Biometrika (1999), **86,** 2, *pp.* 365–379

Printed in Great Britain

Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes

BY CONSTANTINE E. FRANGAKIS AND DONALD B. RUBIN
Department of Statistics, Harvard University, 1 Oxford Street, Cambridge, Massachusetts 02138, U.S.A.
frangaki@hustat.harvard.edu rubin@stat.harvard.edu

SUMMARY

We study the combined impact that all-or-none compliance and subsequent missing outcomes can have on the estimation of the intention-to-treat effect of assignment in randomised studies. In this setting, a standard analysis, which drops subjects with missing outcomes and ignores compliance information, can be biased for the intention-to-treat effect. To address all-or-none compliance that is followed by missing outcomes, we construct a new estimation procedure for the intention-to-treat effect that maintains good randomisation-based properties under more plausible, nonignorable noncompliance and nonignorable missing-outcome conditions: the 'compound exclusion restriction' on the effect of assignment and the 'latent ignorability' of the missing data mechanism. We present both theoretical results and a simulation study. Moreover, we show how the two key concepts of compound exclusion and latent ignorability are relevant in more complicated settings, such as right censoring of a time-to-event outcome.

Some key words: Compound exclusion restriction; Intention-to-treat; Latent ignorability; Noncompliance; Nonignorability; Rubin causal model.

1. INTRODUCTION

Randomised experiments with human subjects often suffer from two major complications, namely noncompliance to treatment assignment and missing outcomes. In general, noncompliance is selective (The Coronary Drug Project Research Group, 1980) in the sense that noncompliers and compliers generally differ in background characteristics. Moreover, missing outcomes, often caused by refusal or loss to follow-up, may also be selective in the analogous sense (Farwell et al., 1990). These complications are rarely fully within the experimenter's control, and there is currently substantial awareness among researchers that such complications in a study compromise the ability to draw clear conclusions.

We assume a simple two-arm randomised experiment comparing a new versus a standard treatment, with access to the new, experimental treatment only in the new treatment arm and all-or-none compliance. A special case is the Zelen randomised single-consent design (Zelen, 1979, 1990a), where those assigned standard treatment cannot receive the new treatment, and those assigned the new treatment either receive it or the standard treatment.

For any subject whose outcome at the end of the study is known, we will say there is a response and refer to the subject as a respondent, whereas if the outcome is missing we will say there is nonresponse and that the subject is a nonrespondent.

The bias of using the actual treatment received as if it had been randomly assigned, 'as-treated' analysis, has been well documented in related settings (The Coronary Drug Project Research Group, 1980; Mark & Robins, 1993; Robins & Greenland, 1994; Sheiner & Rubin, 1995). For this reason, the standard approach to randomised trials with noncompliance has been intention-to-treat analysis, which compares the originally randomised treatment assignment arms, thereby ignoring the observed actual treatment received. When the outcomes of study are observed for each subject, intention-to-treat analysis is valid for estimating the intention-to-treat treatment effect $\bar{Y}_1 - \bar{Y}_0$, that is, the effect of assignment on the population averages of the outcome Y. Commonly, though, observed outcomes are not available for all subjects. Even then, a standard analysis is a respondent-based intention-to-treat analysis, that is, an analysis that is based on the respondents and ignores compliance data, e.g. Farwell et al. (1990), Reuben et al. (1995). As we show, however, the practice of discarding compliance data in the presence of subsequent nonresponse can create a bias even when $\bar{Y}_1 - \bar{Y}_0$ is zero. Other issues that arise with intention-to-treat analyses are discussed by Little & Yau (1996) and Kleinman, Ibrahim & Laird (1998).

We provide a framework that explicitly allows both nonignorable (Rubin, 1976, 1978) noncompliance to the treatment assignment and nonignorable nonresponse on the outcome of study, and offer a new estimation procedure for $\bar{Y}_1 - \bar{Y}_0$ that maintains good randomisation-based properties more generally than respondent-based intention-to-treat analysis.

Our framework is developed in § 2, where we posit two assumptions, namely 'latent ignorability' and 'compound exclusion restriction'. Under our assumptions, we show that respondent-based intention-to-treat analysis, although commonly used to estimate $\bar{Y}_1 - \bar{Y}_0$, for example Lee et al. (1991), can be biased for $\bar{Y}_1 - \bar{Y}_0$, and that the bias can be reduced by using compliance data. In § 3, we discuss identifiability and construct our estimator for $\bar{Y}_1 - \bar{Y}_0$, which is consistent when respondent-based intention-to-treat analysis is consistent and also when respondent-based intention-to-treat analysis is inconsistent under our nonignorable conditions. In § 4, we present simulation results to illustrate some operating characteristics of our new procedure. In § 5 we apply these new ideas to the problem of identifiability in right censoring of a time-to-event outcome, a survival time, following all-or-none compliance. For the case of discrete-time survival, a related approach is independently considered by Baker (1998). Our final section gives concluding remarks. The Appendix provides technical details of our results.

2. FRAMEWORK

2·1. *The data*

In order to address better the different sources of missing information in this problem, we first define potential outcomes separately from a sampling scheme or probabilistic assignment mechanism, an approach dating back to Neyman (1923) in the context of perfect randomised trials, formalised and extended to nonrandomised studies in Rubin (1974, 1978), and referred to as the Rubin causal model (Holland, 1986). Consider a large population of individuals $\{\omega\}$, each of which can potentially participate in a study and be assigned a treatment z, with $z = 1$ for new, 0 for standard. For each individual ω, let $D(\omega, z)$ be the actual treatment received, 1 for new, 0 for standard, if that individual is

assigned treatment z. Also let $Y(\omega, z)$ and $R(\omega, z)$ be, respectively, the outcome and indicator for response, equal to 1 for response, 0 for nonresponse, on outcome Y, if individual ω is assigned treatment z.

A simple random sample of n subjects from $\{\omega\}$, $\omega_1, \ldots, \omega_n$, say, comprises the participants in the study. Each subject is then randomly assigned treatment arm Z_i, 1 for new, 0 for standard, where for simplicity Z_i $(i = 1, \ldots, n)$ are essentially independent, identically distributed Bernoulli variables. Then the observable vectors

$$(Z_i, D(\omega_i, Z_i), Y(\omega_i, Z_i), R(\omega_i, Z_i))$$

$(i = 1, \ldots, n)$ are essentially independent and identically distributed replicates with respect to the random sampling–random assignment mechanism, and will be denoted by (Z_i, D_i, Y_i, R_i).

We assume that the observed data from the study include

(i) the treatment assignments $\{Z_i\}$;
(ii) the actual treatments received $\{D_i\}$, where by assumption $D_i = 0$ if $Z_i = 0$; and
(iii) the indicators for whether or not outcomes are observed, $\{R_i\}$, where outcome Y_i is observed if $R_i = 1$ and missing if $R_i = 0$.

Hence, the observed data are

$$\{(Z_i, D_i, R_i) : i = 1, \ldots, n\} \quad \{Y_i : R_i = 1\}. \tag{2·1}$$

Let $\bar{Y}_z := E\{Y(\omega_i, z)\}$ be the expected outcome when all units are assigned z, for $z = 0, 1$, where the probability measure is the one induced by the random sampling from $\{\omega\}$. We focus on estimating the intention-to-treat treatment effect on the means of Y in $\{\omega\}$, given by

$$\bar{Y}_1 - \bar{Y}_0. \tag{2·2}$$

A more specific question is whether or not the intention-to-treat null hypothesis of no effect of assignment on the outcomes, $H_0 : \bar{Y}_1 - \bar{Y}_0 = 0$, is plausible.

The methods that we will discuss can formally be applied within the levels of other pretreatment covariates, but, to focus on conceptual issues, we assume that no covariate is recorded, or that we are already within a cell defined by such covariates.

2·2. *Role of treatment-noncompliance with missing outcomes*

Let $U(\omega) := D(\omega, 1)$ be the received treatment for individual ω when assigned the new treatment, where $U_i := U(\omega_i)$ for the ith study individual ω_i. Since $U_i = 1$ if person ω_i would comply under both treatment assignments, we refer to such a person as a 'complier', and, because $U_i = 0$ if person ω_i would never take the new treatment, no matter what the assignment, we refer to such a person as a 'never-taker'; see the Harvard Institute of Economic Research Discussion paper #1676 by G. W. Imbens and D. B. Rubin 'Causal inference with instrumental variables'. By definition, the quantity $U(\omega)$ is fixed for individual ω, and therefore it is a covariate, the true compliance status covariate (Angrist, Imbens & Rubin, 1996; Rubin, 1998), though it is only partially observed in the sample; $U_i = D_i$ when $Z_i = 1$, but U_i is unobserved when $Z_i = 0$. Despite being missing in the standard treatment arm, by randomisation, the covariate U_i has the same distribution in the standard treatment arm as in the new treatment arm in the study; see Fig. 1. The observed, post-treatment compliance behaviour, D_i, is completely determined by the values of the covariate U_i and the treatment assignment Z_i. Naive attempts to condition on D_i, the

observed treatment received, generally lead to biased conclusions because D_i is not a true covariate.

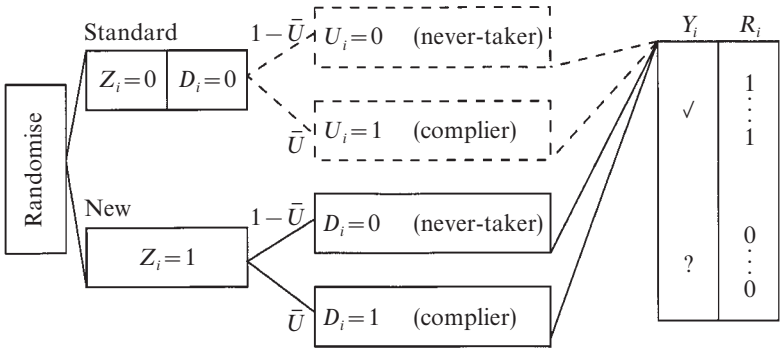

Fig. 1. The study and the unobserved, comparable compliance groups. Dashed lines represent unobserved information. Here, $\bar{U}$ is the proportion of compliers. To each of the four combinations of assignment arm by compliance covariate corresponds a rectangular array $\{(Y_i, R_i)\}$.

In the presence of subsequent missingness of the outcome Y_i, indicated by $R_i = 0$, a strict intention-to-treat analysis cannot be done and the role of the compliance covariate U_i is critical. To describe this role, consider for the moment an investigator who, hypothetically, knows the compliance covariate values U_i for all subjects, that is, whether each subject would comply with the new treatment, a complier, or not, a never-taker, under assignment to take it. Researchers have expressed the desire to have known this information, e.g. Schechtman & Gordon (1988), and consideration of this case suggests how to analyse the data without such information.

With U_i fully observed, the investigator would have a covariate, U_i, and therefore would have the following main options.

Approach 1. Ignore the fully observed covariate, and compare respondents between randomised arms. In order for this approach to be correct, response R_i should be independent of outcomes Y_i before conditioning on the covariates U_i, as when the missing Y values are missing completely at random (Little & Rubin, 1987, Ch. 1).

Approach 2. Do separate analyses for compliers, $U_i = 1$, and for never-takers, $U_i = 0$. Then, combine these analyses, weighted by the proportions of compliers and never-takers since the goal is the overall intention-to-treat effect. This approach allows for response R_i to be ignorable (Rubin, 1976) or, essentially here, independent of outcome Y_i after, but not necessarily before, conditioning on the compliance covariates U_i.

In practice it may be that neither of the above two approaches is correct, but this is not testable. To avoid confounding, it is general practice when confronted with missing outcomes first to condition on important covariates before assuming independent nonresponse. For example, we assume we are already within a cell defined by all other observed covariates, because such conditioning is standard. The clinical trials literature typically regards the latent compliance covariate U_i to be important in the sense of being associated with both (i) background characteristics of health related to the outcome Y_i, as well as (ii) response R_i; see for example The Coronary Drug Project Research Group (1980), Schechtman & Gordon (1988), Farwell et al. (1990), Gordon & Schechtman (1990) and

Zelen (1990b). Therefore, the investigator with knowledge of all values of U_i should regard Approach 2 as more reliable than Approach 1, because Approach 2 conditions on a potentially important covariate whereas Approach 1 does not.

Hence, to formalise this conclusion, we assume that, if U_i were fully observed, we would accept ignorability of the missing data mechanism for Y_i only after conditioning on U_i. That is, we would accept Assumption 1.

Assumption 1: *Latent ignorability*. Potential outcomes and associated potential non-response indicators are independent within each level of the latent compliance covariate:

(a) when assigned new treatment,

$$\mathrm{pr}\{Y(\omega_i, 1) \mid U(\omega_i), R(\omega_i, 1)\} = \mathrm{pr}\{Y(\omega_i, 1) \mid U(\omega_i)\};$$

(b) when assigned standard treatment,

$$\mathrm{pr}\{Y(\omega_i, 0) \mid U(\omega_i), R(\omega_i, 0)\} = \mathrm{pr}\{Y(\omega_i, 0) \mid U(\omega_i)\}.$$

Assumption 1 is related to missing-data mechanisms discussed by Baker (1994) in the context of incomplete covariates.

In the randomised trial, Assumption 1 implies that

$$\mathrm{pr}(Y_i, R_i \mid U_i, Z_i) = \mathrm{pr}(Y_i \mid U_i, Z_i)\,\mathrm{pr}(R_i \mid U_i, Z_i),$$

where, here and in the sequel, the probability measure is the one induced by the random sampling and random assignment mechanism of the trial.

To address complications because the true compliance covariate is missing, we also make Assumption 2.

Assumption 2: *Compound exclusion restriction for never-takers*. If, for individual ω, $U(\omega) = 0$, then

$$\begin{bmatrix} Y(\omega, 0) \\ R(\omega, 0) \end{bmatrix} = \begin{bmatrix} Y(\omega, 1) \\ R(\omega, 1) \end{bmatrix}.$$

Since for the never-takers the treatment actually received would be the same no matter what their treatment assignment, the intervention of assignment within the study is arguably of little relevance to them. Consequently, Assumption 2 asserts that, for the never-takers, who are defined by the covariate $U(\omega) = 0$, there is no effect of assignment on either their outcomes $Y(\omega, z)$ or their response behaviours $R(\omega, z)$. Although Assumption 2 may not be true, it is expected to hold approximately in double-blind trials or often when the outcome is measured long after final exposure to the new treatment for the never-takers has occurred. Assumption 2 is closely related to 'exclusion restriction' assumptions in the traditional instrumental variables approach (Durbin, 1954; Goldberger, 1972; Angrist et al., 1996), also used in biomedical applications, e.g. Baker & Lindeman (1994) and Sommer & Zeger (1991), and which apply to outcomes, Y. Our use of 'compound' implies that the exclusion restriction applies to both the values of Y and the missingness of Y.

An immediate consequence of Assumption 2 is that, in the randomised trial,

$$\mathrm{pr}\left\{\begin{pmatrix} Y_i \\ R_i \end{pmatrix} \middle| U_i = 0, Z_i = 0\right\} = \mathrm{pr}\left\{\begin{pmatrix} Y_i \\ R_i \end{pmatrix} \middle| U_i = 0, Z_i = 1\right\}.$$

Under Assumptions 1 and 2, the outcomes Y_i and response indicators R_i are generally

correlated in the standard treatment arm because no U_i is observed there. That is, for those units assigned standard treatment, there is no set of observed data on which to condition to make nonresponse independent of outcome. In the terminology of Rubin (1976), the missing outcomes $\{Y_i : R_i = 0\}$ are generally not missing at random, and so the nonresponse is nonignorable.

Under Assumptions 1 and 2, we next show that the respondent-based intention-to-treat estimator for $\bar{Y}_1 - \bar{Y}_0$ is generally biased even under H_0. Consistent estimation for $\bar{Y}_1 - \bar{Y}_0$ under Assumptions 1 and 2 will be discussed in § 3.

2·3. *Bias of respondent-based intention-to-treat estimator for intention-to-treat effect*

From the observed data (2·1), standard respondent-based intention-to-treat analysis disregards the actual treatment data $\{D_i\}$ and constructs an estimator for $\bar{Y}_1 - \bar{Y}_0$ based on the respondents within each assignment arm using

$$\tilde{Y}_1^{\text{obs}} - \tilde{Y}_0^{\text{obs}}, \tag{2·3}$$

the difference in sample average outcomes among respondents at each assignment arm:

$$\tilde{Y}_z^{\text{obs}} := \sum_{i=1}^{n} Y_i R_i I(Z_i = z) \Big/ \sum_{i=1}^{n} R_i I(Z_i = z),$$

where $I(.)$ is the indicator function. For technical reasons, assume $E(Y_i^2)$ is finite, and let $\bar{Y}_z^{\text{obs}} := E(Y_i | R_i = 1, Z_i = z)$, the probability limit of $\tilde{Y}_z^{\text{obs}}$ for $z = 0, 1$.

Under Assumptions 1 and 2, however, estimator (2·3) is biased in our setting because, generally, $\bar{Y}_1^{\text{obs}} - \bar{Y}_0^{\text{obs}} \neq \bar{Y}_1 - \bar{Y}_0$ even under H_0. To show this, let

$$\bar{Y}_{u,z} := E(Y_i | U_i = u, Z_i = z), \quad \bar{R}_{u,z} := E(R_i | U_i = u, Z_i = z),$$

the mean outcomes and response probabilities, respectively, within levels of the true compliance covariate and assignment arms, and let $\bar{U} := E(U_i)$, the proportion of compliers. Finally, let $\bar{R}_z := E(R_i | Z_i = z)$, the marginal response rates within assignment arms. We then have the following result.

RESULT 1. *Under Assumptions* 1, 2, *and* H_0,

$$\bar{Y}_1^{\text{obs}} - \bar{Y}_0^{\text{obs}} = \frac{\bar{U}(1 - \bar{U})\bar{R}_{0,1}}{\bar{R}_1 \bar{R}_0} (\bar{R}_{1,1} - \bar{R}_{1,0})(\bar{Y}_{1,1} - \bar{Y}_{0,1}).$$

Proof. Under Assumption 2, the compound exclusion restriction, there is no effect of assignment for the never-takers, $(\bar{Y}_{0,0}, \bar{R}_{0,0}) = (\bar{Y}_{0,1}, \bar{R}_{0,1})$, and so we will denote by $(\bar{Y}_{0,1}, \bar{R}_{0,1})$ the never-taker average outcome and response probability for both treatment assignment arms. Under Assumption 1, latent ignorability, then

$$\begin{aligned} \bar{Y}_z^{\text{obs}} &= E\{E(Y_i | R_i = 1, Z_i = z, U_i) | R_i = 1, Z_i = z\} \\ &= \text{pr}(U_i = 1 | R_i = 1, Z_i = z)\bar{Y}_{1,z} + \text{pr}(U_i = 0 | R_i = 1, Z_i = z)\bar{Y}_{0,1}. \end{aligned} \tag{2·4}$$

From Bayes' theorem, we have

$$\text{pr}(U_i = 1 | R_i = 1, Z_i = z) = \frac{\bar{R}_{1,z}}{\bar{R}_z} \bar{U}, \quad \text{pr}(U_i = 0 | R_i = 1, Z_i = z) = \frac{\bar{R}_{0,1}}{\bar{R}_z} (1 - \bar{U}). \tag{2·5}$$

With $\bar{U} > 0$ we now assume that H_0 is true or, equivalently, $\bar{Y}_{1,1} = \bar{Y}_{1,0}$, by the com-

pound exclusion restriction. If we use (2·5) in (2·4), and with the mixture $\bar{R}_z = \bar{U}\bar{R}_{1,z} + (1-\bar{U})\bar{R}_{0,1}$, Result 1 follows after some algebra. □

In Result 1, $\bar{Y}_1^{\text{obs}} - \bar{Y}_0^{\text{obs}}$ is not zero unless either the response rates for compliers, $\bar{R}_{1,1}$ and $\bar{R}_{1,0}$, are the same, or the mean outcomes between compliers and never-takers, $\bar{Y}_{1,1}$ and $\bar{Y}_{0,1}$, are the same, or $\bar{R}_{0,1} = 0$ so that all never-takers are nonrespondents. Consequently, under Assumptions 1 and 2 and the null hypothesis H_0, the respondent-based intention-to-treat estimator (2·3) is generally biased for the intention-to-treat effect. Although intention-to-treat analysis is generally promoted, in contrast to naive and biased 'as-treated' and 'per-protocol' analyses, to avoid the heterogeneity of differential characteristics between compliers and never-takers, it is this same heterogeneity that may render the missing outcomes nonignorable and hence respondent-based intention-to-treat analysis incorrect. Better procedures, however, do exist. In § 3 we construct an estimator that is randomisation-based consistent for $\bar{Y}_1 - \bar{Y}_0$ under Assumptions 1 and 2.

3. Estimation under latent ignorability and compound exclusion

Each assignment arm comprises a mixture of never-takers and compliers, so

$$\bar{Y}_1 = (1-\bar{U})\bar{Y}_{0,1} + \bar{U}\bar{Y}_{1,1}, \quad \bar{Y}_0 = (1-\bar{U})\bar{Y}_{0,0} + \bar{U}\bar{Y}_{1,0}.$$

Under the compound exclusion restriction, $\bar{Y}_{0,1} = \bar{Y}_{0,0}$, and therefore (2·2) can be written as

$$\bar{Y}_1 - \bar{Y}_0 = \bar{U}(\bar{Y}_{1,1} - \bar{Y}_{1,0}), \tag{3·1}$$

where $\bar{Y}_{1,1} - \bar{Y}_{1,0}$ is the effect of assignment on the compliers; see also Angrist et al. (1996). In the new treatment arm, all true compliance covariates U_i are observed and equal in value to D_i. By virtue of the latent ignorability assumption, the stratified estimator $\tilde{Y}_{1,1} := \sum Y_i R_i D_i Z_i / \sum R_i D_i Z_i$ is then consistent for $\bar{Y}_{1,1}$ by the law of large numbers. The estimation of $\bar{U}$ and $\bar{Y}_{1,0}$ in (3·1) requires only easily estimable quantities.

The key idea in estimation is to realise first that, by randomisation, the mixing proportions of never-takers, $U_i = 0$, and compliers, $U_i = 1$, are directly estimable from the new treatment assignment arm where the covariates U_i are fully observed. Moreover, because of the compound exclusion restriction, for never-takers response R_i is identical under both assignment arms. Hence, in the new treatment arm, $Z_i = 1$, the units can be partitioned into the following: compliers with $Z_i = 1$, $U_i = 1$; responding never-takers with $Z_i = 1$, $U_i = 0$ and $R_i = 1$; and nonresponding never-takers with $Z_i = 1$, $U_i = 0$ and $R_i = 0$. Thus the proportions of these three types in the population are directly estimable from observed data in the new treatment arm. Moreover, the mean outcomes for responding never-takers, assigned new or standard treatment, $\bar{Y}_{0,1}^{\text{obs}}$, are directly estimable in the new treatment arm. Including latent ignorability implies that $\bar{Y}_{0,1}^{\text{obs}} = \bar{Y}_{0,1}$. Latent ignorability analogously would allow direct estimation of the compliers' mean outcome $\bar{Y}_{1,0}$ when assigned standard treatment if the covariates U_i were observed in that arm. Although this specific information is not available, the mixing proportions of compliers, responding never-takers and nonresponding never-takers along with $\bar{Y}_{0,1}$ and the observed $\bar{Y}_0^{\text{obs}}$ are enough to identify the compliers' mean outcome $\bar{Y}_{1,0}$.

The following lemma, proved in the Appendix, allows a method of moments estimator to be constructed for $\bar{Y}_{1,0}$ that can be considered an extension of the econometric instrumental variables estimator (Angrist et al., 1996).

LEMMA 1. *Under Assumptions* 1, 2, *and with* $\bar{U} > 0$,

$$\bar{Y}_{1,0} = \frac{\bar{Y}_0^{\text{obs}} \bar{R}_0 - \bar{Y}_{0,1} \bar{R}_{0,1}(1-\bar{U})}{\bar{R}_0 - \bar{R}_{0,1}(1-\bar{U})}. \tag{3·2}$$

The estimator $\tilde{Y}_{1,0}$ we consider for $\bar{Y}_{1,0}$ is obtained by using the sample analogues for the quantities in the right-hand side of (3·2). Specifically, define $\tilde{U} := \sum D_i Z_i / \sum Z_i$ and

$$\begin{aligned} \tilde{R}_{0,1} &:= \frac{\sum R_i(1-D_i)Z_i}{\sum (1-D_i)Z_i}, \quad \tilde{Y}_{0,1} := \frac{\sum Y_i R_i (1-D_i) Z_i}{\sum R_i (1-D_i) Z_i}, \\ \tilde{R}_0 &:= \frac{\sum R_i (1-Z_i)}{\sum (1-Z_i)}, \quad \tilde{Y}_{1,0} := \frac{\tilde{Y}_0^{\text{obs}} \tilde{R}_0 - \tilde{Y}_{0,1} \tilde{R}_{0,1}(1-\tilde{U})}{\tilde{R}_0 - \tilde{R}_{0,1}(1-\tilde{U})}. \end{aligned} \tag{3·3}$$

Also define $V_{u,z} := \text{var}(Y_i \mid U_i = u, Z_i = z)$, for z and u in $\{0, 1\}$, and

$$V_0^{\text{obs}} := \text{var}(Y_i \mid R_i = 1, Z_i = 0).$$

The following result, proved in the Appendix using the delta method, forms a basis for inference about $\bar{Y}_1 - \bar{Y}_0$.

RESULT 2. *Under Assumptions* 1, 2 *and with* $\bar{U} > 0$,

$$n^{\frac{1}{2}}(\tilde{Y}_{1,0} - \bar{Y}_{1,0}) \to N\left(0, \sum_{k=1}^{5} v_k \delta_k^2\right),$$

in distribution as $n \to \infty$, *where*

$$v = \left[\frac{\bar{U}(1-\bar{U})}{\text{pr}(Z_i = 1)}, \frac{V_{0,1}}{\text{pr}\{Z_i R_i (1-D_i) = 1\}}, \frac{\bar{R}_{0,1}(1-\bar{R}_{0,1})}{\text{pr}\{Z_i(1-D_i) = 1\}}, \right.$$
$$\left. \frac{\bar{R}_0(1-\bar{R}_0)}{\text{pr}(Z_i = 0)}, \frac{V_0^{\text{obs}}}{\text{pr}\{R_i(1-Z_i) = 1\}} \right]',$$

$$\delta = [-\bar{R}_0 \bar{R}_{0,1}(\bar{Y}_0^{\text{obs}} - \bar{Y}_{0,1}) w^2, -\bar{R}_{0,1}(1-\bar{U}) w,$$
$$\bar{R}_0(\bar{Y}_0^{\text{obs}} - \bar{Y}_{0,1})(1-\bar{U}) w^2, -\bar{R}_{0,1}(\bar{Y}_0^{\text{obs}} - \bar{Y}_{0,1})(1-\bar{U}) w^2, \bar{R}_0 w]',$$

$$w = \{\bar{R}_0 - \bar{R}_{0,1}(1-\bar{U})\}^{-1}.$$

By standard results, $n^{\frac{1}{2}}(\tilde{Y}_{1,1} - \bar{Y}_{1,1}) \to N(0, q)$ in distribution as $n \to \infty$, where $q = V_{1,1}/\text{pr}(Z_i D_i R_i = 1)$. Furthermore, it can be shown that the two standardised estimators of interest, $n^{\frac{1}{2}}(\tilde{Y}_{1,0} - \bar{Y}_{1,0})$ and $n^{\frac{1}{2}}(\tilde{Y}_{1,1} - \bar{Y}_{1,1})$, are asymptotically independent, by a Taylor series expansion and Slutsky's theorem. By defining $\tilde{V}_{0,1}$, $\tilde{V}_{1,1}$ and $\tilde{V}_0^{\text{obs}}$ to be the usual sample estimates for $V_{0,1}$, $V_{1,1}$, and V_0^{obs} respectively, and with $\tilde{\delta}$, $\tilde{v}$ and $\tilde{q}$ the resulting sample analogues of δ, v and q, it is a direct consequence of Result 2 and Slutsky's theorem that

$$n^{\frac{1}{2}}\{(\tilde{Y}_{1,1} - \tilde{Y}_{1,0}) - (\bar{Y}_{1,1} - \bar{Y}_{1,0})\} \left(\tilde{q} + \sum_{k=1}^{5} \tilde{v}_k \tilde{\delta}_k^2\right)^{-\frac{1}{2}} \to N(0, 1)$$

in distribution as $n \to \infty$. Finally, if we let $\tilde{Y}_1 - \tilde{Y}_0 := \tilde{U}(\tilde{Y}_{1,1} - \tilde{Y}_{1,0})$, it can be easily shown using the Taylor expansion in the Appendix that

$$n^{\frac{1}{2}}\{(\tilde{Y}_1 - \tilde{Y}_0) - (\bar{Y}_1 - \bar{Y}_0)\} \left\{\tilde{U}^2 \tilde{q} + \tilde{v}_1(\tilde{Y}_{1,1} - \tilde{Y}_{1,0} - \tilde{U}\tilde{\delta}_1)^2 + \tilde{U}^2 \sum_{k=2}^{5} \tilde{v}_k \tilde{\delta}_k^2\right\}^{-\frac{1}{2}} \to N(0, 1) \tag{3·4}$$

in distribution as $n \to \infty$. Therefore, if we use the estimator $\tilde{\bar{Y}}_1 - \tilde{\bar{Y}}_0$, confidence intervals for the intention-to-treat estimand, $\bar{Y}_1 - \bar{Y}_0$, can be constructed based on the normal approximation (3·4). In the following section we examine some finite sample properties of this estimator, and compare it with the standard intention-to-treat estimator (2·3) under various conditions. Although under our assumptions the estimator (2·3) is theoretically inconsistent for $\bar{Y}_1 - \bar{Y}_0$, we investigate the extent of this problem in finite samples, first under hypothetical conditions that follow Assumptions 1 and 2 in § 4·1, and under certain deviations from Assumption 1 in § 4·2.

4. Simulation study

4·1. *Numerical results under both latent ignorability and compound exclusion*

In each of the 12 conditions of Table 1, $n = 500$ individuals are randomised to either the standard or the new treatment arm with $\text{pr}(Z_i = 1) = 0{\cdot}5$ and, independently, compliance covariates U_i are simulated as Bernoulli random variables with probability $\bar{U}$. Outcomes Y_i are simulated from normal distributions with standard deviation equal to two, conditionally on covariates U_i and treatment assignment arms Z_i. The parameters that, for simplicity, we fix across experimental conditions are

(i) the average outcomes for never-takers and compliers at the standard treatment arm, $\bar{Y}_{0,0} = 0$, which equals $\bar{Y}_{0,1}$ by Assumption 2, and $\bar{Y}_{1,0} = 3$ respectively, and
(ii) their response probabilities at the new treatment arm $\bar{R}_{0,1} = \bar{R}_{1,1} = 0{\cdot}5$, for simplicity.

The parameters that we vary are

(iii) the proportion of never-takers, that is $1 - \bar{U}$;
(iv) the average outcome effect for compliers, that is $\bar{Y}_{1,1} - \bar{Y}_{1,0} = 0$ or 1;
(v) the response probability for the compliers at the standard treatment arm; that is $\bar{R}_{1,0} = 0{\cdot}5$ gives ignorable nonresponse, i.e. missing at random, or $\bar{R}_{1,0} = 0{\cdot}8$, which gives nonignorable nonresponse, i.e. not missing at random.

In each case we report the induced value, $\bar{Y}_1 - \bar{Y}_0$, and the ratio of expected observable outcome to true, marginal expected outcome in the standard treatment arm, $\bar{Y}_0^{\text{obs}}/\bar{Y}_0$, as a measure of overall deviation from ignorability within arm $z = 0$.

Table 1 reports coverage rates of nominal 95% confidence intervals for $\bar{Y}_1 - \bar{Y}_0$, and mean squared errors for estimators calculated using (i) the new procedure based on the normal approximation (3·4), labelled ITT^{IV} because it uses a generalisation of the instrumental variables procedure as derived by Angrist et al. (1996), and (ii) the respondent-based intention-to-treat statistic (2·3) with a two-sample t confidence interval, labelled ITT^{obs}. For the special case where the treatment effect is zero within both never-takers and compliers, coverage rates for the null value are also reported using the following: 'as-treated', which compares respondents based on treatment actually received, $\{Y_i : D_i R_i = 1\}$ versus $\{Y_i : (1 - D_i)R_i = 1\}$, using a two-sample t confidence interval; and 'per-protocol', which only considers respondents who adhered to protocol in the study and compares them based on treatment received, $\{Y_i : D_i R_i = 1\}$ versus $\{Y_i : (1 - Z_i)R_i = 1\}$, using a two-sample t confidence interval. Simulations reported in Table 1 are calculated over 10 000 datasets for each experimental condition, to ensure a standard error for the coverage rates no larger than 0·5%.

The as-treated procedure grossly undercovers the true null values, and the per protocol procedure is nearly as dreadful.

In the special case where the outcomes are missing at random, both ITT^{IV} and ITT^{obs}

Table 1. *Inference for* $\bar{Y}_1 - \bar{Y}_0$ *under latent ignorability and compound exclusion restriction: coverage of nominal* 95% *intervals and mean squared error*

	$\bar{R}_{1,0} = 0{\cdot}5$ (MAR)						$\bar{R}_{1,0} = 0{\cdot}8$ (NMAR)					
	$\bar{Y}_{1,1} - \bar{Y}_{1,0} = 0$			$\bar{Y}_{1,1} - \bar{Y}_{1,0} = 1$			$\bar{Y}_{1,1} - \bar{Y}_{1,0} = 0$			$\bar{Y}_{1,1} - \bar{Y}_{1,0} = 1$		
$1 - \bar{U}$	0·2	0·3	0·4	0·2	0·3	0·4	0·2	0·3	0·4	0·2	0·3	0·4
$\bar{Y}_1 - \bar{Y}_0$	0·0	0·0	0·0	0·8	0·7	0·6	0·0	0·0	0·0	0·8	0·7	0·6
$\bar{Y}_0^{obs}/\bar{Y}_0$	1·00	1·00	1·00	1·00	1·00	1·00	1·08	1·13	1·18	1·08	1·13	1·18
Coverage												
as treated	6·1	0·0	0·0	—	—	—	22·1	3·1	0·0	—	—	—
per protocol	46·3	16·8	4·7	—	—	—	66·6	37·3	15·7	—	—	—
ITT^{obs}	94·6	94·4	95·2	95·1	94·4	95·1	88·5	83·8	80·7	89·1	85·5	82·6
ITT^{IV}	94·9	95·1	95·8	95·1	95·1	95·7	95·3	95·0	95·0	95·0	95·2	95·4
MSE												
ITT^{obs}	0·06	0·06	0·06	0·06	0·07	0·07	0·07	0·10	0·12	0·08	0·10	0·12
ITT^{IV}	0·06	0·07	0·07	0·06	0·07	0·08	0·04	0·04	0·04	0·04	0·04	0·04

Asymptotically, the properties of the procedures are not exactly invariant in location shifts of $\bar{Y}_{1,1} - \bar{Y}_{1,0}$.
MAR, missing at random; NMAR, not missing at random

give good coverage. Also, since in this case noncompliance is not critically important, the efficiency that can be gained by capitalising on the compound exclusion restriction with our procedure trades off against the uncertainty about the various additional components that the new procedure estimates, as the mean squared errors indicate.

When, however, the missing outcome mechanism, induced by latent ignorability, is not ignorable, ITT^{obs} considerably undercovers the intention-to-treat effect, even when it is zero, with decreasing coverage as the proportion of never-takers $(1 - \bar{U})$ increases, and is inaccurate even with 80% compliance. In contrast, ITT^{IV} maintains adequate coverage in this case. Moreover, as suggested by the mean squared errors in these settings, ITT^{IV} can have substantially better overall accuracy relative to ITT^{obs} under nonignorability.

4·2. *Numerical results with deviations from Assumption* 1

In Table 2 we compare ITT^{IV} to ITT^{obs} with certain deviations from the latent ignorability of Assumption 1. We only simulate results here under the null hypothesis and under compound exclusion, so we first generate treatment assignments Z_i, compliance covariates U_i and outcomes Y_i conditional on Z_i and U_i, as in the conditions of Table 1 for the null hypothesis. Then, response indicators R_i are generated as in the six conditions of H_0 of Table 1, for each combination of U_i and Z_i, except for the group of compliers in the standard treatment arm. For this group, we create deviations from Assumption 1 by generating response indicators R_i independently of outcomes Y_i conditionally on a covariate G_i imperfectly correlated with U_i. Note that $Y_i(0)$ is a covariate in the sense that it does not change with assignment, and so, for simplicity, we choose G_i to be the indicator for whether or not $Y_i(0)$ exceeds the population average outcome for compliers at $z = 0$, $\bar{Y}_{1,0}$. We generate R_i independently of Y_i conditionally on G_i and with probabilities

$$\text{pr}(R_i = 1 \mid U_i = 1, Z_i = 0, G_i = k) = p_k \quad (k = 0, 1).$$

The odds ratio $r := p_1(1 - p_0)/\{p_0(1 - p_1)\}$ is a simple measure of deviation from Assumption 1. The values of r and $\bar{R}_{1,0}$, the marginal response rate for compliers at $z = 0$, determine the probabilities p_1 and p_0 by a simple relation. We take the parameters in

Table 2. *Inference for* $\bar{Y}_1 - \bar{Y}_0$ *under the deviations from Assumption* 1 *described in* § 4·2: *coverage of nominal* 95% *intervals and mean squared error*

	$1-\bar{U}=0{\cdot}2$				$1-\bar{U}=0{\cdot}3$				$1-\bar{U}=0{\cdot}4$			
	Odds ratio r				Odds ratio r				Odds ratio r			
	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{4}{3}$	2	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{4}{3}$	2	$\frac{1}{2}$	$\frac{3}{4}$	$\frac{4}{3}$	2
					$\bar{R}_{1,0}=0{\cdot}5$							
Coverage												
ITT^{obs}	88·3	93·8	93·7	88·7	90·1	94·0	94·4	90·5	92·0	94·4	93·9	91·3
ITT^{IV}	87·8	93·8	94·2	90·2	89·7	93·9	95·3	92·5	91·3	94·3	95·6	94·3
MSE												
ITT^{obs}	0·08	0·06	0·06	0·06	0·08	0·06	0·06	0·08	0·08	0·07	0·07	0·08
ITT^{IV}	0·08	0·06	0·07	0·07	0·08	0·07	0·07	0·09	0·08	0·07	0·08	0·10
					$\bar{R}_{1,0}=0{\cdot}8$							
Coverage												
ITT^{obs}	93·5	91·0	85·0	80·9	90·3	87·4	80·3	75·9	86·5	83·8	78·2	72·6
ITT^{IV}	93·3	94·7	94·7	93·4	93·5	94·7	94·5	94·1	94·5	94·8	95·1	94·5
MSE												
ITT^{obs}	0·05	0·06	0·08	0·10	0·07	0·08	0·11	0·13	0·09	0·10	0·13	0·15
ITT^{IV}	0·04	0·04	0·04	0·04	0·04	0·04	0·04	0·04	0·04	0·04	0·04	0·04

Assumption 1 holds when $r = 1$; see Table 1.

Table 2 as in the null hypothesis of Table 1 for $\bar{U}$, $\bar{R}_{u,z}$ and $\bar{Y}_{u,z}$, but with $r = \frac{1}{2}, \frac{3}{4}, \frac{4}{3}$ and 2, to represent deviations ranging from half to twice the odds for response for above average versus below average compliers at standard treatment. Sample sizes and number of replications for each condition are as in Table 1.

For conditions with $\bar{R}_{1,0} = 0{\cdot}5$, as with Table 1, ITT^{IV} and ITT^{obs} are practically comparable to each other, although they both experience some bias for the situations with most extreme deviations from Assumption 1, $r = \frac{1}{2}$ and $r = 2$.

When the compliance covariate is associated with missingness, represented by the conditions with $\bar{R}_{1,0} = 0{\cdot}8$, we have two cases. First, values of $r < 1$ generate here deviations in the direction opposite to the bias that is apparent to an investigator who has knowledge of the compliance covariates U_i but not of G_i. Secondly, values of $r > 1$ generate deviations in the same direction as the bias generated by U_i alone. Under the small deviations of the first type, $r < 1$, in Table 2, ITT^{IV} is relatively robust, and there is some cancellation of the bias of ITT^{obs} relative to Table 1. Under the deviations of the second type, $r > 1$, the ITT^{obs} procedure performs notably poorly, whereas ITT^{IV} performs quite well.

These observations also partly indicate what is expected when small deviations from Assumption 1 exist in more than one cell defined by (U, Z). If these deviations are mostly in one direction, the performances of ITT^{IV} and ITT^{obs} will either become more comparable or will further favour ITT^{IV}, relative to analogous cases where Assumption 1 holds. In more likely cases, where small deviations from Assumption 1 can be in different directions for different cells defined by (U, Z), then generally there will be a smaller overall component of additional bias either to cancel or further increase the bias that is due to U_i alone. In these cases, the comparisons between ITT^{IV} and ITT^{obs} are expected to be more analogous to those under Assumption 1.

More generally, in practice, the comparison between ITT^{IV} and ITT^{obs} depends on the unknown, underlying parameter values. In a given study, such a simulation can be con-

ducted around plausible underlying parameters so that the comparison between the two procedures can be more specific to the investigation in question.

5. APPLICATION TO CENSORED DATA

Often incomplete outcomes arise from censoring of a time-to-event, a survival time. In this section we show how the ideas of § 2 can be applied to identify the survival distributions under our framework of all-or-none compliance. For the case of discrete-time survival, Baker (1998) considers a related and independently developed approach.

Let Z_i and $D(\omega_i, z)$ be as before. Now let

$$Y(\omega_i, z),\ R(\omega_i, z),\ X(\omega_i, z) := \min\{Y(\omega_i, z), R(\omega_i, z)\},\quad \Delta(\omega_i, z) := I\{Y(\omega_i, z) < R(\omega_i, z)\}$$

be, respectively, the survival and censoring times, their minimum and the indicator for censoring if individual ω_i is assigned treatment z. Now the observed data are $\{Z_i, D_i, X_i, \Delta_i : i = 1, \ldots, n\}$, where $X_i = X(\omega_i, Z_i)$ and $\Delta_i = \Delta(\omega_i, Z_i)$.

We focus on identifying the intention-to-treat survival curves $S_z(y) := \mathrm{pr}(Y_i > y \mid Z_i = z)$, for times $y > 0$ and assignment arms $z = 0, 1$. In this setting, intention-to-treat analysis based on the Kaplan–Meier estimator (Kaplan & Meier, 1958) assumes independence between censoring and survival times, e.g. Lee et al. (1991). If we argue as in §§ 2·2–2·3, when censoring can be created by loss to follow-up or other potentially selective reasons, it is more plausible if independent censoring is assumed only conditionally on the compliance covariates U_i, as with latent ignorability, rather than unconditionally. This assumption, however, renders the intention-to-treat Kaplan–Meier estimator generally inconsistent for the survival curves. Nevertheless, under the compound exclusion restriction, the survival curves $S_z(y)$ are identifiable.

Let $S_{u,z}(y) := \mathrm{pr}(Y_i > y \mid U_i = u, Z_i = z)$. Under latent ignorability, we have that $S_{0,0}(y) = S_{0,1}(y)$ and

$$S_0(y) = (1 - \bar{U})S_{0,1}(y) + \bar{U}S_{1,0}(y),\quad S_1(y) = (1 - \bar{U})S_{0,1}(y) + \bar{U}S_{1,1}(y). \tag{5·1}$$

Assuming that all $S_{u,z}(y)$ are absolutely continuous, the corresponding net hazard functions $\lambda_{u,z}(y)$ exist. By virtue of latent ignorability, after stratification on the observed covariates U_i within the $z = 1$ arm, the Kaplan–Meier estimators for $S_{0,1}(y)$ and $S_{1,1}(y)$ are consistent. Estimating the remaining curve of interest, $S_{1,0}(y)$, is more subtle because it requires indirect conditioning on the missing covariates U_i within the $z = 0$ arm. To proceed, we express the associated hazard function, $\lambda_{1,0}(y)$, in terms of easily estimable quantities:

$$F_{u,z}(y) := \mathrm{pr}(X_i \leqslant y, \Delta_i = 1 \mid U_i = u, Z_i = z),\quad F_z(y) := \mathrm{pr}(X_i \leqslant y, \Delta_i = 1 \mid Z_i = z),$$
$$H_{u,z}(y) := \mathrm{pr}(X_i \geqslant y \mid U_i = u, Z_i = z),\quad H_z(y) := \mathrm{pr}(X_i \geqslant y \mid Z_i = z),$$

for $u, z \in \{0, 1\}$. Using the law of total probability and the compound exclusion restriction, we have

$$F_0(y) = (1 - \bar{U})F_{0,1}(y) + \bar{U}F_{1,0}(y),\quad H_0(y) = (1 - \bar{U})H_{0,1}(y) + \bar{U}H_{1,0}(y).$$

By standard results, the hazard function sought, $\lambda_{1,0}(y)$, is

$$\begin{aligned}\lambda_{1,0}(y) &= \frac{dF_{1,0}(y)}{dy}\{H_{1,0}(y)\}^{-1} \\ &= \left\{\frac{dF_0(y)}{dy} - (1 - \bar{U})\frac{dF_{0,1}(y)}{dy}\right\}\{H_0(y) - (1 - \bar{U})H_{0,1}(y)\}^{-1}.\end{aligned} \tag{5·2}$$

We may now use empirical estimates for the right-hand side of (5·2) to estimate the survival curve $S_{1,0}(y)$. Define the stochastic processes

$$N_i(y) := I(X_i \leqslant y, \Delta_i = 1), \quad Q_i(y) := I(X_i \geqslant y) \quad (i = 1, \ldots, n)$$

and

$$N_{0,1}(y) := \sum_i N_i(y) I(U_i = 0) I(Z_i = 1), \quad Q_{0,1}(y) := \sum_i Q_i(y) I(U_i = 0) I(Z_i = 1),$$
$$N_0^{\text{obs}}(y) := \sum_i N_i(y) I(Z_i = 0), \quad Q_0(y) := \sum_i Q_i(y) I(Z_i = 0).$$

Under a mild regularity condition, a uniformly consistent estimator for $S_{1,0}(y)$ is

$$\hat{S}_{1,0}(y) = \exp\left[- \int_0^y \frac{dN_0^{\text{obs}}(u)/n_0 - dN_{0,1}(u)/n_1}{Q_0(u)/n_0 - Q_{0,1}(u)/n_1} \right], \tag{5·3}$$

where n_0, n_1 are the numbers of individuals randomised to standard and new treatment arm respectively. An outline of the proof is given in the Appendix. This result establishes identifiability of the intention-to-treat survival curves (5·1) under latent ignorability and compound exclusion even in situations where the intention-to-treat Kaplan–Meier estimators are inconsistent.

6. Final remarks

An alternative approach to the analytical methods discussed here could be to construct bounds for the treatment effect (Robins, 1989; Manski, 1990; Balke & Pearl, 1997). The least favourable bounds are not generally very informative, e.g. Imbens & Rubin (1997), and some additional subject-matter constraints typically are needed to interpret the data from such a study using this approach, e.g. Robins & Greenland (1996). A more general approach is to study sensitivity of the results to deviations from the posited assumptions, e.g. Rosenbaum & Rubin (1983), where the extreme results correspond to bounds.

Our methods, discussed for binary observed compliance for each subject in the new treatment arm, can be directly extended to allow for such compliance behaviour in the standard treatment arm as well, simply by extending Assumptions 1 and 2. For situations where all-or-none observed compliance is not realistic, some further assumptions will be required to address this problem. Barnard et al. (1998) describe a template for applying the ideas presented here to a large social science dataset.

Acknowledgement

The authors are grateful to Stuart Baker and Marvin Zelen for fruitful discussions, to the editor and to anonymous referees for suggestions that improved the focus and presentation of the arguments, and to the U.S. National Science Foundation for financial support.

Appendix

Proofs

Proof of Lemma 1. Using (2·5) to solve (2·4) for $\bar{Y}_{1,z}$, we obtain for $z = 0$

$$\bar{Y}_{1,0} = \frac{\bar{Y}_0^{\text{obs}} \bar{R}_0 - \bar{Y}_{0,1} \bar{R}_{0,1} (1 - \bar{U})}{\bar{R}_{1,0} \bar{U}}.$$

If we use $\bar{R}_0 = \bar{R}_{1,0}\bar{U} + \bar{R}_{0,1}(1-\bar{U})$ to replace $\bar{R}_{1,0}\bar{U}$, Lemma 1 follows. □

Proof of Result 2. Using the expressions in (3·2) for $\bar{Y}_{1,0}$ and $\tilde{Y}_{1,0}$, respectively, to expand $n^{\frac{1}{2}}(\tilde{Y}_{1,0} - \bar{Y}_{1,0})$ in a Taylor series, we obtain

$$n^{\frac{1}{2}}(\tilde{Y}_{1,0} - \bar{Y}_{1,0}) = \delta_1 n^{\frac{1}{2}}(\tilde{U} - \bar{U}) + \delta_2 n^{\frac{1}{2}}(\tilde{Y}_{0,1} - \bar{Y}_{0,1}) + \delta_3 n^{\frac{1}{2}}(\tilde{R}_{0,1} - \bar{R}_{0,1}) + \delta_4 n^{\frac{1}{2}}(\tilde{R}_0 - \bar{R}_0) + \delta_5 n^{\frac{1}{2}}(\tilde{Y}_0^{\text{obs}} - \bar{Y}_0^{\text{obs}}) + o_p(1), \quad \text{(A·1)}$$

where $\delta = (\delta_1, \delta_2, \delta_3, \delta_4, \delta_5)'$ are as defined in Result 2. Under the definitions in §2·3 and the expressions in (3·3), it is not difficult to show that, marginally, all of

$$n^{\frac{1}{2}}(\tilde{U} - \bar{U}), \quad n^{\frac{1}{2}}(\tilde{Y}_{0,1} - \bar{Y}_{0,1}), \quad n^{\frac{1}{2}}(\tilde{R}_{0,1} - \bar{R}_{0,1}), \quad n^{\frac{1}{2}}(\tilde{R}_0 - \bar{R}_0), \quad n^{\frac{1}{2}}(\tilde{Y}_0^{\text{obs}} - \bar{Y}_0^{\text{obs}}),$$

in (A·1), are asymptotically normal with mean 0 and variances $v = (v_1, v_2, v_3, v_4, v_5)'$, respectively, as defined in Result 2. Also, it can be shown that the five quantities in the last expression are asymptotically independent, by further expanding the denominators of their defining expressions, see § 3, in Taylor series and using Slutsky's theorem. Result 2 then follows from (A·1). □

Outline of proof of consistency of $\hat{S}_{1,0}(y)$ *for* $S_{1,0}(y)$. We define the empirical estimates

$$\hat{F}_0(y) = \frac{N_0^{\text{obs}}(y)}{n_0}, \quad \hat{F}_{0,1}(y) = \frac{N_{0,1}(y)}{(1-\bar{U})n_1}, \quad \hat{H}_0(y) = \frac{Q_0(y)}{n_0}, \quad \hat{H}_{0,1}(y) = \frac{Q_{0,1}(y)}{(1-\bar{U})n_1}. \quad \text{(A·2)}$$

Using $S_{1,0}(y) = \exp\{-\int_0^y \lambda_{1,0}(u)\,du\}$, and expressions (5·2) and (A·2), we obtain the estimator $\hat{S}_{1,0}(y)$ of (5·3). By standard results, $\hat{H}_0(y)$ and $\hat{H}_{0,1}(y)$ are uniformly consistent for $H_0(y)$ and $H_{0,1}(y)$ respectively. Then, provided there exists a fixed time y_m such that the curves $S_{u,z}(y)$ and $\text{pr}(R_i > y \mid U_i = u, Z_i = z)$ are bounded away from 0 for all y in $[0, y_m]$ and $u, z \in \{0, 1\}$, uniform consistency of $\hat{S}_{1,0}(y)$ for $S_{1,0}(y)$ in $[0, y_m]$ follows from Theorem 2.4.2 of Gill (1986), i.e. Lenglart's inequality, and a triangle inequality.

REFERENCES

ANGRIST, J. D., IMBENS, G. W. & RUBIN, D. B. (1996). Identification of causal effects using instrumental variables (with Discussion). *J. Am. Statist. Assoc.* **91**, 444–72.

BAKER, S. G. (1994). Regression analysis of grouped survival data with incomplete covariates: nonignorable missing data and censoring mechanisms. *Biometrics* **50**, 821–6.

BAKER, S. G. (1998). Analysis of survival data from a randomized trial with all-or-none compliance: estimating the cost-effectiveness of a cancer screening program. *J. Am. Statist. Assoc.* **93**, 929–34.

BAKER, S. G. & LINDEMAN, K. S. (1994). The paired availability design: a proposal for evaluating epidural analgesia during labor. *Statist. Med.* **13**, 2269–78.

BALKE, A. & PEARL, J. (1997). Bounds on treatment effects from studies with imperfect compliance. *J. Am. Statist. Assoc.* **92**, 1171–6.

BARNARD, J., DU, J., HILL, J. & RUBIN, D. (1998). A broader template for analyzing broken randomized experiments. *Sociol. Meth. Res.* **27**, 285–318.

DURBIN, J. (1954). Errors in variables. *Rev. Int. Statist. Inst.* **22**, 23–32.

FARWELL, J. R., LEE, Y., HIRTZ, D., SULZBACHER, S., ELLENBERG, J. & NELSON, K. (1990). Phenobarbital for febrile seizures—Effects on intelligence and on seizure recurrence. *New Engl. J. Med.* **322**, 364–9.

GILL, R. D. (1986). *Censoring and Stochastic Integrals*, Mathematical Centre Tracts 124, 3rd ed. Amsterdam: Mathematisch Centrum.

GOLDBERGER, A. S. (1972). Structural equation methods in the social sciences. *Econometrica* **40**, 979–1001.

GORDON, M. E. & SCHECHTMAN, K. B. (1990). Patient compliance; a covariate in randomized clinical trials. In *Proc. Biopharm. Sect., Am. Statist. Assoc.*, pp. 24–33. Alexandria, VA: American Statistical Association.

HOLLAND, P. (1986). Statistics and causal inference. *J. Am. Statist. Assoc.* **81**, 945–70.

IMBENS, G. W. & RUBIN, D. B. (1997). Bayesian inference for causal effects in randomized experiments with noncompliance. *Ann. Statist.* **25**, 305–27.

KAPLAN, E. L. & MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Am. Statist. Assoc.* **53**, 457–81.

KLEINMAN, K. P., IBRAHIM, J. G. & LAIRD, N. M. (1998). A Bayesian framework for intent-to-treat analysis with missing data. *Biometrics* **54**, 265–78.

Lee, Y. J., Ellenberg, J. H., Hirtz, D. G. & Nelson, K. B. (1991). Analysis of clinical trials by treatment actually received: is it really an option? *Statist. Med.* **10**, 1595–605.

Little, R. J. A. & Rubin, D. B. (1987). *Statistical Analysis with Missing Data.* New York: Wiley.

Little, R. & Yau, L. (1996). Intent-to-treat analysis for longitudinal studies with drop-outs. *Biometrics* **52**, 1324–33.

Manski, C. F. (1990). Non-parametric bounds on treatment effects. *Am. Econ. Rev., Papers Proc.* **80**, 319–23.

Mark, S. D. & Robins, J. M. (1993). Estimating the causal effect of smoking cessation in the presence of confounding factors using a rank preserving structural failure time model. *Statist. Med.* **12**, 1605–28.

Neyman, J. (1923). On the application of probability theory to agricultural experiments: essay on principles, Section 9. Transl. (1990) *Statist. Sci.* **5**, 465–80.

Reuben, D., Borok, G., Wolde-Tsadik, G., Ershoff, D., Fishman, L., Ambrosini, V., Liu, Y., Rubenstein, L. & Beck, J. (1995). Randomized trial of comprehensive geriatric assessment in the care of hospitalized patients. *New Engl. J. Med.* **332**, 1345–50.

Robins, J. M. (1989). The analysis of randomized and non-randomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. In *Health Service Research Methodology: A Focus on AIDS*, Ed. L. Sechrest, H. Freeman and A. Bailey, pp. 113–59. Washington, DC: National Center for Health Services Research, U.S. Public Health Service.

Robins, J. M. & Greenland, S. (1994). Adjusting for differential rates of prophylaxis therapy for PCP in high- versus low-dose AZT treatment arms in an AIDS randomized trial. *J. Am. Statist. Assoc.* **89**, 737–49.

Robins, J. M. & Greenland, S. (1996). Comment on paper by J. D. Angrist, G. W. Imbens and D. B. Rubin. *J. Am. Statist. Assoc.* **91**, 456–8.

Rosenbaum, P. R. & Rubin, D. B. (1983). Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *J. R. Statist. Soc.* B **45**, 212–8.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psychol.* **66**, 688–701.

Rubin, D. B. (1976). Inference and missing data. *Biometrika.* **63**, 581–92.

Rubin, D. B. (1978). Bayesian inference for causal effects. *Ann. Statist.* **6**, 34–58.

Rubin, D. B. (1998). More powerful randomization-based *p*-values in double-blind trials with noncompliance. *Statist. Med.* **17**, 371–87.

Schechtman, K. & Gordon, M. (1988). The relationship between treatment compliance and the design and analysis of clinical trials. In *Proc. Biopharm. Sect., Am. Statist. Assoc.*, pp. 182–7. Alexandria, VA: American Statistical Association.

Sheiner, L. B. & Rubin, D. B. (1995). Intentions-to-treat analysis and the goal of clinical trials. *Clin. Pharmacol. Therapy* **56**, 6–10.

Sommer, A. & Zeger, S. (1991). On estimating efficacy from clinical trials. *Statist. Med.* **10**, 45–52.

The Coronary Drug Project Research Group (1980). Influence of adherence to treatment and response of cholesterol on mortality in the coronary drug project. *New Engl. J. Med.* **303**, 1038–41.

Zelen, M. (1979). A new design for randomized clinical trials. *New Engl. J. Med.* **300**, 1242–5.

Zelen, M. (1990a). Randomized consent designs for clinical trials: an update. *Statist. Med.* **9**, 645–56.

Zelen, M. (1990b). Discussion of presidential address: 'Biostatistical collaboration in medical research' by J. H. Ellenberg. *Biometrics* **46**, 28–9.

[*Received July* 1997. *Revised November* 1998]