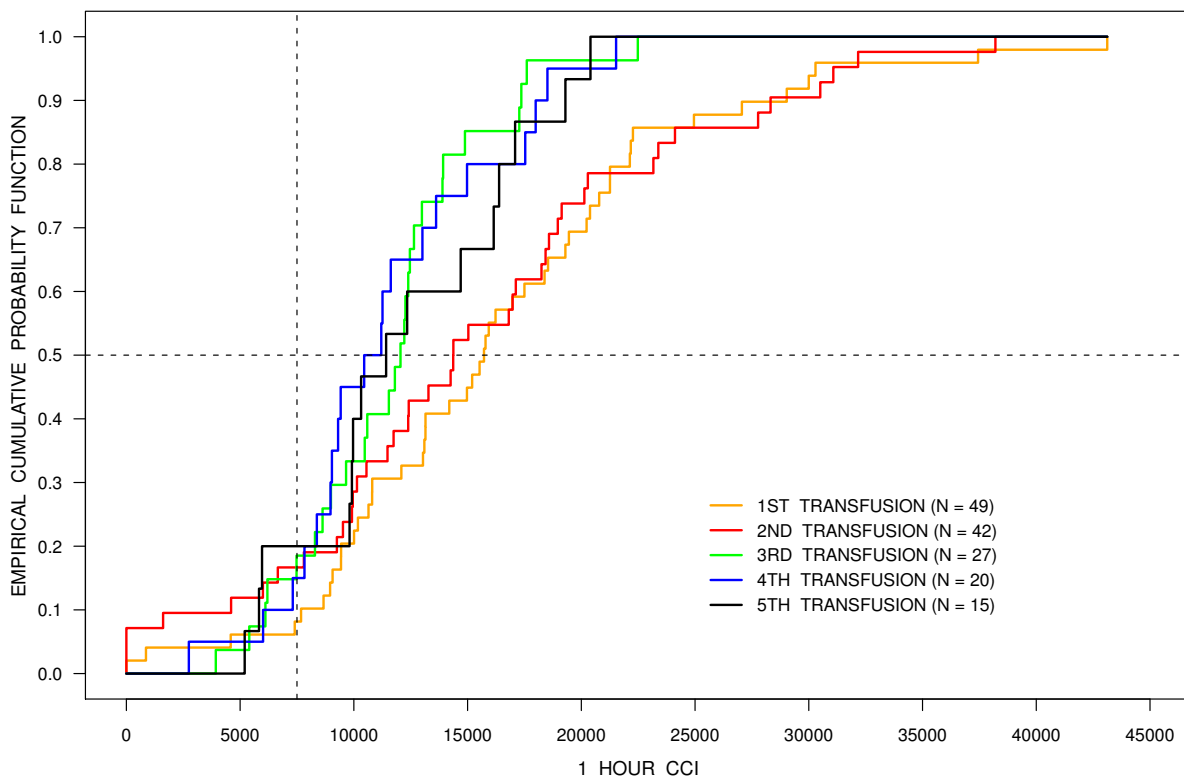


Figure 2: Plots of the empirical cumulative distribution functions for the 1 h corrected (CCI) for the first five transfusions for individuals in the control group (the number of subjects contributing to these estimates are indicated in the legend).

As stipulated in the protocol, secondary analyses were directed at the assessment of trend in the probability of response by transfusion number. There are difficulties in assessing success rates by transfusion number because of selection effects arising from the finite treatment period. Figure 2 which contains the empirical distribution functions for the 1 hour CCI for the first five transfusions among all relevant patients in the control arm, illustrates an apparent trend. A naive interpretation of the plots suggest a trend toward a decreasing median CCI with increasing transfusion number. The

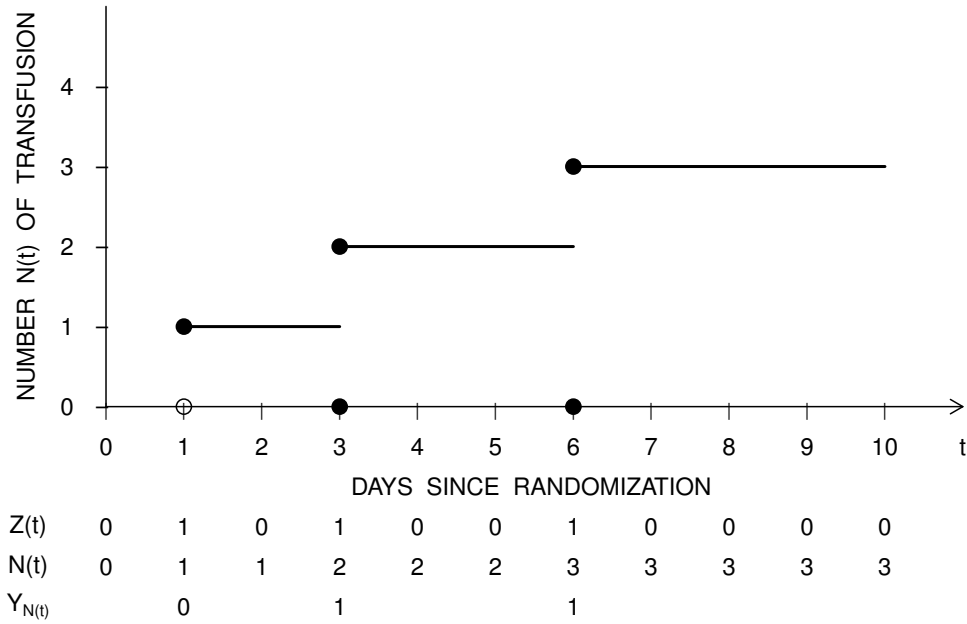


Figure 3: Schematic illustrating the relationship between the different post-randomization variables for a hypothetical individual; on the horizontal axis, open and closed circles correspond to days of unsuccessful and successful platelet transfusions, respectively.

vertical line at 7500 can be used to estimate the empirical success rates, which are not inconsistent with a decreasing success rate with increasing transfusion number. The legend indicates the decreasing number of patients contributing to these estimates, which suggests the possibility of a selection effect driving the apparent trend. We aim to describe strategies for removing this potentially spurious signal of trend, after first formulating a model, which can be used to mimic this phenomenon.

## 2 NOTATION AND MODEL FORMULATION

### 2.1 NOTATION

Let  $X = 1$  if an individual has been randomized to receive the experimental platelet product and  $X = 0$  otherwise, where  $P(X = 1) = 0.5$ . For simplicity, we consider a single additional binary covariate  $V$ , which may be a confounder or effect modifier; we let  $V = 1$  indicate the presence of a factor and  $V = 0$  otherwise, and let  $P(V = 1) = \text{expit}(\nu)$ . We focus on a discrete time stochastic process  $\{Z(t), 0 < t\}$  for the need for transfusion, where time  $t$  is the number of days since randomization; the variable  $Z(t) = 1$  if an individual receives a transfusion on day  $t$  and is zero otherwise. The history for the transfusion process at time  $t$  is  $\bar{Z}(t) = \{Z(1), \dots, Z(t-1)\}$ , which records the number and timing of transfusions up to and including day  $t-1$ . Additional variables related to  $\{Z(s), 0 < s\}$  include  $N(t)$ , which records the cumulative number of transfusions over  $(1, t]$ , and  $T_k$ , which is the day of the  $k$ th transfusions,  $k = 1, \dots$ . Upon receiving the  $k$ th transfusion, we may classify a response as successful, indicated by  $Y_k = 1$ , or unsuccessful, indicated by  $Y_k = 0$ . We denote the history for the outcomes of all transfusions prior to time  $t$  by  $\bar{Y}(t) = \{Y_1, \dots, Y_{N(t-1)}\}$  and we let  $\bar{Y}_k = \{Y_1, \dots, Y_{k-1}\}$  denote the responses to the first  $k-1$  transfusions. We denote the full history at time  $t$  is denoted  $\bar{H}(t) = \{\bar{Y}(t), \bar{Z}(t), X, V\}$ , which includes information on the fixed covariates. Figure 3 contains a schematic diagram demonstrating the meaning of the variables related to the transfusion process, where open and closed circles on the horizontal axis indicate unsuccessful and successful platelet transfusions, respectively.

## 2.2 A MODEL FOR THE TRANSFUSION PROCESS

At the time of the first transfusion, we randomize individuals to a treatment arm, which dictates the platelet product that will be used for the initial and subsequent transfusions. Because all individuals have transfusions on day 1,  $P(Z(1) = 1) = 1$ , but for subsequent days, we require a probability model. We adopt a model on the basis of the discrete intensity  $\lambda(t|\bar{H}(t)) = P(Z(t) = 1|\bar{H}(t))$ ,  $1 < t$ , which is the conditional probability of requiring a transfusion on day  $t$  given the full history. The key features of the history governing the need for transfusions could include, for example, the risk factor  $V$ , whether a transfusion was given on the previous day ( $Z(t-1)$ ), platelet product ( $X$ ), and whether the most recent transfusion was successful ( $Y_{N(t-1)}$ ). In this case, we could write the intensity in the form of a logistic regression model as

$$\lambda(t|\bar{H}(t)) = \text{expit}(\zeta_0 + \zeta_1 V + \zeta_2 Z(t-1) + \zeta_3 X + \zeta_4 Y_{N(t-1)}) . \quad (2.1)$$

Other formulations are of course possible, and we can easily incorporate a dependence on more aspects of the process history.

## 2.3 A MODEL FOR THE RESPONSE PROCESS

Let  $K$  denote the maximum number of transfusions to be analysed in the primary analyses, and let  $Y = (Y_1, \dots, Y_K)'$  denote the multivariate response vector. Clinical interest lies in assessing the effect of the platelet product on the probability of a successful transfusion, and we adopt a model with a marginal mean, which does not depend on the stage of the transfusion process to ensure it is compatible with this objective. Thus, given  $X$  and  $V$ , we take  $Y$  to be a vector of Bernoulli variates with mean  $\pi(X, V) = (\pi_1(X, V), \dots, \pi_K(X, V))$ , where  $\pi_k(X, V) = E(Y_k|X, V)$ . We may let

$$\pi_k(X, V) = \text{expit}(\vartheta_0 + \vartheta_1 X + \vartheta_2 V + \vartheta_3 X V) , \quad (2.2)$$

for example, to reflect the setting in which individuals with different values of  $V$  will experience different effects of treatment. We let  $\text{COV}(Y|X, V) = \Sigma(X, V)$  be a  $K \times K$  matrix comprised variances

$$\text{VAR}(Y_k|X, V) = \pi_k(X, V)(1 - \pi_k(X, V))$$

along the diagonal and covariances

$$\text{COV}(Y_j, Y_k|X, V) = P(Y_j = 1, Y_k = 1|X, V) - \pi_j(X, V)\pi_k(X, V) , \quad (2.3)$$

in the off-diagonal entries. We denote the conditional odds ratio associating the responses from the  $j$ th and  $k$ th transfusions given  $X$  and  $V$  by

$$\varphi = \frac{P(Y_j = 1, Y_k = 1|X, V)/P(Y_j = 0, Y_k = 1|X, V)}{P(Y_j = 1, Y_k = 0|X, V)/P(Y_j = 0, Y_k = 0|X, V)} .$$

We assumed an exchangeable odds ratio here, but other structures are of course possible. We obtain the joint probability in (2.3) as

$$P(Y_j = 1, Y_k = 1|X, V) = \begin{cases} \frac{c(X, V) - [c(X, V)^2 - 4\varphi(\varphi-1)\pi_j(X, V)\pi_k(X, V)]^{1/2}}{2(\varphi-1)} & \text{if } \varphi \neq 1 \\ \pi_j(X, V)\pi_k(X, V) & \text{if } \varphi = 1 , \end{cases}$$

where  $c(X, V) = 1 - (1 - \varphi)(\pi_j(X, V) + \pi_k(X, V))$  [20].

## 2.4 MARGINAL MODELS FOR THE ASSESSMENT OF TREATMENT EFFECTS IN RANDOMIZED TRIALS

We introduce the variable  $V$  to capture variation in the need for transfusions in (2.1) and variation in the probability of a successful response to a transfusion in (2.2). Protocols of most randomized trials, however, typically dictate that analyses of treatment effects be based on simple comparisons in which  $V$  is omitted from the response model. If the marginal mean conditional on  $(X, V)$  is given by (2.2) for example, but we fit a model

$$P(Y_k = 1|X) = \text{expit}(\theta_0 + \theta_1 X) . \quad (2.4)$$

We can derive  $\theta$  in terms of  $\vartheta$  and  $\nu$  through

$$P(Y_k = 1|X; \theta) = E_{V|X} [P(Y_k = 1|X, V)] = \sum_{v=0}^1 P(Y_k = 1|X, V = v; \vartheta) P(V = v; \nu) .$$

Randomization at the point of recruitment justifies the use of  $P(V)$  rather than  $P(V|X)$  in the equation above. Because the true marginal probabilities are independent of  $k$ , the values of  $\theta = (\theta_0, \theta_1)'$  do not depend on  $k$ , and we simply have

$$\theta_0(\vartheta, \nu) = \text{logit}(P(Y_k = 1|X = 0; \vartheta, \nu)) \quad (2.5a)$$

$$\theta_1(\vartheta, \nu) = \text{logit}(P(Y_k = 1|X = 1; \vartheta, \nu)) - \theta_0(\vartheta, \nu) . \quad (2.5b)$$

In this framework the causal parameter of interest is  $\theta_1$ , which represents the log odds ratio for a successful response to transfusion with  $X = 1$  versus  $X = 0$ , averaged over  $V$ .

We can obtain the covariance matrix of  $Y|X$  on the basis of  $P(Y_j = 1, Y_k = 1|X) = E_{V|X}[P(Y_j = 1, Y_k = 1|V, X)]$ , which is

$$\sum_{v=0}^1 P(Y_j = 1, Y_k = 1|V = v, X; \vartheta, \varphi) P(V = v; \nu) .$$

Given this,

$$\text{COV}(Y_j, Y_k|X) = P(Y_j = 1, Y_k = 1|X) - P(Y_j = 1|X) P(Y_k = 1|X) \quad (2.6)$$

and we obtain an exchangeable association structure with the odds ratio associating  $Y_j$  and  $Y_k$  given only  $X$ , as

$$\phi(\vartheta, \varphi, \nu) = \frac{P(Y_j = 1, Y_k = 1|X)/P(Y_j = 0, Y_k = 1|X)}{P(Y_j = 1, Y_k = 0|X)/P(Y_j = 0, Y_k = 0|X)} . \quad (2.7)$$

It is important to note that these measures of association are different than those of Section 2.3 because of the omission of the effect modifier  $V$  from the response model. In the context of a randomized trial, one can legitimately aim to estimate (2.5) and (2.6) or (2.7), and this is routinely done when models are fully specified or assumptions are made about higher order moments in estimating functions.

## 3 ASYMPTOTIC BEHAVIOUR UNDER A DEPENDENT OBSERVATION SCHEME

### 3.1 PRELIMINARY DERIVATIONS FOR A COMPLETE AND AVAILABLE CASE ANALYSIS

The derivations in Section 2.4 gave the joint distribution of  $Y|X$ ; however, not all  $K$  transfusions are necessarily observed when individuals are followed for a fixed period of time. Given  $X$  and  $V$ , the

joint probability for the first  $K$  responses and the event that all are observed over  $[1, C]$  is

$$P(Y, T_K \leq C | X, V) = \sum_{T_K=K}^C \sum_{\bar{Z}(T_K)} P(Z(1) = 1, \dots, Z(T_K) = 1, Y_1, \dots, Y_K | X, V) \quad (3.1)$$

where  $\sum_{\bar{Z}(T_K)}$  means to sum up over all realizations  $Z(2), \dots, Z(T_K - 1)$  subject to the constraint  $\sum_{t=2}^{T_K-1} Z(t) = K - 2$ , where we omit the terms  $Z(1)$  and  $Z(T_K)$  because transfusions occur on day 1 with probability one and  $Z(T_K) = 1$  with probability one by definition. The joint probability in the summand of (3.1) is

$$\prod_{t=2}^{T_K} \left[ P(Z(t) | \bar{Y}(t), \bar{Z}(t), X, V) P(Y_{N(t)} | \bar{Y}(t), Z(t) = 1, \bar{Z}(t), X, V)^{Z(t)} \right] \\ \times P(Y_1 | Z(1) = 1, X, V) P(Z(1) = 1 | X, V)$$

which can be expressed as

$$\prod_{t=2}^{T_K} \left[ \lambda(t | \bar{H}(t))^{Z(t)} (1 - \lambda(t | \bar{H}(t)))^{1-Z(t)} \cdot (\mu_{N(t)}^{Y_{N(t)}} (1 - \mu_{N(t)})^{1-Y_{N(t)}})^{Z(t)} \right] \mu_1^{Y_1} (1 - \mu_1)^{1-Y_1}$$

where  $\mu_k(X, V) = P(Y_k = 1 | X, V, \bar{Y}_k)$  and we suppress the dependence of  $\mu_k(X, V)$  on  $X$  and  $V$  above; see Section 4.1 for details on the form of  $\mu_k(X, V)$ .

The probability that at least  $K$  transfusions are observed over  $[1, C]$  is then simply

$$P(T_K \leq C | X, V) = \sum_{y_1} \dots \sum_{y_K} P(Y_1 = y_1, \dots, Y_K = y_K, T_K \leq C | X, V), \quad (3.2)$$

and the conditional joint probability mass function for  $Y$  given they were observed over  $[1, C]$  is then

$$P(Y_1, \dots, Y_K | T_K \leq C, X, V) = \frac{P(Y_1, \dots, Y_K, T_K \leq C | X, V)}{\sum_{y_1} \dots \sum_{y_K} P(Y_1 = y_1, \dots, Y_K = y_K, T_K \leq C | X, V)}. \quad (3.3)$$

Because we do not wish to assess the treatment effect in a model controlling for  $V$ , we average over  $V$  as in Section 2. The difference here is that for the  $K$ th transfusion, we require the distribution of  $V$  given  $T_K \leq C$  and  $X$ , or equivalently  $P(V | T_K \leq C, X)$ . Thus, for the  $K$ th transfusion, we use

$$P(V | T_K \leq C, X) = \frac{P(T_K \leq C | X, V) P(V)}{\sum_{v=0}^1 P(T_K \leq C | X, V = v) P(V = v)} \quad (3.4)$$

where  $X$  and  $V$  are independent by randomization.

If analyses are restricted to individuals who experienced all  $K$  transfusions, then taking (3.3) and (3.4), we obtain the relevant joint distribution as

$$P(Y_1, \dots, Y_K | T_K \leq C, X) = E_{V | T_K \leq C, X} [P(Y_1, \dots, Y_K | T_K \leq C, X, V)] \quad (3.5)$$

from which we can obtain  $P(Y_k = 1 | T_K \leq C, X)$ ,  $k = 1, 2, \dots, K$ . Moreover,

$$\text{COV}(Y_j, Y_k | T_K \leq C, X) = P(Y_j = 1, Y_k = 1 | T_K \leq C, X) \\ - P(Y_j = 1 | T_K \leq C, X) P(Y_k = 1 | T_K \leq C, X),$$

for  $j < k \leq K$ , and the odds ratio associating  $Y_j$  and  $Y_k$  given  $T_K \leq C$  and  $X$  is then

$$\frac{P(Y_j = 1, Y_k = 1 | T_K \leq C, X) / P(Y_j = 0, Y_k = 1 | T_K \leq C, X)}{P(Y_j = 1, Y_k = 0 | T_K \leq C, X) / P(Y_j = 0, Y_k = 0 | T_K \leq C, X)}. \quad (3.6)$$



This odds ratio is estimable in likelihood analyses or analyses based on generalized estimating equations [21, 22].

A more common approach to analysis would be to use all available information from individuals such that individuals with  $K^\dagger \leq K$  transfusions observed over  $[1, C]$  would contribute data for the  $K^\dagger$  transfusions. In this case, for the  $k$ th transfusion ( $k \leq K^\dagger$ ), the marginal probability of success is

$$P(Y_k = 1|T_k \leq C, X) = E_{V|T_k \leq C, X} [P(Y_k = 1|T_k \leq C, V, X)] , \quad (3.7)$$

and so, it is dependent on  $k$ . For  $j < k \leq K^\dagger$ ,

$$\begin{aligned} \text{COV}(Y_j, Y_k|T_k \leq C, X) &= P(Y_j = 1, Y_k = 1|T_k \leq C, X) \\ &\quad - P(Y_j = 1|T_k \leq C, X)P(Y_k = 1|T_k \leq C, X) \end{aligned}$$

and the corresponding odds ratio is

$$\frac{P(Y_j = 1, Y_k = 1|T_k \leq C, X)/P(Y_j = 0, Y_k = 1|T_k \leq C, X)}{P(Y_j = 1, Y_k = 0|T_k \leq C, X)/P(Y_j = 0, Y_k = 0|T_k \leq C, X)} . \quad (3.8)$$

A key point here is that if the “true” odds ratio for the association conditional on  $X$  is given by (2.7), the exchangeable odds ratio (3.6) is being estimated in an analysis restricted to individuals with  $K$  transfusions. When available data are used, the odds ratio associating  $Y_j$  and  $Y_k$  ( $j < k$ ) depends on  $k$ , as suggested by (3.8), and most standard methods of analyses would not yield consistent estimates of this quantity or the underlying association. Finally, note that the derivations thus far have been based on the assumption that all individuals were observed over the entire treatment period  $(1, C]$ . When the duration of follow-up varies between individuals, the corresponding limiting values of estimators are more complicated to derive and will depend on the distribution of the duration of observation.

### 3.2 UNWEIGHTED AVAILABLE CASE ANALYSIS

If analyses are directed at assessing the probability of response and associated treatment effects by transfusion number, one could fit a binary model with the following estimating equation:

$$I(T_k \leq C)D'_k [\text{VAR}(Y_k|X)]^{-1} (Y_k - \pi_k(X)) \quad (3.9)$$

where  $\pi_k(X) = \text{expit}(\theta_{k0} + \theta_{k1}X)$  is the marginal mean for the  $k$ th transfusion given  $X$  only,  $D_k = d\pi_k(X)/d\theta_k$  is a  $2 \times 1$  vector where  $\theta_k = (\theta_{k0}, \theta_{k1})'$ , and  $\text{VAR}(Y_k|X) = \pi_k(X)(1 - \pi_k(X))$ . With complete observation of all  $K$  responses, the  $I(T_k \leq C)$  term is absent, and estimators of  $\theta_k$  for different values of  $k$  are all estimating  $\theta_0$  and  $\theta_1$  given in (2.5). The condition  $I(T_k \leq C)$  is implicitly present in the “available case” analysis for the  $k$ th transfusion. Thus with an unweighted available case analysis, we are estimating

$$\begin{aligned} \theta_{k0}^*(\zeta, \vartheta, \varphi, \nu) &= \text{logit}(P(Y_k = 1|T_k \leq C, X = 0)) \\ \theta_{k1}^*(\zeta, \vartheta, \varphi, \nu) &= \text{logit}(P(Y_k = 1|T_k \leq C, X = 1)) - \theta_{k0}^*(\zeta, \vartheta, \varphi, \nu) \end{aligned} \quad (3.10)$$

The limiting values in (3.10) are also the limiting values under a generalized estimating equation analysis with a working independence assumption and vectors of regression coefficients that vary by response number.

### 3.3 WEIGHTED ANALYSIS BASED ON A MARGINAL MODEL

Suppose we fit a logistic model with the following contribution to the estimating function from an individual who satisfies the condition that their  $k$ th transfusion occurs over  $[1, C]$ :

$$U_k(\theta_k; \alpha_k) = \frac{I(T_k \leq C)}{P(T_k \leq C|X, V; \alpha_k)} D'_k [\text{VAR}(Y_k|X)]^{-1} (Y_k - \pi_k(X)) \quad (3.11)$$

where  $\pi_k(X) = \text{expit}(\theta_{k0} + \theta_{k1}X)$  is the marginal mean for the  $k$ th transfusion given  $X$  only and  $P(T_k \leq C|X, V; \alpha_k)$  is the probability that this transfusion is observed given  $(X, V)$  is used to specify the unstabilized weights. The marginal response model is the kind of model we would fit if we were interested in assessing trends in response over time, or looking at how the effect of treatment varies by transfusion number. It is marginal in the sense that it does not involve conditioning on any variables other than the treatment variable, but the inverse probability weight ensures that consistent estimates are obtained. The term ‘‘marginal’’ is also appropriate because it may be viewed as based on one of the two potential outcomes in the framework of counterfactuals [23]. At the suggestion of a referee, in Appendix A, we recast this problem within the framework of counterfactual outcomes for readers acquainted with this theory and to justify our use of the term *marginal structural model* [23]. Although marginal structural models are often used to deal with dynamic response-dependent treatment selection [24–26], we show the model here to be simple, robust, and convenient to use. In practice of course, we must estimate the weights, but we obtain consistent estimates of  $\theta_k$  from (3.11) when  $\alpha_k$  is replaced with a  $\sqrt{n}$ -consistent estimate  $\hat{\alpha}_k$ . Appendix B outlines how one can estimate weights and compute suitable variance estimates.

The appeal of (3.11) is that the solution is consistent for  $\theta$  given by (2.5), so the use of the inverse probability weights eliminates any spurious evidence of a trend in the mean response or treatment effects over time. It has the usual form of an inverse probability weighted estimating function used in monotone missing longitudinal data with observation-specific weights proportional to the inverse of the probability; the observation is realized given observed data. A key difference is that in the present setting, the observations are not necessarily ‘‘missing’’, in the same sense they are in longitudinal studies with repeated measures designs, but rather could be observed if the period of observation were longer. It is apparent from the form of the weight in (3.11) that the weighting aims to create a pseudosample mimicking the sample at the time of randomization when  $X \perp V$ ; in equation (A.2), the term  $P(X|V, T_k < C)$  is analogous to the probability of exposure used to inverse weight in observation studies. We comment more on this in Section 5.

If an analysis is planned on the basis of generalized estimating equations, we may adopt a marginal mean response model  $\text{logit}(P(Y_k = 1|X)) = \theta_{k0} + \theta_{k1}X$ . We may allow the intercept to vary with transfusion number as well as the effect of treatment, which would be particularly appealing if interest was in the assessment of trend. Alternatively, a parsimonious way of assessing treatment effects is to constrain  $\theta_{k1} = \theta_1$ , so there is a single odds ratio for treatment. For generality, we let  $\theta$  denote the vector of parameters in the mean model, which here is of dimension  $p$ . We remark that it is unlikely that the correct covariance matrix would be specified given the dependence of the covariance on the number of transfusions and suppose a working independence assumption is made, in part for simplicity but also because there can be convergence issues when working covariance matrices are grossly misspecified [27, 28]. In this case, each individual would have a contribution to an estimating function of the form

$$U(\theta; \alpha) = \mathbb{D}' \mathbb{V}^{-1} \Delta (Y - \pi(X)) = 0$$

where  $\pi = \pi(X) = E(Y|X)$  is a  $K \times 1$  vector of marginal means with  $\pi(X) = (\pi_1(X), \dots, \pi_K(X))'$ ,  $\mathbb{V} = \text{diag}(\pi_k(1 - \pi_k), k = 1, \dots, K)$  is the  $K \times K$  covariance matrix,  $\mathbb{D} = d\pi(X)/d\theta$  is a  $K \times p$  matrix of derivatives,  $\Delta = \text{diag}(I(T_k \leq C)/P(T_k \leq C|X, V; \alpha_k), k = 1, 2, \dots, K)$  is a weight matrix,

and  $\alpha = (\alpha_1, \dots, \alpha_K)'$  is the full vector of parameters for the weights. Under the working independence assumption, it is clear that the previous arguments for consistency hold in this setting with the indicated weights. The denominators of the entries in the weight matrix  $\Delta$  are key to ensuring that consistent estimators of  $\theta$  are obtained, but we can use stabilized weights [29] to improve efficiency. In this case, the diagonal entries in  $\Delta$  may be replaced by  $I(T_k \leq C)P(T_k \leq C)/P(T_k \leq C|X, V)$ , or  $I(T_k \leq C)P(T_k \leq C|X)/P(T_k \leq C|X, V)$ ,  $k = 1, 2, \dots, K$ ; we consider the latter in the next section.

## 4 EMPIRICAL STUDIES AND AN APPLICATION

### 4.1 SIMULATION STUDIES

Here, we consider some particular parameter settings to study the factors governing the development of spurious trends from the selection effects. We let  $\vartheta_0 = 0$  in (2.2),  $\vartheta_1 = \log(0.5)$  to correspond to a 50% reduction in the odds of success with the experimental product,  $\vartheta_2 = \log(0.2)$  to reflect the effect of  $V$  as increasing the odds of success, and  $\vartheta_3 = 0$  and  $\log(5)$  to correspond to no effect modification and a strong effect modifying role of  $V$ . Preisser et al. [30] give an algorithm for modeling multivariate binary data with a given marginal mean and covariance structure on the basis of conditional probabilities. Let  $\mu_1(X, V) = P(Y_1 = 1|X, V) = \pi_1(X, V)$  and  $\mu_k(X, V, \bar{Y}_k) = P(Y_k = 1|X, V, \bar{Y}_k)$  where  $k = 2, \dots, K$  denote conditional probabilities. Let

$$\Sigma_k(X, V) = \begin{pmatrix} \text{VAR}(Y_1|X, V) & \text{COV}(Y_1, Y_2|X, V) & \dots & \text{COV}(Y_1, Y_{k-1}|X, V) \\ \text{COV}(Y_2, Y_1|X, V) & \text{VAR}(Y_2|X, V) & \dots & \text{COV}(Y_2, Y_{k-1}|X, V) \\ \vdots & \vdots & \ddots & \vdots \\ \text{COV}(Y_{k-1}, Y_1|X, V) & \text{COV}(Y_{k-1}, Y_2|X, V) & \dots & \text{VAR}(Y_{k-1}|X, V) \end{pmatrix}$$

denote the  $(k-1) \times (k-1)$  covariance matrix of the subvector  $(Y_1, \dots, Y_{k-1})'$  given  $X$  and  $V$ . If we let

$$\psi_k^*(X, V) = (\text{COV}(Y_1, Y_k|X, V), \text{COV}(Y_2, Y_k|X, V), \dots, \text{COV}(Y_{k-1}, Y_k|X, V))'$$

be a  $(k-1) \times 1$  vector and

$$\psi_k(X, V) = [\Sigma_k(X, V)]^{-1} \psi_k^*(X, V),$$

then the conditional probability of success for the  $k$ th transfusion given  $X, V, \bar{Y}_k$  is

$$\mu_k(X, V) = P(Y_k = 1|X, V, \bar{Y}_k) = \pi_k(X, V) + \sum_{j=1}^{k-1} \psi_j(X, V)(Y_j - \pi_j(X, V)),$$

$k = 2, \dots, K$ , and we obtain

$$P(Y_1, \dots, Y_K|X, V; \theta) = \prod_{j=1}^K \mu_j(X, V)^{Y_j} (1 - \mu_j(X, V))^{1-Y_j}.$$

We followed this algorithm to generate the responses to the first  $K$  transfusions, and focus on the setting where  $\varphi = 1$ ; other values of the association parameter yielded similar results, which we do not report here. We next address the generation of the transfusion times over a treatment period of  $C = 10$  days, and suppose  $K = 5$  transfusions are to be analysed. For the transfusion process, we consider (2.1) with  $\zeta_3 = \zeta_4 = 0$ , set  $\zeta_2 = \log(0.5)$ , so a transfusion on the previous day reduces the

Table 1: Frequency properties of the estimated  $\theta_{k0}$  and  $\theta_{k1}$  for  $k = 1, 2, \dots, 5$ ; data are generated with a common censoring time,  $C_i = 10$ ; a marginal model was fit with  $P(Y_k = 1|X) = \text{expit}(\theta_{k0} + \theta_{k1}X)$ ; number of subjects  $n = 2, 000$ ; 1000 simulations.<sup>†</sup>

		Inverse Probability Weighted Analyses																				
$e^{\theta_3}$	$k$	Unweighted							Estimated Weights $P(T_k \leq C X, V)$							Estimated Stabilized Weights $P(T_k \leq C X, V)/P(T_k \leq C X)$						
		TRUE VALUE	BIAS	ASE	ESE	BIAS	ESE	ECP	BIAS	ESE	BIAS	ESE	ECP	BIAS	ESE	ECP	BIAS	ESE	ECP			
$\theta_{k0}$																						
1.0	1	-0.6931	0.0000	0.0671	0.0697	0.0007	0.0671	93.6	0.0007	0.0697	0.0671	93.6	0.0007	0.0697	0.0671	93.6	0.0007	0.0697	0.0671	93.6		
	2	-0.6931	0.0036	0.0673	0.0659	0.0019	0.0673	95.9	0.0019	0.0659	0.0673	95.8	0.0018	0.0659	0.0673	95.9	0.0018	0.0659	0.0673	95.9		
	3	-0.6931	0.0309	0.0682	0.0699	0.0033	0.0683	90.5	0.0033	0.0701	0.0683	93.6	0.0034	0.0701	0.0682	93.6	0.0034	0.0701	0.0683	93.6		
	4	-0.6931	0.1184	0.0720	0.0749	-0.0002	0.0727	60.8	-0.0002	0.0756	0.0727	94.4	-0.0004	0.0747	0.0725	94.9	-0.0004	0.0747	0.0725	94.9		
	5	-0.6931	0.2758	0.0813	0.0798	0.2742	0.0869	8.9	-0.0004	0.0867	0.0869	95.5	-0.0020	0.0844	0.0856	95.1	-0.0020	0.0844	0.0857	95.1		
$\theta_{k1}$																						
	1	-0.6190	0.0000	0.1025	0.1075	-0.0032	0.1025	92.8	-0.0032	0.1075	0.1025	92.8	-0.0032	0.1075	0.1025	92.8	-0.0032	0.1075	0.1025	92.8		
	2	-0.6190	-0.0001	0.1026	0.1025	0.0014	0.1026	95.3	0.0014	0.1025	0.1026	95.2	0.0014	0.1024	0.1026	95.5	0.0014	0.1024	0.1026	95.5		
	3	-0.6190	-0.0011	0.1040	0.1033	-0.0056	0.1041	95.7	-0.0044	0.1036	0.1041	95.7	-0.0046	0.1036	0.1040	95.5	-0.0046	0.1036	0.1040	95.6		
	4	-0.6190	-0.0054	0.1090	0.1111	-0.0058	0.1101	95.1	0.0001	0.1124	0.1101	95.1	-0.0001	0.1113	0.1098	95.6	-0.0001	0.1113	0.1099	95.6		
	5	-0.6190	-0.0172	0.1220	0.1208	-0.0177	0.1299	94.7	-0.0019	0.1311	0.1299	94.8	-0.0013	0.1257	0.1282	95.4	-0.0013	0.1257	0.1283	95.4		
$\theta_{k0}$																						
5.0	1	-0.6931	0.0000	0.0671	0.0697	0.0007	0.0671	93.6	0.0007	0.0697	0.0671	93.6	0.0007	0.0697	0.0671	93.6	0.0007	0.0697	0.0671	93.6		
	2	-0.6931	0.0036	0.0673	0.0659	0.0019	0.0673	95.9	0.0019	0.0659	0.0673	95.8	0.0018	0.0659	0.0673	95.9	0.0018	0.0659	0.0673	95.9		
	3	-0.6931	0.0309	0.0682	0.0699	0.0033	0.0683	90.5	0.0033	0.0701	0.0683	93.6	0.0034	0.0701	0.0682	93.6	0.0034	0.0701	0.0683	93.6		
	4	-0.6931	0.1184	0.0720	0.0749	-0.0002	0.0727	60.8	-0.0002	0.0756	0.0727	94.4	-0.0004	0.0747	0.0725	94.9	-0.0004	0.0747	0.0725	94.9		
	5	-0.6931	0.2758	0.0813	0.0798	0.2742	0.0869	8.9	-0.0004	0.0867	0.0869	95.5	-0.0020	0.0844	0.0856	95.1	-0.0020	0.0844	0.0857	95.1		
$\theta_{k1}$																						
	1	0.0000	0.0000	0.0950	0.0970	-0.0012	0.0950	94.5	-0.0012	0.0970	0.0950	94.5	-0.0012	0.0970	0.0950	94.5	-0.0012	0.0970	0.0950	94.5		
	2	0.0000	-0.0036	0.0951	0.0961	-0.0021	0.0951	95.4	0.0014	0.0961	0.0951	95.5	0.0015	0.0961	0.0951	95.4	0.0015	0.0961	0.0951	95.4		
	3	0.0000	-0.0309	0.0968	0.0952	-0.0353	0.0968	93.7	-0.0043	0.0954	0.0968	95.4	-0.0044	0.0955	0.0968	95.5	-0.0044	0.0955	0.0968	95.5		
	4	0.0000	-0.1184	0.1048	0.1048	-0.1193	0.1039	78.3	-0.0006	0.1063	0.1039	94.8	-0.0003	0.1056	0.1037	95.0	-0.0003	0.1056	0.1038	95.0		
	5	0.0000	-0.2758	0.1172	0.1144	-0.2747	0.1260	35.5	0.0006	0.1241	0.1260	96.0	0.0021	0.1222	0.1251	96.1	0.0021	0.1222	0.1252	96.1		

<sup>†</sup> ESE is the empirical standard error, ASE is the average sample standard error, and ECP is the empirical coverage probability

Table 2: Frequency properties of the estimated  $\theta_{k0}$  and  $\theta_{k1}$  for  $k = 1, 2, \dots, 5$ ; data are generated with varying censoring times; a marginal model was fit with  $P(Y_k = 1|X) = \text{expit}(\theta_{k0} + \theta_{k1}X)$ ; number of subjects  $n = 2,000$ ; 1000 simulations.<sup>†</sup>

		Inverse Probability Weighted Analyses																					
$e^{\theta_3}$	$k$	Unweighted					Known Weights $P(T_k \leq C X, V)$					Estimated Weights $P(T_k \leq C X, V)$					Estimated Stabilized Weights $P(T_k \leq C X, V)/P(T_k \leq C X)$						
		TRUE VALUE	BIAS	ASE	ESE	ECP	BIAS	ESE	ASE	ECP	BIAS	ESE	ASE	ECP	BIAS	ESE	ASE	ECP	BIAS	ESE	ASE	ECP	
<b>ASYMPTOTIC</b>																							
$\theta_{k0}$																							
1.0	1	-0.6931	0.0000	0.0007	0.0679	0.0671	94.4	0.0007	0.0679	0.0671	94.4	0.0007	0.0679	0.0671	94.4	0.0007	0.0679	0.0671	94.4	0.0007	0.0679	0.0671	94.4
	2	-0.6931	0.0036	0.0158	0.0713	0.0692	92.7	0.0036	0.0729	0.0704	93.9	0.0038	0.0729	0.0704	94.0	0.0026	0.0714	0.0693	94.7	0.0026	0.0714	0.0693	94.7
	3	-0.6931	0.0309	0.0446	0.0716	0.0725	90.3	-0.0033	0.0773	0.0793	95.0	-0.0033	0.0775	0.0791	94.7	-0.0025	0.0717	0.0727	95.1	-0.0025	0.0717	0.0727	95.1
	4	-0.6931	0.1184	0.1318	0.0773	0.0784	59.9	0.0014	0.1003	0.0980	95.4	0.0010	0.1002	0.0964	95.4	-0.0025	0.0777	0.0795	95.7	-0.0025	0.0777	0.0795	95.7
	5	-0.6931	0.2758	0.2923	0.0909	0.0903	10.2	0.0039	0.1434	0.1249	94.5	0.0060	0.1462	0.1164	93.4	0.0038	0.0947	0.0959	96.0	0.0038	0.0947	0.0959	96.0
$\theta_{k1}$																							
	1	-0.6190	0.0000	0.0023	0.1003	0.1024	95.5	0.0023	0.1003	0.1024	95.5	0.0023	0.1003	0.1024	95.5	0.0023	0.1003	0.1024	95.5	0.0023	0.1003	0.1024	95.5
	2	-0.6190	-0.0001	0.0000	0.1074	0.1057	94.5	-0.0008	0.1088	0.1074	94.8	-0.0012	0.1086	0.1074	94.5	0.0002	0.1076	0.1058	94.3	0.0002	0.1076	0.1058	94.3
	3	-0.6190	-0.0011	0.0002	0.1045	0.1104	95.5	0.0013	0.1124	0.1197	96.4	0.0007	0.1123	0.1194	96.1	0.0019	0.1047	0.1108	95.9	0.0019	0.1047	0.1108	95.9
	4	-0.6190	-0.0054	-0.0020	0.1208	0.1187	94.9	0.0007	0.1540	0.1464	94.6	0.0000	0.1548	0.1440	93.7	0.0044	0.1213	0.1203	95.9	0.0044	0.1213	0.1203	95.9
	5	-0.6190	-0.0172	-0.0268	0.1378	0.1355	94.0	-0.0040	0.2141	0.1893	94.2	-0.0033	0.2121	0.1751	92.7	-0.0049	0.1421	0.1437	95.0	-0.0049	0.1421	0.1437	95.0
$\theta_{k0}$																							
5.0	1	-0.6931	0.0000	0.0007	0.0679	0.0671	94.4	0.0007	0.0679	0.0671	94.4	0.0007	0.0679	0.0671	94.4	0.0007	0.0679	0.0671	94.4	0.0007	0.0679	0.0671	94.4
	2	-0.6931	0.0036	0.0158	0.0713	0.0692	92.7	0.0036	0.0729	0.0704	93.9	0.0038	0.0729	0.0704	94.0	0.0026	0.0714	0.0693	94.7	0.0026	0.0714	0.0693	94.7
	3	-0.6931	0.0309	0.0446	0.0716	0.0725	90.3	-0.0033	0.0773	0.0793	95.0	-0.0033	0.0775	0.0791	94.7	-0.0025	0.0717	0.0727	95.1	-0.0025	0.0717	0.0727	95.1
	4	-0.6931	0.1184	0.1318	0.0773	0.0784	59.9	0.0014	0.1003	0.0980	95.4	0.0010	0.1002	0.0964	95.4	-0.0025	0.0777	0.0795	95.7	-0.0025	0.0777	0.0795	95.7
	5	-0.6931	0.2758	0.2923	0.0909	0.0903	10.2	0.0039	0.1434	0.1249	94.5	0.0060	0.1462	0.1164	93.4	0.0038	0.0947	0.0959	96.0	0.0038	0.0947	0.0959	96.0
$\theta_{k1}$																							
	1	0.0000	0.0000	0.0022	0.0926	0.0949	95.9	0.0022	0.0926	0.0949	95.9	0.0022	0.0926	0.0949	95.9	0.0022	0.0926	0.0949	95.9	0.0022	0.0926	0.0949	95.9
	2	0.0000	-0.0036	-0.0140	0.0998	0.0981	94.0	-0.0024	0.1008	0.1000	95.0	-0.0026	0.1009	0.0999	94.8	-0.0010	0.1001	0.0982	94.4	-0.0010	0.1001	0.0982	94.4
	3	0.0000	-0.0309	-0.0460	0.0983	0.1030	93.4	0.0011	0.1071	0.1134	96.1	0.0011	0.1075	0.1132	96.2	0.0013	0.0990	0.1034	95.1	0.0013	0.0990	0.1034	95.1
	4	0.0000	-0.1184	-0.1324	0.1102	0.1121	79.3	-0.0036	0.1476	0.1424	94.5	-0.0032	0.1495	0.1403	94.5	0.0015	0.1115	0.1141	96.0	0.0015	0.1115	0.1141	96.0
	5	0.0000	-0.2758	-0.2941	0.1307	0.1303	38.3	-0.0068	0.2093	0.1880	95.3	-0.0079	0.2067	0.1728	93.4	-0.0041	0.1365	0.1405	95.4	-0.0041	0.1365	0.1405	95.4

<sup>†</sup> ESE is the empirical standard error, ASE is the average sample standard error, and ECP is the empirical coverage probability

odds of needing one on a current day, and set  $\zeta_0$  and  $\zeta_1$  such that  $P(N(C) \geq K|V = 0) = 0.6$  and  $\text{logit}(P(N(C) \geq K|V = 1)) - \text{logit}(P(N(C) \geq K|V = 0)) = \log(0.1)$ . This gives  $\zeta_0 = 0.624$  and  $\zeta_1 = -0.953$ . We let  $P(V = 1) = 0.5$  and  $P(X = 1) = 0.5$ . We simulated one thousand data sets with samples of  $n = 2000$  each.

Table 1 summarizes the results where the asymptotic bias is defined as the parameters in (3.10) minus the corresponding parameters in (2.5). The case with  $\exp(\vartheta_3) = 1$  is in the top half, and here, we see the unweighted analysis yields estimates consistent for the limiting values derived in Section 3; we report the empirical biases as the mean estimate minus the true value. Because we know the true parameter values, we can conduct weighted analyses with the true weights, and we report the corresponding results in the second set of columns. These demonstrate that inverse weighting yields estimators with negligible empirical bias. When weights are estimated on the basis of stratified discrete time survival models, the empirical biases are again negligible and as one would expect, there is a very slight decrease in the empirical standard errors. When stabilized weights are used, the results are again very similar to the nonstabilized weights. Table 2 gives analogous results for the case in which subjects are randomly censored over the course of the treatment period. Because of variable follow-up, the asymptotic calculations in Section 3 no longer apply, and the empirical biases for the naive analyses are slightly different from the asymptotic biases computed on the basis of common follow-up. Again, the use of true or estimated weights reduces the bias to a negligible level. In this setting, however, the reduction in the asymptotic and empirical standard errors from use of stabilized weights is much more substantial because of the greater variation in the weights employed (i.e., some of the weights are smaller in this setting because the  $k$ th transfusion is less likely to be experienced within a shorter period of follow-up).

#### 4.2 APPLICATION TO A PLATELET TRANSFUSION TRIAL

Here, we report on an illustrative analysis of the Mirasol trial discussed in Section 1.2 to compare the effectiveness of pathogen-inactivated platelets to standard platelets in patients with chemotherapy-induced thrombocytopenia. The cumulative distribution function for the 1 h CCI is given by transfusion number in Figure 2 where it can be seen that the median CCI decreases by transfusion number and the empirical probability of success ( $\text{CCI} > 7500$ ) decreases with transfusion number.

The available potential confounding variables we adjusted for in the inverse probability weight were treatment, sex, ABO group, baseline platelet count (above median of 43,500 or not), and systolic (above median of 120 or not) and diastolic (above median of 70 or not) blood pressure. Proportional odds regression models were used for each transfusion number with individuals censored at 28 days or the time of withdrawal. We considered stabilized weights with the numerators given by  $P(T_k \leq C|X)$ . Figure 4 gives the estimated probability of a successful transfusion by transfusion number for reference patients, with the solid circles corresponding to naive estimates based on

$$\sum_{i=1}^n I(T_{ik} \leq C_i)Y_{ik}(1 - X_i)W_{ik} / \sum_{i=1}^n I(T_{ik} \leq C_i)(1 - X_i)W_{ik}$$

where  $W_{ik} = 1$ . We present the inverse weighted estimates as diamonds and are computed them on the basis of stabilized weights with  $W_{ik} = P(T_{ik} \leq C_i|X_i)/P(T_{ik} \leq C_i|X_i, V_i)$ . It is readily apparent that the suggestion of a trend based on the unweighted estimates is greatly attenuated or eliminated with the weighted analyses. Figure 5 gives the box plots ( $\log_{10}$  scale) of the distribution of stabilized weights, where it can be seen that there is greater variation in the weights for higher transfusion numbers.

Next, a response model of the form

$$\text{logit}(\pi(X)) = \theta_{k0} + \theta_1 X \quad (4.1)$$

was fit using generalized estimating equations under a working independence covariance structure. The common regression coefficient in (4.1) was used to synthesize the evidence of a treatment effect over all transfusions. An unweighted analysis yielded  $\hat{\theta}_1 = -0.73$  (s.e.  $(\hat{\theta}_1) = 0.38$ ) giving an odds ratio estimate of 0.48 (95% CI: (0.23,1.02); p-value = 0.056). The inverse probability weighted generalized estimating equation gave  $\hat{\theta}_1 = -0.84$  (s.e.  $(\hat{\theta}_1) = 0.40$ ) and a slightly smaller estimate of the odds ratio at 0.43 (95% CI: (0.20,0.94); p-value = 0.034). There are at most two individuals with particularly large weights contributing to the analyses of the sixth and seventh transfusions, so we conducted a sensitivity analysis in which weights were truncated [31] at 10 and then 5, and the conclusions did not materially change ( $\tilde{\theta} = -0.85$  and  $-0.86$ , respectively).

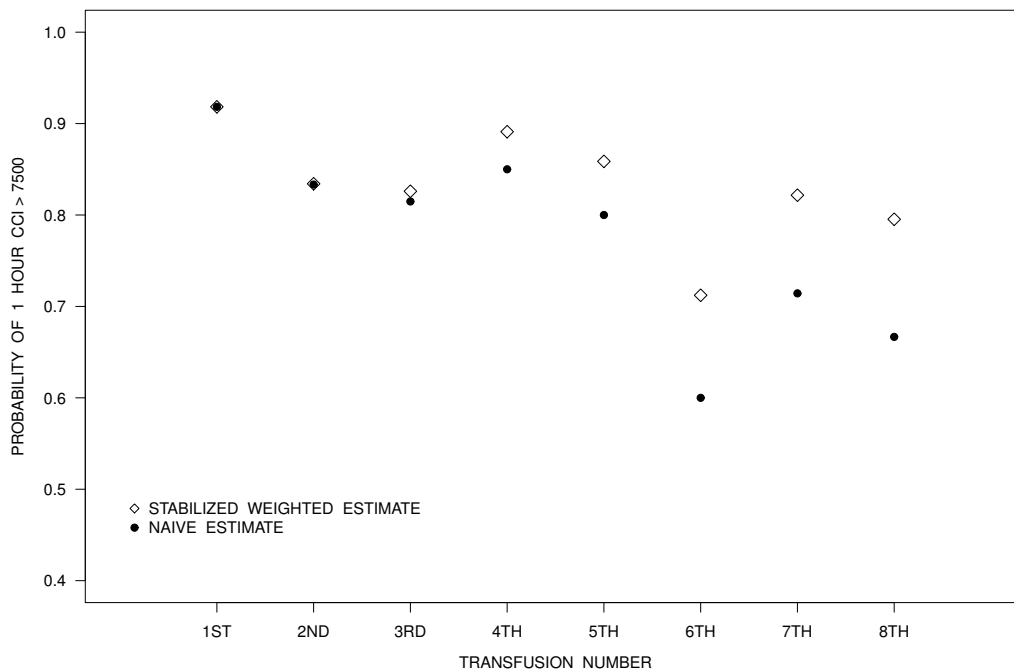


Figure 4: Unweighted and stabilized weighted estimates of the probabilities of success for patients in the reference group.

## 5 DISCUSSION

This article highlighted the need for care when analysing responses to administration of therapeutic interventions to treat episodic conditions. There are numerous clinical settings where the observations we make are relevant, and the findings should lead to alternative designs or frameworks for analysis. Examples of such settings include migraine studies, which involve randomizing patients to an experimental therapy or standard care, and interest lies in assessing the effectiveness of the experimental intervention relative to standard care [32, 33]. Studies of new rescue medications for exacerbations in chronic bronchitis aim to assess effects on the severity or duration of symptoms [34, 35]. In studies of therapies for erectile dysfunction, trials aim to evaluate treatment effects based in part on success rates in attempts at sexual intercourse [36, 37]. The distinguishing feature of these settings is that past responses may alter the probability of observing future responses by changing the distribution of times between critical events (e.g., exacerbations in respiratory disease or need for transfusions in thrombocytopenia).

Although data resulting from such designs have similar features to repeated measures data, there is an underlying stochastic mechanism, which leads to the need for transfusions, and this can mean that

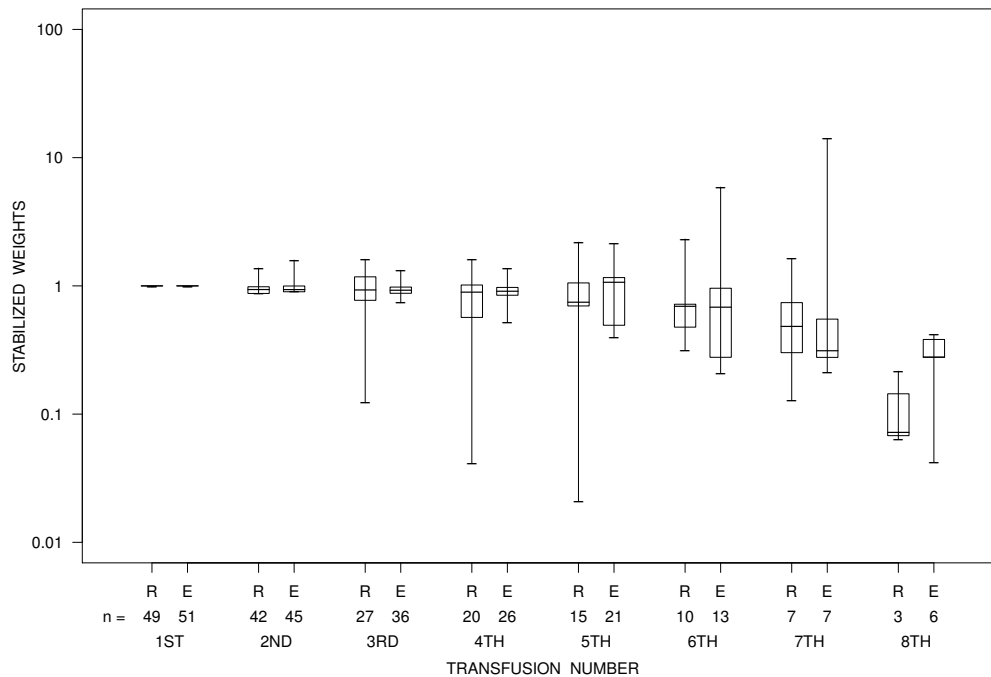


Figure 5: Box plots ( $\log_{10}$  scale) for the distribution of stabilized weights ( $P(T_k \leq C|X)/P(T_k \leq C|X, V)$ ) for patients in the reference group ( $R$ ) and experimental group ( $E$ ).

naive estimators are consistent for uninterpretable parameters. There are subtle issues related to the following: (i) heterogeneity in baseline platelet levels between patients; (ii) between patient heterogeneity in the response rates to transfusions; and (iii) the relationship between response to transfusion and the time of the next transfusion, all of which complicate analyses. Whether it is sensible to conceptualize a “response rate” across all transfusions, or indeed a common treatment effect across all transfusions, warrants careful consideration. The interest in simple summary measures of intervention effects motivated us to adopt a framework where the true response rates did not change over time. Even in this setting, patterns or trends may be apparent from naive analyses, but these are spurious and a reflection of heterogeneity in need for transfusions and the consequent selection effects. This work was motivated in part by apparent trends in response to transfusion by transfusion number. Although clinical researchers have suggested there may be a decreasing effect of transfusions over time [19], the models we present here offer alternative explanations.

There is a large literature on methods for incomplete longitudinal data. The focus of this work has been primarily on studies where follow-up assessments are scheduled at particular visit times and individuals may withdraw from a study before completing all of these assessments (i.e., monotone missing data problems). Inverse weighted estimating functions are perhaps most widely adopted for semiparametric analyses, and much of the literature is devoted to “missing at random” mechanisms [38]. In this context, these weights are estimated by modeling the conditional probability of remaining in the study until the next visit given completion of the current visit and the observed “history” [39, 40]. Less attention has been directed to the setting in which no visit schedule is planned and assessments happen in a “fully stochastic” fashion. One exception is Lin et al. [41] who consider intensity-based models for a continuous time visit process and estimating functions for marginal semiparametric models of the responses obtained upon assessment, which are inverse weighted by the intensity of the visit process. Zhu et al. [42] extend this methodology to deal with time-varying regression coefficients for the response model. Sun et al. [43, 44] develop joint models for the response process and assessment process through the use of shared frailty models. Chen et al. consider joint



models for the observation process and responses in the context of progressive models in continuous-time [45] and discrete-time [46] processes, respectively.

The approach we describe is somewhat similar in spirit with these approaches, but has key differences. First, rather than modeling the visit intensity (the instantaneous conditional probability of an assessment at time  $t$  given the history), we model the (not instantaneous) probabilities that a particular intervention ( $k$ th) will be required within the predetermined window of observation given covariates. Second, we adopt marginal models in the sense that we fit these models separately for the  $k$ th and  $\ell$ th assessments; that is, we do not make use of the fact that the  $k$ th assessment cannot happen if the  $(k - 1)$ st assessment has not yet happened for reasons of robustness. This is similar to the marginal approach of Pepe and Cai [47], which offers a robust approach to analysis. Third, we obtain further robustness with discrete covariates in which case one may stratify on the covariate values; we have demonstrated empirically that this performs well. Fourth, because we model the distributions of the times to the interventions rather than intensities, we do not need to adopt any kernel smoothing of baseline intensities. Finally, the observation process is fundamentally different in the present setting because it relates to a need for an intervention and so is directly tied to the underlying disease process; in much of the previous work, there are latent responses that are unobserved between assessments.

Inverse probability weighting is essential for fitting marginal models because these can be used to obtain estimates of response rates and treatment effects for second and subsequent transfusions, which are compatible with the estimates from the first transfusion. In practice, the weights for such analyses are often difficult to specify, and although model diagnostics offer one approach for assessment of the model for the weights, it is clear that this approach to the analysis is within realm of observational analyses and so will not furnish evidence of the same level as more simple randomized trials. An alternative approach to the analysis would be to match, stratify, or otherwise control for the propensity score at each transfusion [48]. This would require developing different propensity scores for each transfusion and would yield an estimate of the treatment effect conditional on the associated factors, which has a different interpretation than the estimate from the marginal structural model [49].

In some randomized trials, it is possible to randomize patients each time an intervention (e.g., transfusion) is required [50, 51], and this addresses the effect of confounding variables. There remains a subtle issue of whether the probability of response, or the effect of the intervention, is stable over successive transfusions, but we have assumed this in our current formulation; when this is not plausible, it is even more challenging to think about suitable objects of inference and analysis strategies.

## APPENDIX A REFORMULATION OF THE PROBLEM IN TERMS OF COUNTERFACTUALS

Discussions regarding causal inference often involve the conceptualization of counterfactual or “potential” outcomes. Although the need to adopt this framework is somewhat contentious [52], many view it to be a useful construct [53], and Robins [23] defines marginal structural models in terms of counterfactuals. Herein, we make some related remarks at the suggestion of a referee. Adapting the notation of Section 2.1, we let  $Z^{[x=1]}(t)$  and  $Z^{[x=0]}(t)$  denote the indicators that an individual would receive a transfusion on day  $t$  if they were randomized to the experimental or standard platelet product respectively; the counterfactual transfusion history at time  $t$  is given by  $\bar{Z}^{[x]}(t)$ . We let  $Y_k^{[x]}$  be the indicator that the  $k$ th transfusion under treatment  $x$  yields a successful response and let the histories in discrete time and in terms of transfusion number be  $\bar{Y}^{[x]}(t)$  and  $\bar{Y}_k^{[x]}$ , respectively. Under the *consistency* assumption [24], the observed transfusion and outcome processes correspond to the counterfactuals for the treatment actually received, that is,  $Z(t) = Z^{[x]}(t)$  and  $Y_k = Y_k^{[x]}$  if  $X = x$ .

The response model in (2.4) is compatible with the marginal structural model

$$P(Y_k^{[x]} = 1) = \text{expit}(\theta_0^\dagger + \theta_1^\dagger x) . \quad (\text{A.1})$$

Because this is a model for the potential outcome  $Y_k^{[x]}$  the parameter  $\theta_1^\dagger$  has a causal interpretation. In particular  $\exp(\theta_1^\dagger)$  is the causal odds ratio for a successful response when an individual is given the experimental platelet product versus the standard. The model in (A.1) can be extended to allow the parameters  $\theta_0^\dagger$  and/or  $\theta_1^\dagger$  to depend on transfusion number; when the intercept depends on  $k$ , we denote this by  $\theta_{k0}^\dagger$  and are stratifying on transfusion number, and when the treatment coefficient depends on  $k$ , we write  $\theta_{k1}^\dagger$  to allow the treatment effect to vary by transfusion number.

Although treatment  $X$  is randomized and therefore unconfounded, the probability of observing the  $k$ th transfusion before the end of follow-up on day  $C$  may not be the same in each treatment arm. We handle this administrative censoring by inclusion of weights  $w_k(t) = [P(T_k \leq C|X, V)]^{-1}$  in (3.11). Under the assumptions of *positivity* and *no unmeasured confounders* [25] and assuming the model for  $P(T_k \leq C|X, V)$  is correctly specified, adjustment via these weights allows us to estimate the causal odds ratio for experimental versus standard platelet products for the setting in which subjects in each treatment arm are at the same risk of receiving  $k$  transfusions.

To see this, we now show that (3.11) is an unbiased estimating function for  $\theta$  by taking its expectation. First note that

$$P(T_k \leq C|X, V) = \frac{P(X|V, T_k \leq C)P(T_k \leq C|V)P(V)}{P(X)P(V)} = 2 \cdot P(X|V, T_k \leq C)P(T_k \leq C|V) . \quad (\text{A.2})$$

Thus, we can write

$$U_k(\theta_k; \alpha_k) = \frac{I(T_k \leq C)}{2 \cdot P(X|V, T_k \leq C)P(T_k \leq C|V)} D'_k[\text{VAR}(Y_k|X)]^{-1} [Y_k - \pi_k(X)]$$

We can show that the first element of this estimating equation has expectation zero by first taking the expectation with respect to  $Y_k|X, V, T_k \leq C$  to obtain

$$\frac{I(T_k \leq C)}{2 \cdot P(X|V, T_k \leq C)P(T_k \leq C|V)} [E(Y_k|X, V, T_k \leq C) - \pi_k(X)] .$$

Next, we take the expectation with respect to  $X|V, T_k \leq C$

$$\begin{aligned} & \sum_{x_0=0}^1 \frac{I(T_k \leq C)}{2 \cdot P(X = x_0|V, T_k \leq C)P(T_k \leq C|V)} [E(Y_k|X = x_0, V, T_k \leq C) - \pi_k(x_0)] P(X = x_0|V, T_k \leq C) \\ &= \sum_{x_0=0}^1 \frac{I(T_k \leq C)}{2 \cdot P(T_k \leq C|V)} [E(Y_k|X = x_0, V, T_k \leq C) - \pi_k(x_0)] . \end{aligned}$$

Next, we take the expectation with respect to  $I(T_k \leq C)|V$  or equivalently  $I_k|V$  where  $I_k = I(T_k \leq C)$ . This gives

$$\begin{aligned} & \sum_{i_0=0}^1 \sum_{x_0=0}^1 \frac{i_0}{2 \cdot P(I_k = i_0|V)} [E(Y_k|X = x_0, V, I_k = i_0) - \pi_k(x_0)] P(I_k = i_0|V) \\ &= \sum_{x_0=0}^1 \frac{1}{2} [E(Y_k|X = x_0, V) - \pi_k(x_0)] , \end{aligned}$$

because when  $i_0 = 0$ , there is no contribution. The net effect of the cancellation of  $P(I_k = 1|V)$  previously is to “average” over  $I_k$  in the “setting” where the probability of observing the  $k$ th transfusion

is 1. This is the setting we would be working in if we did not have a finite period of observation, and is the setting compatible with the specification of the marginal response model. Hence, we remove the condition  $I_k = 1$ . Finally, we can take the expectation with  $V$  as

$$\sum_{v_0=0}^1 \sum_{x_0=0}^1 \frac{1}{2} [E(Y_k|X = x_0, V = v_0) - \pi_k(x_0)] P(V = v_0) = \sum_{x_0=0}^1 \frac{1}{2} [E(Y_k|X = x_0) - \pi_k(x_0)] = 0$$

The estimating equation for the previous marginal structural model can be fit through use of weighted generalized estimating equations as given in (3.11); large sample properties of the resulting estimators are discussed in Appendix B.

## APPENDIX B DERIVATION OF ASYMPTOTIC VARIANCE FROM WEIGHTED ANALYSES

The following derivation is based on the general theory layed out by Newey and McFadden [54] and Robins et al. [39]. Here, we introduce a subscript  $i$  to index individuals because we must compute empirical averages over individuals for variance estimation. Let  $C_i$  denote the last observation time for individual  $i$  and  $C_i(t) = I(t \leq C_i)$  and  $C_{ik}^\dagger(t) = I(t \leq T_{ik})$  and  $\bar{C}_{ik}(t) = C_i(t)C_{ik}^\dagger(t)$ . Let  $T_{ik}$  denote the time of the  $k$ th transfusion for individual  $i$ , and let  $\mathcal{F}_k(t|x, v) = P(T_k > t|x, v; \alpha)$ . Then, suppose we let

$$p_k(t|x, v) = P(T_k = t|T_k \geq t, x, v)$$

be the discrete hazard for an individual with covariate vector  $(x, v)$  at  $t \geq k$ , and

$$\text{logit}(p_k(t|x, v)) = u'_k(t)\alpha_k(x, v)$$

where  $u_k(t) = (I(t = k + 1), \dots, I(t = C))$  is a  $(C - k + 1) \times 1$  vector of time-dependent indicators  $\alpha_k(x, v) = (\alpha_{kk}(x, v), \dots, \alpha_{kC}(x, v))'$  is a vector of coefficients where  $\alpha_{kk}(x, v) = \text{logit}(p_k(k|x, v))$  and  $\alpha_{k\ell}(x, v) = \text{logit}(p_k(\ell|x, v)) - \alpha_{kk}(x, v)$ ,  $\ell = k + 1, \dots, C$ . Then

$$\widehat{\mathcal{F}}_k(t|x, v) = \prod_{s=k}^t (1 - \widehat{p}_k(s|x, v)). \quad (\text{B.1})$$

Let  $dN_{ik}(t) = 1$  if  $k$ th transfusion occurs on day  $t$  for individual  $i$  and zero otherwise. Consider the estimating function

$$S_{ikr}(\alpha_{kr}(x_i, v_i)) = \sum_x \sum_v \bar{C}_{ik}(r) \cdot (dN_{ik}(r) - p_{ik}(r|x, v)) \cdot I(x = x_i, v = v_i)$$

and the vector

$$S_{ik}(\alpha_k(x_i, v_i)) = \begin{pmatrix} S_{ikk}(\alpha_k(x_i, v_i)) \\ S_{ik,k+1}(\alpha_k(x_i, v_i)) \\ \vdots \\ S_{ikC}(\alpha_k(x_i, v_i)) \end{pmatrix}$$

Setting  $S_{ik}(\alpha_k(x_i, v_i)) = 0$  gives  $\widehat{\alpha}_k(x_i, v_i)$ , which can be plugged in to get estimate (B.1).

The derivation of the asymptotic variance of  $\widehat{\theta}_k$  from the weighted estimating equation is as follows. Let  $\psi_k = (\alpha'_k, \theta'_k)'$ . Let us proceed first by just considering the weighted estimating function for the response to the  $k$ th transfusion

$$G_k(\psi_k) = \sum_{i=1}^n G_{ik}(\psi_k) = \begin{pmatrix} S_{ik}(\psi_k) \\ U_{ik}(\psi_k) \end{pmatrix} = \sum_{i=1}^n \begin{pmatrix} S_{ik}(\alpha_k) \\ U_{ik}(\theta_k, \alpha_k) \end{pmatrix}.$$

Then

$$G_k(\widehat{\psi}_k) = G_k(\psi_k) + \frac{dG_k(\psi_k)}{d\psi_k} (\widehat{\psi}_k - \psi_k) + o_p\left(\frac{1}{\sqrt{n}}\right)$$

and

$$\sqrt{n} (\widehat{\psi}_k - \psi_k) = - \left[ \frac{1}{n} \frac{dG_k(\psi_k)}{d\psi_k} \right]^{-1} \left[ \frac{1}{\sqrt{n}} G_k(\psi_k) \right] + o_p(1) \quad (\text{B.2})$$

where

$$\frac{dG_k(\psi_k)}{d\psi'_k} = \begin{bmatrix} dS_k(\alpha_k)/d\alpha'_k & dS_k(\alpha_k)/d\theta'_k \\ dU_k(\psi_k)/d\alpha'_k & dU_k(\psi_k)/d\theta'_k \end{bmatrix}.$$

Note

$$\frac{1}{n} \frac{dG_k(\psi_k)}{d\psi_k} = \frac{1}{n} \sum_{i=1}^n \frac{dG_{ik}(\psi_k)}{d\psi_k}$$

which converges in probability to  $E(dG_{ik}(\psi_k)/d\psi_k)$  where

$$E \left( \frac{dG_{ik}(\psi_k)}{d\psi_k} \right) = \begin{bmatrix} E(dS_{ik}(\alpha_k)/d\alpha'_k) & 0 \\ E(dU_{ik}(\psi_k)/d\alpha'_k) & \Gamma(\psi_k) \end{bmatrix}$$

where  $\Gamma(\psi_k) = E(dU_{ik}(\psi_k)/d\theta'_k)$ . From (B.2) then  $\sqrt{n} (\widehat{\psi}_k - \psi_k)$  converges to

$$\begin{aligned} & - \left[ E \left( \frac{dG_{ik}(\psi_k)}{d\psi'_k} \right) \right]^{-1} \left[ \frac{1}{\sqrt{n}} \begin{pmatrix} S_k(\alpha_k) \\ U_k(\psi_k) \end{pmatrix} \right] \\ = & - \frac{1}{\sqrt{n}} \begin{bmatrix} [E(dS_{ik}(\alpha_k)/d\alpha'_k)]^{-1} S_k(\alpha_k) \\ -\Gamma^{-1}(\psi_k) E(dU_{ik}(\psi_k)/d\alpha_k) [E(dS_{ik}(\alpha_k)/d\alpha'_k)]^{-1} S_k(\alpha_k) + \Gamma^{-1}(\psi_k) U_k(\psi_k) \end{bmatrix} \end{aligned}$$

Thus,

$$\sqrt{n} (\widehat{\theta}_k - \theta_k) \rightarrow - \frac{1}{\sqrt{n}} \Gamma^{-1}(\psi_k) \left\{ U_k(\psi_k) - E \left( \frac{dU_{ik}(\psi_k)}{d\alpha'_k} \right) \left[ E \left( \frac{dS_{ik}(\alpha_k)}{d\alpha'_k} \right) \right]^{-1} S_k(\alpha_k) \right\}$$

which converges in distribution to

$$N(0, \Gamma^{-1}(\psi_k) E [Q_{ik}(\psi_k) Q'_{ik}(\psi_k)] [\Gamma^{-1}(\psi_k)]')$$

where

$$Q_{ik}(\psi_k) = U_{ik}(\psi_k) - E \left( \frac{dU_{ik}(\psi_k)}{d\alpha'_k} \right) \left[ E \left( \frac{dS_{ik}(\alpha_k)}{d\alpha'_k} \right) \right]^{-1} S_{ik}(\alpha_k).$$

As is commonly carried out, we estimate the requisite expectations with their empirical counterparts, and hence

$$\widehat{\text{asvar}}(\sqrt{n}(\widehat{\psi}_k - \psi_k)) = \widehat{\Gamma}^{-1}(\widehat{\psi}_k) \widehat{B}_k(\widehat{\psi}_k) \left[ \widehat{\Gamma}^{-1}(\widehat{\psi}_k) \right]'$$

where

$$\begin{aligned} \widehat{\Gamma}(\widehat{\psi}_k) &= \frac{1}{n} \sum_{i=1}^n \frac{dU_{ik}(\psi_k)}{d\theta'_k} \Big|_{\psi_k=\widehat{\psi}_k}, \\ \widehat{B}_k(\widehat{\psi}_k) &= \frac{1}{n} \sum_{i=1}^n \widehat{Q}_{ik}(\psi_k) \widehat{Q}'_{ik}(\psi_k) \Big|_{\psi_k=\widehat{\psi}_k} \end{aligned}$$

and

$$\widehat{Q}_{ik}(\psi_k) = U_{ik}(\psi_k) - \left[ \frac{1}{n} \sum_{i=1}^n \frac{dU_{ik}(\psi)}{d\alpha'_k} \right] \left[ \frac{1}{n} \sum_{i=1}^n \frac{dS_{ik}(\alpha_k)}{d\alpha'_k} \right]^{-1} S_{ik}(\alpha_k).$$

The derivation of the asymptotic variance of the estimates of the parameters in the marginal structural model obtained by inverse weighting is similar to Hernan et al. [25].

## ACKNOWLEDGEMENTS

This research was supported by Discovery Grants to RJC (RGPIN 155849) and CAC (RGPIN 402474) and a Graduate Student Scholarship (MC) from the Natural Sciences and Engineering Research Council of Canada and a grant to RJC from the Canadian Institutes for Health Research (FRN 13887). Richard Cook is a Canada Research Chair in Statistical Methods for Health Research. The authors thank Professor Nancy Heddle for helpful discussion and Ray Goodrich of CaridianBCT for permission to use the data from the platelet transfusion trial.

## REFERENCES

- [1] Webert K, Cook RJ, Sigouin CS, Rebullia P, Heddle NM. The risk of bleeding in thrombocytopenic patients with acute myeloid leukemia. *Haematologica* 2006; **91**: 1530–1537.
- [2] Rock GA. Management of thrombotic thrombocytopenic purpura. *British Journal of Haematology* 2000; **109**: 496–507. DOI:10.1046/j.1365-2141.2000.01941.x
- [3] Warkentin TE, Levine MN, Hirsh J, Horsewood P, Roberts RS, Gent M, Kelton JG. Heparin-induced thrombocytopenia in patients treated with low-molecular-weight heparin or unfractionated heparin. *New England Journal of Medicine* 1995; **332**: 1330–1336. DOI:10.1056/NEJM199505183322003
- [4] Yoto Y, Kudoh T, Suzuki N, Katoh S, Matsunaga Y, Chiba S. Thrombocytopenia induced by human parvovirus B19 infections. *European Journal of Haematology* 1993; **50**: 255–257. DOI:10.1111/j.1600-0609.1993.tb00158.x
- [5] Franck H, Rau R, Herborn G. Thrombocytopenia in patients with rheumatoid arthritis on long-term treatment with low dose methotrexate. *Clinical Rheumatology* 1996; **15**: 163–167. DOI:10.1007/BF02229705
- [6] Elting LS, Rubenstein EB, Martin CG, Kurtin D, Rodriguez S, Laiho E, Kanesan K, Cantor SB, Benjamin RS. Incidence, cost, and outcomes of bleeding and chemotherapy dose modification among solid tumor patients with chemotherapy-induced thrombocytopenia. *Journal of Clinical Oncology* 2001; **19**: 1137–1146.
- [7] Rodeghiero F, Stasi R, Gernsheimer T, Michel M, Provan D, Arnold DM, Bussel JB, Cines DB, Chong BH, Cooper N, Godeau B, Lechner K, Mazzucconi MG, McMillan R, Sanz MA, Imbach P, Blanchette V, Kuhne T, Ruggeri M, George JN. Standardization of terminology, definitions and outcome criteria in immune thrombocytopenic purpura of adults and children: report from an international working group. *Blood* 2009; **113**: 2386–2393. DOI:10.1182/blood-2008-07-162503

- [8] van Rhenen D, Gulliksson H, Cazenave J-P, Pamphilon D, Ljungman P, Klüter H, Vermeij H, Kappers-Klunne M, de Greef G, Laforet M, Lioure B, Davis K, Marblie S, Mayaudon V, Flament J, Conlan M, Lin L, Metzel P, Buchholz D, Corash L. Transfusion of pooled buffy coat platelet components prepared with photochemical pathogen inactivation treatment: the euroSPRITE trial. *Blood* 2003; **101**: 2426–2433. DOI:10.1182/blood-2002-03-0932
- [9] McCullough J, Vesole DH, Benjamin RJ, Slichter SJ, Pineda A, Snyder E, Stadtmauer EA, Lopez-Plaza I, Coutre S, Strauss RG, Goodnough LT, Fridey JL, Raife T, Cable R, Murphy S, Howard F, Davis K, Lin J-S, Metzel P, Corash L, Koutsoukos A, Lin L, Buchholz DH, Conlan MG. Therapeutic efficacy and safety of platelets treated with a photochemical process for pathogen inactivation: the SPRINT trial. *Blood* 2004; **104**: 1534–1541. DOI:10.1182/blood-2003-12-4443
- [10] Janetzko K, Cazenave J-P, Klüter H, Kientz D, Michel M, Beris P, Lioure B, Hastka J, Marblie S, Mayaudon V, Lin L, Lin J-S, Conlan MG, Flament J. Therapeutic efficacy and safety of photochemically treated apheresis platelets processed with an optimized integrated set. *Transfusion* 2005; **45**: 1443–1452. DOI:10.1111/j.1537-2995.2005.00550.x
- [11] Cazenave J-P, Folléa G, Bardiaux L, Boiron J-M, Lafeuillade B, Debost M, Lioure B, Harousseau J-L, Tabrizi R, Cahn J-Y, Michallet M, Ambruso D, Schots R, Tissot J-D, Sensebé L, Kondo T, McCullough J, Rebullia P, Escolar G, Mintz P, Heddle NM, Goodrich RP, Bruhwylér J, Le C, Cook RJ, Stouch B, for members of the The Mirasol Clinical Evaluation Study Group. A randomized controlled clinical trial evaluating the performance and safety of platelets treated with MIRASOL pathogen reduction technology. *Transfusion* 2010; **50**: 2362–2375. DOI: 10.1111/j.1537-2995.2010.02694.x
- [12] Kerkhoffs J-LH, van Putten WLJ, Novotny VMJ, Te Boekhorst PAW, Schipperus MR, Zwaginga JJ, van Pampus LC, de Greef GE, Lutén M, Huijgens PC, Brand A, van Rhenen DJ, on behalf of the Dutch-Belgian HOVON Cooperative Group. Clinical effectiveness of leucoreduced, pooled donor platelet concentrates, stored in plasma or additive solution with and without pathogen reduction. *British Journal of Haematology* 2010; **150**: 209–217. DOI:10.1111/j.1365-2141.2010.08227.x
- [13] Lozano M, Knutson F, Tardivel R, Cid J, Maymó RM, Löf H, Roddie H, Pelly J, Docherty A, Sherman C, Lin L, Propst M, Corash L, Prowse C. A multi-centre study of therapeutic efficacy and safety of platelet components treated with amotosalen and ultraviolet A pathogen inactivation stored for 6 or 7 d prior to transfusion. *British Journal of Haematology* 2011; **153**: 393–401. DOI:10.1111/j.1365-2141.2011.08635.x
- [14] Corash L, Sherman CD. Evaluation of platelet transfusion trials. *British Journal of Hematology* 2011; **153**: 529–530. DOI:10.1111/j.1365-2141.2010.08413.x
- [15] Kerkhoffs J-LH. Evaluation of platelet transfusion clinical trials - response to Corash and Sherman. *British Journal of Hematology* 2011; **153**: 530–531. DOI:10.1111/j.1365-2141.2010.08415.x
- [16] Heddle NM, Cook RJ, Webert KE, Sigouin CS, Rebullia P. Methodological issues in the use of bleeding outcomes in transfusion medicine studies. *Transfusion* 2003; **43**: 742–752. DOI:10.1046/j.1537-2995.2003.00418.x
- [17] Davis KB, Slichter SJ, Corash L. Corrected count increment and percent platelet recovery as measures of posttransfusion platelet response: problems and a solution. *Transfusion* 1999; **39**: 586–592. DOI:10.1046/j.1537-2995.1999.39060586.x

- [18] Yusuf S, Wittes J, Probstfield J, Tyroler HA. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *Journal of the American Medical Association* 1991; **266**: 93–98. DOI:10.1001/jama.266.1.93
- [19] Cazenave J-P, Davis K, Corash L. Design of clinical trials to evaluate the efficacy of platelet transfusion: The euroSPRITE trial for components treated with helinx technology. *Seminars in Hematology* 2001; **38**(11): 46–54. DOI:10.1016/S0037-1963(01)90123-4
- [20] Lipsitz SR, Laird NM, Harrington DP. Generalized estimating equations for correlated binary data: Using the odds ratio as a measure of association. *Biometrika* 1991; **78**: 153–160. DOI:10.1093/biomet/78.1.153
- [21] Liang KL, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986; **73**: 13–22. DOI:10.1093/biomet/73.1.13
- [22] Zeger SL, Liang KL. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986; **42**: 121–130. DOI:10.2307/2531248
- [23] Robins JM. Association, causation, and marginal structural models. *Synthese* 1999; **121**: 151–179. DOI:10.1023/A:1005285815569
- [24] Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000; **11**(5):550–560. DOI:10.1097/00001648-200009000-00011
- [25] Hernan MA, Brumback B, Robins JM. Marginal structural models to estimate the joint causal effect of nonrandomized treatments. *Journal of the American Statistical Association* 2001; **96**(454):440–448. DOI:10.1198/016214501753168154
- [26] Joffe MM, Ten Have TR, Feldman HI, Kimmel SE. Model selection, confounder control, and marginal structural models: review and new applications. *The American Statistician* 2004; **58**(4):272–279. DOI:10.1198/000313004X5824
- [27] Crowder M. On the use of a working correlation matrix in using generalized linear models for repeated measures. *Biometrika* 1995; **82**, 407–410. DOI:10.1093/biomet/82.2.407
- [28] Sutradhar BC, Das K. On the efficiency of regression estimators in generalised linear models for longitudinal data. *Biometrika* 1999; **86**, 459–465. DOI:10.1093/biomet/86.2.459
- [29] Hernan MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 2000; **11**(5):561–570. DOI:10.1097/00001648-200009000-00012
- [30] Preisser JS, Lohman KK, Rathouz PJ. Performance of weighted estimating equations for longitudinal binary data with drop-outs missing at random. *Statistics in Medicine* 2002; **21**: 3035–3054. DOI:10.1002/sim.1241
- [31] Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. *American Journal of Epidemiology* 2008; **168**: 656–664. DOI:10.1093/aje/kwn164
- [32] Pascual J, Falk RM, Piessens F, Prusinski A, Docekal P, Robert M, Ferrer P, Luria X, Segarra R, Zayas JM. Consistent efficacy and tolerability of almotriptan in the acute treatment of multiple migraine attacks: results of a large, randomized, double-blind, placebo-controlled study. *Cephalalgia* 2000; **20**: 588–596. DOI: 10.1046/j.1468-2982.2000.00091.x

- [33] Ahonen K, Hamalainen ML, Eerola M, Hoppu K. A randomized trial of rizatriptan in migraine attacks in children. *Neurology* 2006; **67**: 1135–1140. DOI: 10.1212/01.wnl.0000238179.79888.44
- [34] Grossman R, Mukherjee J, Vaughan D, Eastwood C, Cook R J, LaForge J, Lampron N. The Canadian Ciprofloxacin Health Economic Study Group. A 1-year community-based health economic study of ciprofloxacin vs usual antibiotic treatment in acute exacerbations of chronic bronchitis. *CHEST* 1998; **113**: 131–141. DOI:10.1378/chest.113.1.131
- [35] Wilson R, Allegra L, Huchon G, Izquierdo J, Jones P, Schaberg T, Sagnier P. Short-term and long-term outcomes of Moxifloxacin compared to standard antibiotic treatment in acute exacerbations of chronic bronchitis. *CHEST* 2004; **125**: 953–964. DOI: 10.1378/chest.125.3.953
- [36] Moncada I, Damber J, Mirone V, Wespes E, Casariego J, Chan M, Varanese L. Sexual intercourse attempt patterns with two dosing regimens of tadalafil in men with erectile dysfunction: results from the SURE study in 14 European countries. *The Journal of Sexual Medicine* 2005; **2**: 668–674. DOI:10.1111/j.1743-6109.2005.00102.x
- [37] van Ahlen H, Wahle K, Kupper W, Yassin A, Reblin T, Neureither M. Safety and efficacy of vardenafil, a selective phosphodiesterase 5 inhibitor, in patients with erectile dysfunction and arterial hypertension treated with multiple antihypertensives. *The Journal of Sexual Medicine* 2005; **2**: 856–864. DOI:10.1111/j.1743-6109.2005.00150.x
- [38] Little RJA, Rubin DB. *Statistical Analysis with Missing Data* (2nd ed.), 2002, New York: Wiley.
- [39] Robins JM, Rotnitzky A, Zhao LP. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data, *Journal of the American Statistical Association* 1995; **90** 106–121. DOI:10.1080/01621459.1995.10476493
- [40] Robins JM, Rotnitzky A. Semiparametric efficiency in multivariate regression models with missing data, *Journal of the American Statistical Association* 1995; **90**, 122–129. DOI:10.1080/01621459.1995.10476494
- [41] Lin H, Scharfstein DO, Rosenheck RA. Analysis of longitudinal data with irregular, outcome-dependent follow-up. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 2004; **66**: 791–813. DOI: 10.1111/j.1467-9868.2004.b5543.x
- [42] Zhu L, Sun J, Tong X, Pounds S. Regression analysis of longitudinal data with informative observation times and application to medical cost data. *Statistics in Medicine* 2011; **30**: 1429–1440. DOI:10.1002/sim.4198
- [43] Sun J, Park D-H, Sun L, Zhao X. Semiparametric regression analysis of longitudinal data with informative observation times. *Journal of the American Statistical Association* 2005; **471**: 882–889. DOI:10.1198/016214505000000060
- [44] Sun J, Sun L, Liu D. Regression analysis of longitudinal data in the presence of informative observation and censoring times. *Journal of the American Statistical Association* 2007; **102**: 1397–1406. DOI:10.1198/016214507000000851
- [45] Chen B, Yi GY, Cook RJ. Analysis of interval-censored disease progression data via multi-state models under a nonignorable inspection process. *Statistics in Medicine* 2010; **29**: 1175–1189. DOI:10.1002/sim.3804



- [46] Chen B, Yi GY, Cook RJ. Progressive multi-state models for informatively incomplete longitudinal data. *Journal of Statistical Planning and Inference* 2011; **141**: 80–93. DOI:10.1016/j.jspi.2010.05.014
- [47] Pepe MS, Cai J. Some graphical displays and marginal regression analyses for recurrent failure times. *Journal of the American Statistical Association* 1993; **88**: 811–820. DOI:10.1080/01621459.1993.10476346
- [48] Austin PC, Mamdani MM, Stukel TA, Anderson GM, Tu JV. The use of the propensity score for estimating treatment effects: administrative versus clinical data. *Statistics in Medicine* 2005; **24**: 1563–1578. DOI:10.1002/sim.2053
- [49] Ertefaie A, Stephens DA. Comparing approaches to causal inference for longitudinal data: inverse probability weighting versus propensity scores. *The International Journal of Biostatistics* 2010; **6**: Article 14. DOI:10.2202/1557-4679.1198
- [50] Heddle NM, Klama L, Singer J, Richards C, Fedak P, Walker I, Kelton JG. The role of the plasma from platelet concentrates in transfusion reactions. *New England Journal of Medicine* 1994; **331**: 625–628. DOI:10.1056/NEJM199409083311001
- [51] Heddle NM, Cook RJ, Blajchman RL, Sigouin CS, Boye DM, Nelson EJ, Kelton JG. Assessing the effectiveness of whole blood-derived platelets stored as a pool: a randomized block noninferiority trial. *Transfusion* 2005; **45**: 896–903. DOI:10.1111/j.1537-2995.2005.04233.x
- [52] Dawid AP. Causal inference without counterfactuals. *Journal of the American Statistical Association* 2000; **450**: 407–424. DOI:10.1080/01621459.2000.10474210
- [53] Cox DR, Wermuth N. Causality: a statistical view. *International Statistical Review* 2004; **72**: 285–305. DOI:10.1111/j.1751-5823.2004.tb00237.x
- [54] Newey WK, McFadden D. Large sample estimation and hypothesis testing. *Handbook of Econometrics* 1994; **4**:2111-2245. DOI: 10.1016/S1573-4412(05)80005-4.