Practical Properties of Some Structural Mean Analyses of the Effect of Compliance in Randomized Trials

Krista Fischer-Lapp, PhD, and Els Goetghebeur, PhD
Institute of Mathematical Statistics, University of Tartu, Tartu, Estonia (K. F.-L.) and Applied Mathematics and Informatics, University of Ghent, Ghent, Belgium (E.G.)

ABSTRACT: We can use the structural mean model (SMM) to estimate the mean effect of dose-timing patterns of active treatment actually taken by patients in a randomized placebo-controlled trial. An SMM therefore models the expected difference between a patient’s potential response on the treatment arm and potential response on the placebo arm as a function of observed compliance on the treatment arm and baseline predictors. It accounts for the possibly selective nature of noncompliance without needing to model that aspect directly. It nevertheless enjoys the intention-to-treat property of protecting the α level when we are testing the hypothesis of no treatment effect.

In the presence of selective compliance, classical regression methods lead to inconsistent and seriously biased estimates of the effects of treatment actually taken. The SMM is designed to reduce these problems. This paper studies selectivity and addresses some practical properties of the SMM estimator. Specifically, we use a blood pressure trial to explore the precision of the estimates in practical cases. We also compare mean squared errors (MSEs) of an SMM and the ordinary least-squares (OLS) estimator. We study the effect of baseline covariates on the precision of the SMM estimator and describe the potential role of a run-in period in this regard. Control Clin Trials 1999;20:531–546 © Elsevier Science Inc. 1999

KEY WORDS: Causal inference, estimating equations, noncompliance, placebo-controlled clinical trial, structural nested mean models

1. INTRODUCTION

An intention-to-treat analysis, the standard approach for analyzing a randomized placebo-controlled clinical trial, estimates the effect of assigning patients to the active treatment instead of placebo. This effect estimate is unbiased, assuming there are no missing data in the observed outcome variables and valid randomization. Given compliance data in such a trial, one may also want to estimate the expected effect of actually observed dose-timing patterns of active treatment. The difficulty is that noncompliance is a postrandomization
variable. In a simple association analysis, such as regression of response on observed compliance in the treatment arm, confounding between compliance and baseline prognosis is likely to occur. Early work recognized this fact [1-3], but did not yet possess the newer methodologic tools developed over the last decade. Special methods to overcome the bias problem have been proposed but may gain consistency and asymptotic unbiasedness at the cost of low precision [4, 5].

Efron and Feldman have approached the problem by assuming that the expected treatment-free outcome, conditional on a given quantile of the compliance distribution, is the same in different arms of the trial [6]. Albert and DeMets have investigated robustness of this approach, finding that severe biases can occur when the assumptions of comparable complier groups are violated [7].

Many authors have developed methods for the case when both compliance and outcome are binary. For a control arm without any treatment and for binary (“all or nothing”) compliance on the treatment arm, Sommer and Zeger have estimated consistently the average effect of treatment in the subgroup of compliers [8]. Angrist, Imbens, and Rubin used the instrumental variables approach with some additional assumptions to find a causal parameter for the placebo-controlled trial with binary compliance and binary outcomes [5]. Cuzick et al. explicitly allow for contaminators in the population, that is, persons on the placebo arm who take the experimental treatment [9]. Goetghebeur and Molenberghs have estimated efficacy at observed dose levels for ordered categorical compliance with binary outcomes and a monotone dose-response [10].

Robins and Tsiatis have developed a semiparametric rank preserving structural failure time model for the analysis of survival data [11]. Later, Robins proposed related structural nested mean models to find causal parameters in a quite general complex setting with repeated continuous outcomes [12]. His methodology used estimating equations (M-estimators) based on a parametric model for the treatment benefit and left unspecified the possibly selective nature of noncompliance in terms of baseline prognosis. He has proved that the resulting estimator is consistent and asymptotically normally distributed. For such linear causal models, Goetghebeur and Lapp have derived closed-form expressions for parameter estimates and their variance [13].

The main purpose of this paper is to investigate the practical potential and limitations of such estimators in the case of continuous outcome and continuous treatment compliance measures (e.g., percentage of prescribed drug actually taken).

We consider two possible outcome variables for each subject in a two-arm placebo-controlled trial: one observed when assigned to treatment and the other when assigned to placebo. The expected difference between these two potential responses at fixed levels of treatment compliance and baseline covariates is considered as a measure of causal effect: the benefit of treatment compared to that of placebo.

As a motivating example, we consider a placebo-controlled double-blind trial on the reduction of blood pressure. The trial has randomized 300 patients with mild hypertension in the United Kingdom over one of two treatments or placebo. (A detailed analysis appears in Goetghebeur and Lapp [13]).
prescribed dosing regimen consists of one tablet daily at breakfast. Patients are scheduled to have five clinic visits, each a fortnight apart. The period until visit 3 is designed as the run-in (and washout) period, during which all subjects receive placebo tablets. Baseline characteristics (including assigned treatment, sex, height, weight, BMI, and age) and blood pressure are recorded at visit 1 while previous antihypertensive treatment is discontinued at that point. Blood pressure is measured again at visit 2 and at visit 3, when the randomized treatment is started. After ineligible cases have been excluded, patients are randomized with equal probability to nebivolol, atenolol, or placebo. Visit 3 also indicates the start of MEMS monitoring: a device in the pill container measures exact times at which patients open and close the container, one of the most reliable compliance measures currently available.

In this trial we wish to find the effect of the percentage of assigned active dose that subjects actually took (assessed by MEMS) on the reduction of diastolic blood pressure. Classical methods like OLS regression would yield biased estimates, because treatment compliance is likely to be selective: an association may exist between a patient’s treatment compliance and potential placebo response. We investigate the issue of selectivity both in general terms and by the example of this trial. Analysis by SMM shows that even quite small sample sizes are sufficient to detect an important effect of treatment dose actually taken and its interaction with a baseline covariate.

Next, we investigate the causes and magnitude of bias when we use OLS regression to estimate the same effects. An evaluation of the ratio of MSEs of OLS and SMM estimators will shed some light on when one can prefer one estimator to another.

We also study in more detail the effect of baseline predictors on the precision of the SMM estimates. The results indicate that including good predictors not only for placebo response but also for treatment compliance will reduce the sample size needed to obtain sufficient power for detecting the effect of compliance on benefit of treatment. When one uses SMM to estimate several structural parameters, predictability of treatment compliance is essential for their identifiability. Past compliance is often a good predictor for a patient’s treatment compliance. One can sometimes obtain data on compliance during a run-in period on placebo. To increase the precision of the structural parameter estimates, one may then choose to extend the run-in period rather than increase the total sample size.

2. FORMULATION OF THE MODEL

We consider the following potential data for each patient $i$:

- $Y_i^T (Y_i^P)$: response to the assigned active treatment (placebo). We can think of the response from each patient that would follow each possible flip of the randomization coin, regardless of which is actually observed or not.
- $C_i$: a vector containing the compliance measures to the assigned active treatment.

1 MEMS is short for Medication Event Monitoring System, a device developed for measuring compliance with drug regimens.
Xi: a vector of baseline characteristics (e.g., sex, age, weight, height, and some clinical measures of the patient’s health status at the beginning of the trial).

For subjects in the trial, either \((Y_i^T, C_i^T)\) or \(Y_i^P\) is missing. We let \(R_i\) be the indicator of treatment assignment, equal to 1 (0) for treatment (placebo) arm patients. We aim to estimate \(E(Y_i^T - Y_i^P|C_i^T, X_i)\), the expected benefit of treatment, given treatment compliance and baseline characteristics.

Denote by \(Z_i^T\) a vector containing all measured (time-dependent) covariates thought to “explain” the shift \(Y_i^T - Y_i^P\). Possible components are a summary of treatment compliance \(C_i^T\) and its interaction with baseline covariates. We assume that a linear regression model explains the treatment effect:

\[ Y_i^T - Y_i^P = \beta_i Z_i^T + \epsilon_i, \quad \text{where } E(\epsilon_i|X_i, Z_i^T) = 0. \] (1)

The \(\epsilon\) distribution is not further parameterized. Model 1, which involves latent variables, is a special case of the structural nested mean model discussed by Robins [12]. We call this model the structural mean model (SMM).

### 3. SELECTIVE COMPLIANCE: BIAS OF CLASSICAL METHODS

Note that

\[ E(Y_i^T - Y_i^P|Z_i^T, X_i) = E(Y_i^T|Z_i^T, X_i) - E(Y_i^P|Z_i^T, X_i). \]

For the observed response \(Y_i = Y_i^T R_i + Y_i^P (1 - R_i)\), model (1) implies

\[ Y_i = [E(Y_i^T|Z_i^T, X_i) + \epsilon_i^T] + \beta_i Z_i^T R_i + \epsilon_i R_i, \] (2)

where \(\epsilon_i^T = Y_i^P - E(Y_i^P|Z_i^T, X_i)\) and hence \(E(\epsilon_i^T + \epsilon_i R_i|X_i, Z_i^T, R_i) = 0\).

In general, the first term \(E(Y_i^T|Z_i^T, X_i)\) is not directly estimable from the data. For the problem at hand, the treatment compliance is reduced to \(Z\). Here we distinguish between the following identifying assumptions:

1. We call the treatment compliance nonselective if \(E(Y_i^P|Z_i^T) = E(Y_i^P)\) and selective otherwise. A typical case of selective compliance occurs when health status affects treatment compliance.
2. We call the treatment compliance explainable if \(E(Y_i^P|Z_i^T, X_i) = E(Y_i^P|X_i)\). This case is particularly interesting when compliance is selective but some baseline covariates \(X_i\) can explain the association between treatment compliance and placebo response.

Suppose treatment compliance is explainable and we can apply a linear regression model to predict the placebo response, that is, \(Y_i^P = \alpha + \delta X_i + \epsilon_i^P\) where \(E(\epsilon_i^P|X_i^T, Z_i^T) = 0\). Then, the model for observed response is

\[ Y_i = \alpha + \delta^T X_i + \beta_i Z_i^T R_i + \epsilon_i R_i + \epsilon_i^P \quad \text{where } E(\epsilon_i^P + \epsilon_i R_i|X_i, Z_i^T, R_i) = 0. \] (3)

Here the parameters are consistently estimated by ordinary or weighted least-squares regression, yielding the most efficient estimator if the error term is normally distributed. If treatment compliance is not explainable but we still use model 3 to estimate \(\beta_{S_i}\), the resulting estimate will be biased.

We shall investigate the effect of selectivity in the case of the blood pressure trial. Although we do not assume that compliance in the placebo arm is exchangeable with compliance in the active treatment arm, it is tempting to use
Table 1  Average Diastolic Blood Pressure at Visit 5 in Two Randomized Groups

<table>
<thead>
<tr>
<th>Randomized Group</th>
<th>Group Average DBP</th>
<th>Subgroup Average DBP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>90.4 (−11.2)</td>
<td>93 (−8.6) 89.8 (−11.8)</td>
</tr>
<tr>
<td>Placebo</td>
<td>99.0 (−3.1)</td>
<td>93.5 (−9.3) 100.5 (−1.4)</td>
</tr>
</tbody>
</table>

Abbreviation: DBP, diastolic blood pressure.

* The number in brackets is the average change of DBP between visits 3 and 5.
* C is the percentage of the prescribed dose actually taken.

compliance data on the placebo arm to get some idea of possible selectivity of the treatment compliance. We shall not use data on placebo compliance in the actual analysis, however.

The intent-to-treat analysis (Table 1) indicates that the mean diastolic blood pressure at the end of the trial as well as the mean blood pressure reduction differ between both arms, showing a significant beneficial effect of treatment at the level of 5%. On the other hand, the marginal association between response and compliance (Table 1 and Figure 1) shows almost no association on the treatment arm but some indication of an association on the placebo arm: better compliers tend to have higher diastolic blood pressure at the end of the trial. If treatment and placebo compliance were exchangeable, this relation would suggest the presence of a selection effect. Under those conditions, we could check explainable noncompliance through inspection of the plot of residuals from the linear regression of response on baseline characteristics in the placebo arm against placebo compliance. The data still show a clear slope; hence we would expect the OLS estimate of the structural parameter $\beta_3$ to be biased.

We must caution against overinterpretation of these plots, as they constitute merely a description of observed associations. Indeed, a patient’s compliance to placebo and active treatment need not be the same, even though (while unable to check on this) we may expect them to be correlated. Hence, Figure 1 need not reflect selectivity, as the latter concerns itself strictly with the association between active treatment compliance and latent placebo response.

4. THE SMM ESTIMATOR

4.1 Estimator, Derived From Estimating Equations

Estimability of the SMM parameters becomes clear when we note that (1) implies

$$E(Y_i | \beta, Z_i | X_i) = E(Y_i | X_i).$$

(4)

Given the randomization assumption

$$[Y_i, Y_i, X_i, C_i] \perp R_0,$$

(5)

we can find a consistent estimator under model 1 as a solution to a set of unbiased estimating equations.
Figure 1  Scatter plots with fitted simple OLS regression lines: (a) diastolic blood pressure at visit 5 (DBP5) versus compliance (% of prescribed dose) in the active treatment group; (b) DBP5 versus compliance in the placebo group; and (c) OLS residuals from the fitted model $Y_i^p = \alpha + \delta' X_i + \epsilon_i^p$ in the placebo group versus compliance.

$$\sum_i g(R_i, X_i)[Y_i - \beta_i Z_i'R_i - q(X_i)] = 0, \text{ with } E[g(R_i, X_i)|X_i] = 0.$$  

In fact, these equations equate weighted averages of $Y_i^p - \beta_i Z_i'R_i$ in the treatment arm to weighted averages of $Y_i^p$ in the placebo arm. Robins proposed the functions $g_{opt}(R_i, X_i)$ and $q_{opt}(X_i)$ [12], being optimal forms for $g(R_i, X_i)$ and $q(X_i)$ in the sense that they reach the semi-parametric efficiency bound (defined for example by Newey [14]). These functions $g_{opt}(R_i, X_i)$ and $q_{opt}(X_i)$, involve the unknown quantities $E(Y_i^p|X_i)$ and $E(Z_i'R_i|X_i)$. Robins proposed to estimate them iteratively; instead, we estimate them from linear regressions on the basis of the placebo or treatment group data, respectively.

If $X$ and $Z_i'$ are matrices with the $i$th row $(1; X_i)$ and $Z_i'$, respectively, $Y$ is a column vector with $i$th element $Y_i$, the observed response of the $i$th patient, and $R$ is the diagonal matrix with $i$th diagonal element $R_i$, then we can write the resulting SMM estimator in matrix form as
Effect of Compliance

\[ \hat{\beta}_s = (\hat{g}_{opt} R Z^T)^{-1} \hat{g}_{opt} (Y - \hat{q}_{opt}), \]  

(6)

with \( \hat{q}_{opt} = X' [I - R)^{-1} X' (I - R) Y \) and \( \hat{g}_{opt} = (I - 2R) X' RX)^{-1} X' R Z^T \) being our estimates of the optimal weights.

As shown by Robins, this estimator is asymptotically normally distributed and asymptotically unbiased. A consistent variance estimator is

\[ \frac{1}{n} \sum_i (\hat{g}_{opt} (H_i - q_{opt})) | \hat{g}_{opt}' (H_i - q_{opt}) |^2 (\hat{g}_{opt} R Z^T)^{-1}, \]

where \( H_i = Y_i - \hat{\beta} Z_i R_i \).

4.2 An Alternate Formulation

An equivalent expression of the SMM estimator (6) is

\[ \hat{\beta}_s = (\hat{Z}^T R Z^T)^{-1} \hat{Z}^T R (\hat{Y}^T - \hat{Y}^T \hat{\beta}), \]

(7)

with \( \hat{Z}^T = X' RX)^{-1} X' R Z^T, \)

\( \hat{Y}^T = X' [I - R)^{-1} X' (I - R) Y \)

and \( \hat{Y}^T = X' RX)^{-1} X' R Y \).

Thus the SMM estimate is an OLS estimate, regressing \( \hat{Y}^T - \hat{Y}^T \hat{\beta} \) on \( \hat{Z} \). Note that this is regression through the origin (it is unrealistic to assume that \( Z \) involves a constant column, because there can be no treatment effect at zero dose).

4.3 Justification in a Simple Case

Suppose an SMM (1) holds, where there is a linear effect of one univariate compliance measure \( C^T_i \), thus:

\[ Y_i^T - Y_i^T = \beta C^T_i + \epsilon_i^{[1]}, \]

with \( E[\epsilon_i^{[1]}|X_i, C^T_i] = 0, \)

(8)

and a linear regression model holds for the placebo response:

\[ Y_i^T = \alpha + \delta X_i + \epsilon_i^{[2]}, \]

with \( E[\epsilon_i^{[2]}|X_i] = 0. \)

(9)

Here \( \epsilon_i^{[2]} \) may depend on \( C^T_i \), allowing for selectivity. Finally, let compliance \( C^T_i \) satisfy:

\[ C^T_i = \gamma X_i + \epsilon_i^{[3]} \]

with \( E[\epsilon_i^{[3]}|X_i] = 0. \)

(10)

Now

\[ Y_i^T = \delta X_i + \beta \gamma X_i + \epsilon_i^{[3]} + \epsilon_i^{[1]} + \beta \epsilon_i^{[2]}, \]

with \( E[\epsilon_i^{[1]} + \epsilon_i^{[2]} + \beta \epsilon_i^{[3]}|X_i] = 0. \)

(11)

We can fit the models (9), (10), and (11) by OLS, with corresponding predictions \( \hat{Y}_i^T, \hat{Y}_i^T, \) and \( \hat{C}_i \), unbiased estimates for \( E[Y_i^T|X_i], E[Y_i^T|X_i], \) and \( E[C_i|X_i] \) respectively. Because

\[ E[Y_i^T|X_i] - E[Y_i^T] = \beta E[C_i|X_i], \]

we can estimate \( \beta \) by regressing \( \hat{Y}_i^T - \hat{Y}_i^T \) on \( \hat{C}_i \). This turns out to be the SMM estimate (6) of \( \beta \).
Table 2  Parameter Estimates and Estimated Standard Errors of the Fitted SMM Models*

<table>
<thead>
<tr>
<th>Model</th>
<th>( Z^T_i )</th>
<th>( \hat{\beta}_i )</th>
<th>se(( \hat{\beta}_i ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>( C^T_i )</td>
<td>-7.41</td>
<td>1.84</td>
</tr>
<tr>
<td>2.</td>
<td>( C^T_i )</td>
<td>-7.61</td>
<td>1.73</td>
</tr>
<tr>
<td></td>
<td>( C^T_i W_i )</td>
<td>0.36</td>
<td>0.14</td>
</tr>
</tbody>
</table>

* Data from the blood pressure trial. \( Y_i^T - Y_i^T = Z^T_i \beta_i + \epsilon_i \), with \( C^T_i \) = percentage of prescribed active dose taken and \( W_i \) = weight - sample average weight. All available baseline covariates were used in the estimation (see Section 4.4).

The equivalence also holds in the case of a more complex SMM, involving a vector \( Z^T_i \) that may contain different treatment compliance measures and interaction terms of treatment compliance with the baseline characteristics. Notice that the error term \( \epsilon_i^{(2)} \) in model (9) may incorporate any possible selectivity mechanisms, without leading to incorrect estimates of \( \beta \). We can study the importance of baseline predictors most easily in formulation (7) of the SMM estimator. We do so in Section 6.

Given that one does not observe covariates and outcomes jointly in the regression model (1), one worries naturally about the information content. For practical purposes, the question of efficiency is very important: can we detect significant effects with a realistic sample size? Can SMM lead to more precise estimates than some classical methods, for example, a well-chosen classical regression model?

4.4 Example of the Blood Pressure Trial

Consider the trial described in Section 1. The three-arm blood pressure trial enrolled 300 patients; however, the compliance data of only 54 patients on the active treatment arm and 51 patients on the placebo arm were available. The primary reason we have such a small subset for our analysis is that some of the MEMS data were missing. The investigators’ assertion that technical problems with the MEMS device were responsible would support the assumption that data are missing completely at random. Still, in caution, we consider our analysis as an example of how we can apply the methodology rather than relying too strongly on the practical conclusions for this particular trial. Here we focus on comparing one of the active treatment arms with placebo. For a separate analysis for both treatment arms, see [13].

First we fit model (1) with \( Z^T_i = C^T_i \) = “percentage of the active dose prescribed to the treatment arm that was actually taken.” The vector of baseline predictors \( X \) consists of sex, height, weight, diastolic blood pressure at visits 2 and 3 (dbp2, dbp3), sex × dbp2, and height × dbp2. A graphical inspection of the modeling assumptions of this SMM shows that we must allow for an interaction of treatment compliance with baseline body weight: let \( Z^T_i = (C^T_i, C^T_i W_i) \), with \( W_i = \) weight - sample average (weight). This interaction effect is significant at the level of 5%; the same dose has less effect for heavier persons. We show results of the fitted models in Table 2.
5. RELATIVE EFFICIENCY OF SMM AND OLS ESTIMATORS

Although non-explainable treatment compliance leads to biased OLS estimates whereas SMM estimates are asymptotically unbiased under model (1), one could still prefer the classical estimator if it led to a much smaller variance and a smaller MSE. We investigate the precision of the two estimators for the simplest possible SMM: a linear effect in $C_i^T$ and no baseline covariates. Thus, assume that (1) holds:

$$Y_i^T - Y_i^O = \beta_s C_i^T + \epsilon_i \quad \text{with } E(\epsilon_i|C_i^T) = 0. \quad (12)$$

The SMM estimate for $\beta_s$ is:

$$\hat{\beta}_s = \left( \sum_i R_i C_i \right)^{-1} \left[ \sum_i Y_i - \frac{p}{1 - p} \sum_i (1 - R_i) Y_i \right], \quad (13)$$

where $p$ is the proportion of patients randomized to the active treatment. This $\hat{\beta}_s$ is now the difference between the average responses in two randomized groups divided by average treatment compliance in the active treatment group.

The model for observed response (2) becomes

$$Y_i = \text{E}(Y_i^O|C_i) \pm \beta_s C_i^T R_i + \epsilon_i R_i + \epsilon_i^o,$$

where $\epsilon_i^o = Y_i^O - \text{E}(Y_i^O|C_i)$ and $E(\epsilon_i^o + \epsilon_i R_i|C_i^T, R_i) = 0$.

The corresponding OLS estimate for $\beta_s$ is (if we assume, perhaps incorrectly, that $E(Y_i^O|C_i) = E(Y_i) = \mu_{yp}$)

$$\hat{\beta}_O = \frac{n \sum_i R_i C_i Y_i - \sum_i Y_i \sum_i R_i C_i}{n \sum_i R_i C_i - (\sum_i R_i C_i)^2}.$$

Suppose there is a linear selection effect of $C_i^T$, that is $Y_i^O = \alpha + \gamma C_i^O + \epsilon_i^o$, with $E(\epsilon_i^o|C_i^T) = 0$. Then the approximate bias of the OLS estimate of $\beta_s$ is

$$\gamma \left[ \frac{\sigma_{\epsilon_i^o}}{\sigma_{\epsilon_i^o}^2 + (1 - \pi) \mu_{\epsilon_i^o}^2} \right] = \rho \frac{\sigma_{\epsilon_i^o} \sigma_{\gamma \epsilon_i^o}}{\sigma_{\epsilon_i^o}^2 + (1 - \pi) \mu_{\epsilon_i^o}^2},$$

where $\mu_{\epsilon_i} = E(\epsilon_i)$, $\sigma_{\gamma}^2 = \text{Var}(\gamma)$, $\sigma_{\epsilon_i}^2 = \text{Var}(Y_i^O)$, $\rho = \text{Cor}(Y_i^O, C_i^O)$ and $\pi = P(R_i = 1)$. The SMM estimate will be asymptotically unbiased with a small finite sample bias, as illustrated in Section 7.

We compare the precision of the two estimators in terms of asymptotic MSE. We investigate how the ratio of their MSEs depends on selectivity of treatment compliance and on the correlation between $Y_i^O$ and $C_i^O$ in the case of a linear dependence. Setting marginal means and variances equal to their observed value in the example, Figure 2 shows the asymptotic bias of the OLS estimator and ratio of the asymptotic MSEs as a function of Cor($Y_i^O, C_i$), the Pearson correlation between placebo response and treatment compliance.

Even if a relatively small selection effect is present, Cor($Y_i^O, C_i$) > 0.14, SMM leads to estimates with lower MSE. This finding remains true when baseline covariates are involved, as long as the treatment compliance is not fully explained by $X$.

As an aside, note that Var($Y_i^O$) = 120 > Var($Y_i^O - \beta_s C_i^O$) = 50, the latter being equal to Var($Y_i^O + \epsilon_i$). This relationship highlights a feature of the model: $E(\epsilon_i Y_i^O) \neq 0$ in general. Here the change of DBP on the placebo arm has a much
Figure 2  Plots of the asymptotic bias of the OLS estimator and of the ratio $\text{MSE}(\hat{b}_S)/\text{MSE}(\hat{b}_O)$, where $E(\hat{C}_\delta) = 0.86$, $\text{Var}(\hat{C}_\delta) = 0.1$, $n = 100$, $\text{Var}(Y_i^\delta) = 120$, $\text{Var}(Y_i^\delta - \hat{\beta}_S \hat{C}_\delta) = 50$.

larger variance than on the treatment arm. This may be due to the fact that treatment was reducing the blood pressure of most patients to “normal limits,” but on the placebo we see a wider range of negative and positive changes. The fit of our model suggests that random errors in equation (6) account for the fact that persons with positive placebo response (increased blood pressure) would benefit more from the treatment than those who would have a spontaneous reduction of blood pressure while on placebo.

6. POWER CONSIDERATIONS OF THE SMM ESTIMATOR

6.1 The Role of Baseline Predictors

Expression (7) reveals that baseline predictors play a very important role in SMM estimation. If no baseline predictors for $\hat{C}_\delta$ are available, the estimate of $\hat{\beta}$ is given by 13. When patients are randomized to two groups with equal probability, this equals the sample average of $Y_i^\delta - Y_i^\delta$ divided by average treatment compliance. Next, we study precision increase when one non-interacting baseline covariate is included in estimation.

6.2 Asymptotic Variance of a Univariate $\hat{\beta}_S$ in the Case of a Single Baseline Covariate

We study the variance of the SMM estimator in the simple case, where only one baseline covariate is present and that covariate has a linear mean effect on $C_i$. We estimate univariate dose effect as follows:

$$Y_i^\delta - Y_i^\delta = \beta_S C_i^\delta + \epsilon_i$$

with $E(\epsilon_i | C_i^\delta, X_i) = 0$. (14)

Now the asymptotic variance of the SMM estimator $\hat{\beta}_S$ (under randomization and homoscedasticity assumptions) satisfies:

$$n \text{Var}(\hat{\beta}_S | X) = \frac{1}{\left( \mu^2 + \rho^2 \sigma^2 \right)} \left( \frac{\sigma^2}{\pi} + \frac{\sigma^2}{1 - \pi} \right)$$

Here $\mu^2$ and $\sigma^2$ are the mean and variance of $C_i^\delta$; $\sigma^2 = \text{Var}(Y_i^\delta - \hat{\beta}_S C_i^\delta | X_i) = \text{Var}(Y_i^\delta + \epsilon_i | X_i)$, $\sigma^2 = \text{Var}(Y_i^\delta | X_i)$, and $\rho = \text{Cor}(X_i, C_i^\delta)$. 

Table 3  Results of the Fitted SMM $Y_i^T - Y_i^C = C_i^T\beta_i + \epsilon_i$, with $E[\epsilon_i|X, C_i^T] = 0$ on the Data of the Blood Pressure Trial$^a$

<table>
<thead>
<tr>
<th>$X$</th>
<th>$\hat{\text{Cor}}(X, Y_i^C)$</th>
<th>$\hat{\text{Cor}}(X, C_i^T)$</th>
<th>$\hat{\beta_i}$</th>
<th>$\hat{SE}(\hat{\beta_i})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>-8.92</td>
<td>2.13</td>
</tr>
<tr>
<td>dbp$^c$</td>
<td>-0.4</td>
<td>0.12</td>
<td>-9.39</td>
<td>2.01</td>
</tr>
<tr>
<td>dbp, height, dbp2 $\times$ height</td>
<td>0.11$^b$</td>
<td>0.41$^b$</td>
<td>-7.91</td>
<td>2.01</td>
</tr>
<tr>
<td>dbp2, height, dbp2 $\times$ height, dbp3</td>
<td>0.39$^b$</td>
<td>0.41$^b$</td>
<td>-7.53</td>
<td>1.86</td>
</tr>
<tr>
<td>All</td>
<td>0.58$^b$</td>
<td>0.45$^b$</td>
<td>-7.41</td>
<td>1.84</td>
</tr>
</tbody>
</table>

$^a$ Pearson correlations of baseline predictors with placebo response and treatment compliance, and parameter estimates $\hat{\beta_i}$ with estimated standard error $\hat{SE}(\hat{\beta_i})$.

$^b$ Correlations between $YP_i^C$ and the linear predictor for $YP_i^C$, based on $X$.

$^c$ dbp2 and dbp3 are diastolic blood pressures at visits 2 and 3, respectively.

This equation shows that, in order to minimize the variance of the estimator, one needs baseline predictors that are correlated not only with $Y_i^C$ (minimizing $\sigma_2^2$ and $\sigma_2^2$), but also with $C_i^T$. Thus careful consideration of predictors at the design stage can reduce the number of patients needed.

6.3 More Than One Structural Parameter

Predictability of compliance measures from baseline covariates is necessary for identifiability of more than one structural parameter. Indeed, in (7) we need the matrix $\tilde{Z}^T \tilde{Z}^T$ to be nonsingular. If columns in $Z^T$ represent different compliance measures, then at least $k-1$ compliance measures must be predictable by at least $k-1$ independent covariates in $X$ so that resulting predictions (columns of $\tilde{Z}^T$) are also linearly independent. This implies that we may predict no more than one of the compliance measures by its mean only; different compliance measures need to be predictable by different independent baseline covariates or by independent linear combinations of the same covariates. Columns of $Z^T$ that stand for interaction terms between treatment compliance and baseline characteristics form a lesser problem here as univariately they are automatically correlated with part of $X$.

If the columns of $\tilde{Z}^T$ are linearly independent but highly correlated, we have a problem of multicollinearity: the variance of $\hat{\beta}$ will be large. We are likely to have this problem when the components of $Z^T$ are highly correlated or when we have too few baseline covariates to predict compliance measures.

6.4 Effect of Covariates in the Blood Pressure Trial

We examine the effect of covariates in the SMM analysis of the blood pressure trial. Table 3 shows results of the fitted model with univariate dose effect for five different sets of baseline predictors. In this trial, pairwise correlations between baseline covariates and $Y_i^C$ (or $C_i^T$) were quite small. Correlation with a linear predictor involving a set of covariates $X$ are shown in columns $\text{Cor}(X, Y_i^C)$ and $\text{Cor}(X, C_i^T)$ of Table 3.

By increasing the set of baseline predictors to all available covariates associated with either $Y_i^C$ or $Y_i^C$ we obtain more precise estimates. The variance decrease achieved in this way corresponds to a sample size reduction of 25%.
Table 4  Average Estimated Pearson Correlation Matrix of 1000 Simulated Datasets

<table>
<thead>
<tr>
<th></th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$Y^p$</th>
<th>$C^t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>1</td>
<td>0</td>
<td>0.71</td>
<td>0</td>
</tr>
<tr>
<td>$X_2$</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.88</td>
</tr>
<tr>
<td>$Y^p$</td>
<td>0.71</td>
<td>0</td>
<td>1</td>
<td>0.21</td>
</tr>
<tr>
<td>$C^t$</td>
<td>0</td>
<td>0.88</td>
<td>0.21</td>
<td>1</td>
</tr>
</tbody>
</table>

Using only a limited set of baseline predictors (e.g., baseline DBP and body weight) did not yield enough precision to declare the interaction between body weight and treatment compliance significant at the level of 5% (see Table 2). This illustrates the theoretical result (see Section 6.3) that good predictors are important in the estimation of more complex structural models.

6.5 Simulation Results

In the blood pressure trial, no baseline covariate was highly correlated only with treatment compliance and not with placebo response. To illustrate the practical effect of covariates that are predictors for either placebo response or treatment compliance, we simulated 1000 datasets with: $n = 200$ (100 subjects on each arm); univariate measure of treatment compliance $C^t$; two baseline covariates $X_1, X_2$; placebo response $Y^p$; and treatment response $Y^t$. We generated data to satisfy the following conditions:

1. The values of covariates $X_1$ and $X_2$ are independent normal variates with means 100 and 10, standard deviations 10 and 3, respectively.
2. Placebo response $Y^p$ satisfies $Y^p = X_1 + \epsilon^p$, where $\epsilon^p$ are independent normal variates with mean 0 and standard deviation 10.
3. Next, $C^t = 0.6 + (X_2 - 10)/10 + 0.01 \epsilon^t + \epsilon^c$, with $\epsilon^c$ independent mean 0 normal variates with standard deviation 0.1. We set to 0 and 1 values of $C^t$ that were less than 0 or greater than 1, respectively.
4. Treatment response $Y^t$ satisfies $Y^t = Y^p - 10C^t + \epsilon$, with values of $\epsilon$ being independent normal variates with mean 0 and standard deviation 2.

As a result, $X_1$ is correlated with placebo response but not with treatment compliance; $X_2$ is correlated with treatment compliance but independent of placebo response; treatment compliance is selective and non-explainable: $\text{Cor}(Y^p, C^t) = 0.21$ and $\text{Cor}(Y^p, C^t|X_1, X_2) = 0.63$ (coefficient of partial correlation given the values of $X_1$ and $X_2$). The mean and variance of treatment compliance are $\text{E}(C^t) = 0.58$ and $\text{Var}(C^t) = 0.08$. Table 4 gives the estimated correlation matrix (average over all 1000 datasets). We present in Table 5 the resulting SMM estimates (which are obtained by using different sets of baseline predictors).

Adding a covariate $X_2$ that is highly correlated with treatment compliance in model 1 without covariates reduces the variance by 16%, and adding it to the model with $X_1$ (predictor for placebo response) reduces the variance by 15%. The approximate maximum possible reduction of variance, when $\text{Cor}(C^t, X_2) = 1$, is 19.5%. In comparison to the model without baseline covariates, including both $X_1$ and $X_2$ reduces the variance by more than 50%.
Table 5  Summary of the Fitted SMM $Y_i^f - Y_i^d = C_i^f \beta_i + \epsilon_i$, with $E[\epsilon_i|X_i, C_i^f] = 0^*$

<table>
<thead>
<tr>
<th>( X )</th>
<th>( \bar{E}(\beta_i) )</th>
<th>( \sqrt{\text{Var}(\beta_i)} )</th>
<th>( \text{Var}(\beta_i) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>—</td>
<td>—10.06</td>
<td>3.44</td>
<td>11.87</td>
</tr>
<tr>
<td>( X_2 )</td>
<td>—10.06</td>
<td>3.16</td>
<td>9.96</td>
</tr>
<tr>
<td>( X_1 )</td>
<td>—10.03</td>
<td>2.54</td>
<td>6.44</td>
</tr>
<tr>
<td>( X_1, X_2 )</td>
<td>—10.05</td>
<td>2.35</td>
<td>5.51</td>
</tr>
</tbody>
</table>

*Based on 1000 simulated datasets: the average point estimate, estimates of standard error, and variance.

6.6 Importance of a Run-in Period

We can reduce sample size by using covariates correlated with treatment compliance if we measure compliance during a run-in period on placebo, even if placebo compliance does not perfectly predict treatment compliance. The gain in efficiency will then depend on the length of such run-in. As we retain more information, better predictors are possible, but going back too far in time may alter the relevance for future predictions. The optimal length will acknowledge both processes. Theoretically, by including a baseline covariate correlated with treatment compliance but not with placebo response, we can reduce the asymptotic variance of a single structural parameter \( \beta_s \) maximally by a factor \((\mu_c^2 + \sigma_e^2)/\mu_c^2 \) (if that covariate has correlation one with treatment compliance). In the blood pressure example, this factor equals 1.14. Suppose we plan a trial in which we expect treatment compliance to have the same mean and variance as in the blood pressure trial described here. Suppose also that, without using any baseline covariates, we would require a sample size of 100 patients on both arms to estimate a single SMM parameter \( \beta_s \) with sufficient power. Then, by including a covariate with correlation about 0.9 (\( \rho^2 = 0.81 \)) with treatment compliance and no correlation with placebo response, we would only need 90 patients on each arm to have the same variance for the structural parameter. If this covariate is the compliance measure during a run-in period of a specific length, we can decide whether including that run-in period costs us more than having 200 patients instead of 180. We can note that the reduction of variance depends heavily on the mean and variance of treatment compliance (here \( \mu_c = 0.86, \sigma_c^2 = 0.1 \)). If, for instance, the mean and variance of treatment compliance were \( \mu_c = 0.7 \) and \( \sigma_c^2 = 0.15 \), the same covariate would reduce the required sample size by about 20%. Although in practice this correlation might be less than 0.9, we could still expect an even bigger gain in efficiency, because compliance during a run-in period is likely to be correlated also with placebo response.

In an exploratory fashion, we looked at data from a Belgian study of blood pressure reduction with MEMS compliance measures [15]. There the correlation between compliance measures taken during the first period of the trial and compliance during the following 14 days (both expressed as a percentage of prescribed dose) reached its maximum 0.84, when data on the first 9 days of the trial were used as linear predictors. This suggests that patients tend to have similar compliance during two consecutive time periods on the same treatment.
If, however, compliance is related to treatment (side) effects and patients receive a different treatment during the second period (placebo during a run-in and active treatment later), there is an extra source of variation in compliance between different periods. Nevertheless, it is quite natural that the same unmeasurable patient characteristic influences both run-in placebo compliance and trial treatment compliance.

### 7. FINITE SAMPLE BIAS OF SMM ESTIMATES

From the matrix form of the SMM estimator (6), we find its finite sample bias:

\[
E[\hat{\beta}_s - \beta_s | X] = E[(\hat{g}'_{opt} R Z)^{-1} \hat{g}'_{opt} (Y^p - \hat{q}_{opt}) | X].
\]

This converges to 0 in probability [12]. To see the magnitude of the bias in small samples, we performed a simulation study. We generated data to satisfy the simple SMM (14) (using only one baseline covariate \(X\) in the estimation). We ran simulations under six different sets of conditions on explainability and predictability of treatment compliance, with three different sample sizes (20, 40, 100) and with probability 0.5 of a patient being randomized to treatment or placebo. Table 6 shows the empirical bias and standard errors of the simulated estimates. These results indicate that the finite sample bias of the structural mean model estimates, although theoretically present, is realistically very close to zero. On the other hand, even a very weak selection effect leads to large enough bias in the OLS estimate to produce incorrect inferences.

### 8. DISCUSSION

Our findings suggest that structural estimators may aid in the analysis of clinical trials in the presence of noncompliance. Classical regression analysis can lead to strongly biased estimates if treatment compliance is selective and...
non-explainable. In practice, non-explainability arises as an association between treatment compliance and response that is not part of the dose–response relationship, and not caused by measured baseline covariates (e.g., the effect of outcome on compliance). Such association is realistic in many trials, including the blood pressure trial analyzed here. Given a structural model, the SMM methodology leads to estimates unconfounded by these selection effects. The SMM estimator is then asymptotically normally distributed and unbiased under any selection mechanism. Furthermore, the implied test of no treatment effect retains the $\alpha$ level of the intent-to-treat analysis.

Analysis of the blood pressure trial shows that with reasonable baseline covariates, a mere sample size of about 100 (50 on each arm) enables us to detect a significant effect not only of a univariate compliance measure, but also of its interaction with a baseline covariate. The fitting algorithm has a simple matrix form and is easy to implement with flexible software such as S-PLUS (Mathsoft, Inc.). In comparison to the OLS estimators, the SMM leads to a smaller asymptotic MSE for the estimation of a single parameter under a range of realistic scenarios.

For a useful interpretation of the structural parameters, it is important that the assumed structural model is correct. Yet whether the structural (linear) model should be questioned can only be addressed through the implied equality on observables: $E(Y^T - Z^T\hat{\beta}|X) = E(Y^T|X)$. Hence, plots of $Y^T - Z^T\hat{\beta}$ versus individual $X$ variables must show the same regression line as plots of $Y^T$ versus $X$. In an earlier paper [13], we proposed such graphs to check whether one should include interaction terms between treatment compliance and baseline covariates. Alternatively, one can always fit more complex nested structural models and eliminate nonsignificant terms from them. None of the more complex models we tried in our example was found superior to the two-parameter model we presented. A systematic approach toward a sensitivity analysis in this context is the topic of further work.

The extent to which we can increase power by involving baseline predictors for placebo response and treatment compliance can be studied on the basis of formulae like (15). Investigators should consider the issue at the stage of design. One way to reduce the required sample size is to include a run-in period during which investigators take compliance measures on a placebo or the standard regimen for all patients in the trial. Increasing the length of such a run-in (up to an “optimal” length) increases the correlation between a summary of run-in compliance and treatment compliance during the trial. Thus one can trade off a prolongation of the run-in period with reduction of required sample size.

We are grateful to the Janssen Research Foundation for giving us the opportunity to explore this data set and Jef Dom for practical help and to Jamie Robins and Mārt Møls for helpful discussions. We also thank Thomas Chalmers Student Scholarship Committee of the Society for Clinical Trials for helpful comments on the paper and for awarding the first author with the Thomas Chalmers Student Scholarship Prize.

REFERENCES


