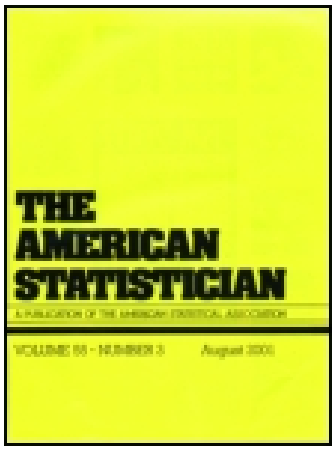

The American Statistician
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/utas20

Efficiency Study of Estimators for a Treatment Effect in a Pretest-Posttest Trial

Li Yang[a] & Anastasios A Tsiatis[a]

[a] Li Yang is a Research Statistician, SAS Institute Inc., SAS Campus Drive, R5241, Cary NC 27513 . Anastasios A. Tsiatis is Professor, Department of Statistics, Box 8203, North Carolina State University, Raleigh, NC 27695-8203 . This research was supported by grants R01 AI-31789 from the National Institute of Allergy and Infectious Disease and R01 CA-51962 from the National Cancer Institute. The authors also thank Marie Davidian for her advice on GEE models.
Published online: 01 Jan 2012.

Efficiency Study of Estimators for a Treatment Effect in a Pretest–Posttest Trial

Li YANG and Anastasios A. TSIATIS

Several possible methods used to evaluate treatment effects in a randomized pretest–posttest trial with two treatment groups are the two-sample t test, the paired t test, analysis of covariance I (ANCOVA I), the analysis of covariance II (ANCOVA II), and generalized estimating equations (GEE). The ANCOVA I includes treatment and baseline response as covariates in a linear model and ANCOVA II additionally includes an interaction term between the baseline response and treatment indicator as a covariate. The parameters in the ANCOVA I and ANCOVA II models are generally estimated using ordinary least squares. In this article, a semiparametric model, which makes no assumptions about the response distributions, is used. The asymptotic properties of the estimators derived from these five methods and their relative efficiencies are discussed under this semiparametric model. We show that all these methods yield consistent estimators for the treatment effect which have asymptotically normal distributions under the semiparametric model. The GEE and the ANCOVA II estimators are asymptotically equivalent and the most efficient. The estimators from other three methods are less efficient except under some special conditions which are outlined in the article.

KEY WORDS: Analysis of covariance; GEE model; Paired t test; Semiparametric model; Two-sample t test.

1. INTRODUCTION

In biomedical as well as other scientific fields, a pretest–posttest trial is an important and frequently employed method to evaluate treatment effects. In a pretest–posttest trial, the responses (outcomes) are measured both at baseline and at follow-up. Commonly a randomized pretest–posttest trial is used, in which subjects are randomly assigned to the two groups—the treatment group and the control group. The focus of such a trial is to test whether a treatment affects responses at follow-up. The pretest responses are also made at baseline for comparison.

Currently, there are several methods that have been proposed to estimate the treatment difference or test that there is no difference. These methods include the two-sample t test, the paired t test, and the analysis of covariance. The analysis of covariance which includes baseline response as a covariate was mentioned by Follmann (1991). Stein (1989) suggested that the analysis of covariance which includes baseline response as a covariate with different slopes for the two groups could also be used.

Each method has its own characteristics. The two-sample t test estimates the treatment effect using only the responses at follow-up. It does not use any information at baseline which may be useful in increasing efficiency if the baseline and follow-up outcomes are correlated. The paired t test, where the differences between the follow-up and baseline responses in the two treatment groups are compared, may lose some information by using an average additive relationship between the follow-up and baseline responses. The analysis of covariance has also been advocated as being more efficient than the other two methods. Feldt (1958) proved that the ANCOVA (the analysis of covariance) is more efficient than the change scores (the paired t test) under the normal distribution. A number of other articles—Pigrache and Graham (1976), Crager (1987), and Laird (1983)—discuss the ANCOVA in finite sample situations under specific distributional assumptions such as normality, equal variance or a fixed baseline effect. In addition, a conditional linear relationship between the follow-up and baseline responses is assumed in the analysis of covariance.

The objective of this article is to carefully study the large sample properties of the various estimators for treatment difference under very general conditions where only the first and second moments of the baseline and follow-up response variables are assumed to exist. The only constraint we impose on the model is that the mean and variance of the response at baseline is the same for both treatments, a condition that must hold because of randomization. Subject to the constraints of the first two moments given above, we allow the joint distribution of the responses at baseline and follow-up, conditional on treatment, to be arbitrary. Such models are often referred to as restricted moment semiparametric models.

The analysis of covariance has been separated into two cases: In addition to the treatment indicator, ANCOVA I includes only baseline response as a covariate and ANCOVA II includes both baseline response and its interaction with the treatment indicator as covariates. We also consider estimates based on the generalized estimating equation (GEE). GEEs are widely used to model the marginal dependence of multivariate outcomes in longitudinal studies since Zeger and Liang (1986) first developed this method. To the best of our knowledge the properties of GEEs have not been studied carefully in the context of pretest–posttest trials, except for a brief mention by Singer and Andrade (1997) in their discussion section.

2. LARGE SAMPLE PROPERTIES OF THE FOUR TRADITIONAL METHODS AND THE GEE METHOD

Data on N subjects from a randomized pretest–posttest trial can be represented by the independent and identically distributed random vectors $(Y_{2i}, Y_{1i}, Z_i), i = 1, \ldots, N$, where for the ith

Li Yang is a Research Statistician, SAS Institute Inc., SAS Campus Drive, R5241, Cary NC 27513 (E-mail: li.yang@sas.com). Anastasios A. Tsiatis is Professor, Department of Statistics, Box 8203, North Carolina State University, Raleigh, NC 27695-8203 (E-mail: tsiatis@stat.ncsu.edu). This research was supported by grants R01 AI-31789 from the National Institute of Allergy and Infectious Disease and R01 CA-51962 from the National Cancer Institute. The authors also thank Marie Davidian for her advice on GEE models.

individual, Y_{1i} is the response at baseline, Y_{2i} is the response at follow-up, and Z_i is the treatment indicator with 1 denoting the treatment group and 0 representing the control group. Because we are considering a randomized study, we will assume that patient i is assigned to the treatment group with probability δ independently of their baseline response Y_{1i}. Otherwise, we will allow the conditional distribution of (Y_{2i}, Y_{1i}) given Z_i to be arbitrary, subject to the existence of the first two moments. Specifically, letting $\boldsymbol{Y}_i = \begin{pmatrix} Y_{2i} \\ Y_{1i} \end{pmatrix}$ be the outcome vector, then its conditional expectation is

$$E(\boldsymbol{Y}_i|Z_i) = \boldsymbol{v}_i = \begin{pmatrix} \mu_2 + \beta Z_i \\ \mu_1 \end{pmatrix}, \quad i = 1, 2, \ldots, N. \quad (1)$$

The conditional expectation of the response at baseline does not depend on the treatment indicator because of the randomization assumption. The parameter β represents the treatment effect which is the difference between the mean follow-up response of the treatment group and that of the control group and is the parameter of primary interest.

The outcome vector is assumed to have finite second moments. Because of randomization, the variances of the baseline response by treatment are equal. No additional assumptions are made on the second moments. We parameterize the second moments as

$$\text{var}(\boldsymbol{Y}_i|Z_i = 1) = \begin{pmatrix} \sigma_{22}^{(1)} & \sigma_{12}^{(1)} \\ \sigma_{12}^{(1)} & \sigma_{11} \end{pmatrix}, \quad (2)$$

$$\text{var}(\boldsymbol{Y}_i|Z_i = 0) = \begin{pmatrix} \sigma_{22}^{(0)} & \sigma_{12}^{(0)} \\ \sigma_{12}^{(0)} & \sigma_{11} \end{pmatrix}, \quad (3)$$

where σ_{11} is the variance of the response at baseline, $\sigma_{22}^{(1)}$ is the variance of the response at follow-up in the treatment group, and $\sigma_{22}^{(0)}$ is the variance of the response at follow-up in the control group. $\sigma_{12}^{(1)}$ and $\sigma_{12}^{(0)}$ correspond to the covariance between the responses at baseline and follow-up in the treatment group and control group, respectively.

We will now study the asymptotic properties of the estimator for the treatment effect β derived for the five methods where we only assume the moment restrictions given by (1), (2), and (3).

2.1 The Two-Sample *t* Test

In the two-sample t test, the test statistic is based on the estimated treatment difference using only the data at follow-up. Let $\overline{Y}_{21} = (\sum_{i=1}^N Z_i)^{-1} \sum_{i=1}^N Z_i Y_{2i}$ and $\overline{Y}_{20} = \{\sum_{i=1}^N (1 - Z_i)\}^{-1} \sum_{i=1}^N (1-Z_i) Y_{2i}$ be the averages of responses at follow-up in the treatment and the control groups, respectively, then the estimator of the treatment effect will be $\hat{\beta}_t = \overline{Y}_{21} - \overline{Y}_{20}$. Since the estimator is a difference of two independent sample averages, the expectation is $E(\hat{\beta}_t) = E(\overline{Y}_{21} - \overline{Y}_{20}) = \beta$. A simple application of the central limit theorem yields

$$N^{\frac{1}{2}}(\hat{\beta}_t - \beta) \xrightarrow{D} N(0, \Sigma_t) \quad \text{as} \quad N \to \infty,$$

where $\xrightarrow{D}$ denotes convergence in distribution and $\Sigma_t = \frac{1}{\delta}\sigma_{22}^{(1)} + \frac{1}{(1-\delta)}\sigma_{22}^{(0)}$.

2.2 The Paired *t* Test

The paired t test uses the differences between the responses at follow-up and baseline to evaluate the treatment effect. Denote these differences as $D_i = Y_{2i} - Y_{1i}$, $i = 1, 2, \ldots, N$. The averages of these differences in the treatment and the control groups are, respectively: $\overline{D}_1 = (\sum_{i=1}^N Z_i)^{-1} \sum_{i=1}^N Z_i D_i$ and $\overline{D}_0 = \{\sum_{i=1}^N (1 - Z_i)\}^{-1} \sum_{i=1}^N (1 - Z_i) D_i$. Correspondingly, the estimator for β is $\hat{\beta}_p = \overline{D}_1 - \overline{D}_0$. As for the two-sample test, the estimator for β used in the paired t test is a difference of two independent sample averages. Hence, the expectation of $\hat{\beta}_p$ is $E(\hat{\beta}_p) = E(\overline{D}_1) - E(\overline{D}_0) = \beta$, and

$$N^{\frac{1}{2}}(\hat{\beta}_p - \beta) \xrightarrow{D} N(0, \Sigma_p) \quad \text{as} \quad N \to \infty,$$

where

$$\Sigma_p = \frac{1}{\delta}\sigma_{22}^{(1)} + \frac{1}{(1-\delta)}\sigma_{22}^{(0)} + \frac{1}{\delta(1-\delta)}\sigma_{11} - \frac{2}{\delta(1-\delta)}\{(1-\delta)\sigma_{12}^{(1)} + \delta\sigma_{12}^{(0)}\}.$$

2.3 The Analysis of Covariance I

For the analysis of covariance models, a linear model is used to model the follow-up response Y_{2i} to the baseline response Y_{1i} and the treatment indicator Z_i, where the regression parameter associated with the treatment indicator is used to represent treatment effect. Typically, the parameters are estimated using ordinary least squares (OLS). In the next two subsections we will study the large sample properties of the least squares estimator of treatment effect for the two analysis of covariance models only assuming the moment restrictions given by (1), (2), and (3).

It will be convenient to express our models in terms of centered random variables. Thus, for Y_{2i}, Y_{1i}, and Z_i, we will denote by Y_{2i}^*, Y_{1i}^*, and Z_i^*, the centered random variables $Y_{2i} - E(Y_{2i})$, $Y_{1i} - E(Y_{1i})$, and $Z_i - E(Z_i)$, and by Y_{2i}^c, Y_{1i}^c, and Z_i^c, the empirically centered variables $Y_{2i} - \bar{Y}_2$, $Y_{1i} - \bar{Y}_1$, and $Z_i - \bar{Z}$, where $\bar{Y}_2 = \sum_{i=1}^N Y_{2i}/N$, $\bar{Y}_1 = \sum_{i=1}^N Y_{1i}/N$, and $\bar{Z} = \sum_{i=1}^N Z_i/N$.

The analysis of covariance I model assumes:

$$Y_{2i}^* = \beta Z_i^* + \alpha Y_{1i}^* + \epsilon_i, \quad i = 1, 2, \ldots, N, \quad (4)$$

where the errors ϵ_i are assumed to be iid random variables with mean zero. Because centered, these variables have mean zero; hence, we do not include an intercept term in the model. Let $\boldsymbol{X}_{1i}^* = \left(Z_i^*, Y_{1i}^*\right)^T$ and $\boldsymbol{X}_{1i}^c = \left(Z_i^c, Y_{1i}^c\right)^T$, then the OLS estimators for the parameters are

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\alpha}_1 \end{pmatrix} = \left(\sum_{i=1}^N \boldsymbol{X}_{1i}^c \boldsymbol{X}_{1i}^{cT}\right)^{-1} \left(\sum_{i=1}^N \boldsymbol{X}_{1i}^c Y_{2i}^c\right). \quad (5)$$

The estimators for $\hat{\beta}_1$ and $\hat{\alpha}_1$ are the same as the OLS estimators which include the intercept term, but where the variables are not centered.

Standard asymptotic arguments could be used to prove that

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\alpha}_1 \end{pmatrix} \xrightarrow{p} \left\{E\left(\boldsymbol{X}_{1i}^* \boldsymbol{X}_{1i}^{*T}\right)\right\}^{-1} E\left(\boldsymbol{X}_{1i}^* Y_{2i}^*\right),$$

where

$$E\Big(\boldsymbol{X}^*_{1i}\boldsymbol{X}^{*T}_{1i}\Big) = \begin{pmatrix} E(Z^*_i)^2 & E(Z^*_iY^*_{1i}) \\ E(Z^*_iY^*_{1i}) & E(Y^*_{1i})^2 \end{pmatrix},$$

$E(Z^*_i)^2 = \delta(1-\delta)$, $E(Z^*_iY^*_{1i}) = 0$, $E(Y^*_{1i})^2 = \sigma_{11}$. Because $E\Big(\boldsymbol{X}^*_{1i}\boldsymbol{X}^{*T}_{1i}\Big)$ is a diagonal matrix, the least square estimator

$$\hat{\beta}_1 \xrightarrow{p} \{E(Z^*_i)^2\}^{-1}E(Z^*_iY^*_{2i}). \tag{6}$$

Since

$$E(Z^*_iY^*_{2i}) = E\{Z^*_iE(Y^*_{2i}|Z_i)\} = E\{Z^*_i(\mu_2+\beta Z_i)\} = \delta(1-\delta)\beta,$$

we can use (6) to show that $\hat{\beta}_1$ converges in probability to β, where $\beta = E(Y_{2i}|Z_i=1) - E(Y_{2i}|Z_i=0)$ is the treatment effect of interest. Similarly, we can show that $\hat{\alpha}_1$ converges to $\alpha = E(Y^*_{2i}Y^*_{1i})/E(Y^*_{1i})^2$.

To derive the asymptotic normality of the OLS estimators, standard asymptotic arguments can be used to show that

$$N^{\frac{1}{2}}\begin{pmatrix}\hat{\beta}_1-\beta\\ \hat{\alpha}_1-\alpha\end{pmatrix} = N^{-1/2}\sum_{i=1}^{N}\Big\{E\Big(\boldsymbol{X}^*_{1i}\boldsymbol{X}^{*T}_{1i}\Big)\Big\}^{-1} \times\boldsymbol{X}^*_{1i}\left(Y^*_{2i} - \boldsymbol{X}^{*T}_{1i}\begin{pmatrix}\beta\\ \alpha\end{pmatrix}\right) + o_p(1),$$

where $o_p(1)$ denotes a term that converges in probability to zero. Again, because $E\Big(\boldsymbol{X}^*_{1i}\boldsymbol{X}^{*T}_{1i}\Big)$ is a diagonal matrix,

$$N^{1/2}(\hat{\beta}_1-\beta) = N^{-1/2}\sum_{i=1}^{N}\{\delta(1-\delta)\}^{-1} \times Z^*_i(Y^*_{2i} - \beta Z^*_i - \alpha Y^*_{1i}) + o_p(1).$$

Since $Z^*_i(Y^*_{2i} - \beta Z^*_i - \alpha_1 Y^*_{1i})$ has mean zero, a simple application of the central limit theorem can be used to show that $N^{1/2}(\hat{\beta}_1-\beta)$ converges to a normal distribution with mean zero and variance equal to $\{\delta(1-\delta)\}^{-2}E\{Z^*_i(Y^*_{2i}-\beta Z^*_i-\alpha Y^*_{1i})\}^2$. After some algebra we derive the asymptotic variance of $\hat{\beta}_1$ to equal

$$\Sigma_1 = \frac{1}{(1-\delta)}\sigma^{(0)}_{22} + \frac{1}{\delta}\sigma^{(1)}_{22} + \frac{1}{\sigma_{11}\delta(1-\delta)}\big\{(1-\delta)\sigma^{(0)}_{12} + \delta\sigma^{(1)}_{12}\big\} \times\big\{(1-3\delta)\sigma^{(0)}_{12} + (3\delta-2)\sigma^{(1)}_{12}\big\}. \tag{7}$$

2.4 The Analysis of Covariance II

In the analysis of covariance II model, treatment indicator, baseline response and its interaction with the treatment indicator are included as covariates. Such a model would generally be used if there was a belief that the treatment effect may vary with baseline status. In terms of the centered variables, this model is given as

$$Y^*_{2i} = \beta Z^*_i + \alpha_1 Y^*_{1i} + \alpha_2 Z^*_iY^*_{1i} + \epsilon_i, \quad i=1,\ldots,N, \tag{8}$$

where the errors ϵ_i are assumed to be iid random variables with mean zero. Let $\boldsymbol{X}^*_{2i} = \Big(Z^*_i, Y^*_{1i}, Z^*_iY^*_{1i}\Big)^T$ and $\boldsymbol{X}^c_{2i} = \Big(Z^c_i, Y^c_{1i}, Z^c_iY^c_{1i}\Big)^T$, then the OLS estimators for the parameters are:

$$\begin{pmatrix}\hat{\beta}_2\\ \hat{\alpha}_{21}\\ \hat{\alpha}_{22}\end{pmatrix} = \left(\sum_{i=1}^{N}\boldsymbol{X}^c_{2i}\boldsymbol{X}^{cT}_{2i}\right)^{-1}\left(\sum_{i=1}^{N}\boldsymbol{X}^c_{2i}Y^c_{2i}\right). \tag{9}$$

Similar to the arguments used for the analysis of covariance I model, we can show that

$$\begin{pmatrix}\hat{\beta}_2\\ \hat{\alpha}_{21}\\ \hat{\alpha}_{22}\end{pmatrix} \xrightarrow{p} \Big\{E\Big(\boldsymbol{X}^*_{2i}\boldsymbol{X}^{*T}_{2i}\Big)\Big\}^{-1}E\Big(\boldsymbol{X}^*_{2i}Y^*_{2i}\Big),$$

where

$$E\Big(\boldsymbol{X}^*_{2i}\boldsymbol{X}^{*T}_{2i}\Big) = \begin{pmatrix} E(Z^*_i)^2 & E(Z^*_iY^*_{1i}) & E((Z^*_i)^2Y^*_{1i}) \\ E(Z^*_iY^*_{1i}) & E(Y^*_{1i})^2 & E(Z^*_i(Y^*_{1i})^2) \\ E((Z^*_i)^2Y^*_{1i}) & E(Z^*_i(Y^*_{1i})^2) & E((Z^*_i)^2(Y^*_{1i})^2) \end{pmatrix}.$$

Because of the independence of Z_i and Y_{1i} induced by randomization, $E\Big(\boldsymbol{X}^*_{2i}\boldsymbol{X}^{*T}_{2i}\Big)$ is a diagonal matrix. Consequently, it is easy to show that $\hat{\beta}_2$ converges to $\beta = E(Y_{2i}|Z_i=1) - E(Y_{2i}|Z_i=0)$, the treatment effect of interest, $\hat{\alpha}_{21}$ converges to $\alpha_1 = E(Y^*_{2i}Y^*_{1i})/E(Y^*_{1i})^2$ and $\hat{\alpha}_{22}$ converges to $\alpha_2 = E\Big(Z^*_iY^*_{1i}Y^*_{2i}\Big)/E\Big(Z^*_iY^*_{1i}\Big)^2$. At first, it may seem surprising that the estimator for the parameter β in the ANCOVA II model, which also includes an interaction term, is a consistent estimator for overall treatment effect. However, this follows from the fact that we use the centered baseline response variable which, by randomization, is independent of the treatment indicator. Therefore, the mean of the interaction term given in formula (8) is zero.

Arguments, similar to those for the analysis of covariance I model, can be used to derive asymptotic normality. However, because of the interaction term, the derivation is a little more delicate. Standard asymptotic expansions can be used to show that

$$N^{\frac{1}{2}}\begin{pmatrix}\hat{\beta}_2-\beta\\ \hat{\alpha}_{21}-\alpha_1\\ \hat{\alpha}_{22}-\alpha_2\end{pmatrix} = N^{-1/2}\sum_{i=1}^{N}\Big\{E\Big(\boldsymbol{X}^*_{2i}\boldsymbol{X}^{*T}_{2i}\Big)\Big\}^{-1} \times\left\{\boldsymbol{X}^*_{2i}\left(Y^*_{2i} - \boldsymbol{X}^{*T}_{2i}\begin{pmatrix}\beta\\ \alpha_1\\ \alpha_2\end{pmatrix}\right) \times + \begin{pmatrix}\alpha_2\delta(1-\delta)Y^*_{1i}\\ \alpha_2\sigma_{11}Z^*_i\\ 0\end{pmatrix}\right\} + o_p(1).$$

Consequently,

$$N^{1/2}(\hat{\beta}_2-\beta) = N^{-1/2}\sum_{i=1}^{N}\{\delta(1-\delta)\}^{-1} \times\{Z^*_i(Y^*_{2i}-\beta Z^*_i-\alpha_1Y^*_{1i}-\alpha_2Z^*_iY^*_{1i})+\alpha_2\delta(1-\delta)Y^*_{1i}\}+o_p(1).$$

Since $Z^*_i(Y^*_{2i} - \beta Z^*_i - \alpha_1 Y^*_{1i} - \alpha_2 Z^*_iY^*_{1i}) + \alpha_2\delta(1-\delta)Y^*_{1i}$ has mean zero, a simple application of the central limit theorem

can be used to show that $N^{1/2}(\hat{\beta}_2 - \beta)$ converges to a normal random variable with mean zero and variance equal to $\{\delta(1-\delta)\}^{-2}E\{Z_i^*(Y_{2i}^* - \beta Z_i^* - \alpha_1 Y_{1i}^* - \alpha_2 Z_i^* Y_{1i}^*) + \alpha_2\delta(1-\delta)Y_{1i}^*\}^2$. After some algebra we derive the asymptotic variance of $\hat{\beta}_2$ to equal

$$\Sigma_2 = \frac{1}{(1-\delta)}\sigma_{22}^{(0)} + \frac{1}{\delta}\sigma_{22}^{(1)} - \frac{1}{\delta(1-\delta)\sigma_{11}}\left\{(1-\delta)\sigma_{12}^{(1)} + \delta\sigma_{22}^{(0)}\right\}^2. \quad (10)$$

2.5 GEE—Generalized Estimating Equation

The theory of generalized estimating equations (GEE), developed by Zeger and Liang (1986), is widely used to model the marginal dependence of multivariate outcomes in longitudinal studies. Here, we consider the outcomes at baseline and follow-up together as a bivariate vector and only assume the restricted moment model. The estimators for the mean parameters will be obtained through the following equations:

$$\sum_{i=1}^{N} \boldsymbol{D}_i^T \boldsymbol{V}_i^{-1}(\boldsymbol{Y}_i - \hat{\boldsymbol{\upsilon}}_i) = 0, \quad (11)$$

where $\hat{\boldsymbol{\upsilon}}_i = \begin{pmatrix} \hat{\mu}_2 + \hat{\beta}_e Z_i \\ \hat{\mu}_1 \end{pmatrix}$, $\boldsymbol{V}_i = \text{var}(\boldsymbol{Y}_i | Z_i)$ is the conditional covariance matrix of the outcome vector given the treatment indicator, and $\boldsymbol{D}_i = \frac{\partial \boldsymbol{\upsilon}_i}{\partial \theta^T} = \begin{pmatrix} Z_i & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ is the derivative of $\boldsymbol{\upsilon}_i$ with respect to the mean parameters θ.

Because Equation (11) involves the second moments which are not known, the statistics from the equations will not be computable. We therefore substitute the empirical second moments $\hat{\boldsymbol{V}}_i$ for $\boldsymbol{V}_i$ in Equation (11) to get:

$$\sum_{i=1}^{N} \boldsymbol{D}_i^T \hat{\boldsymbol{V}}_i^{-1}(\boldsymbol{Y}_i - \hat{\boldsymbol{\upsilon}}_i) = 0. \quad (12)$$

The estimators from Equation (12) are

$$\begin{pmatrix} \hat{\beta}_e \\ \hat{\mu}_2 \\ \hat{\mu}_1 \end{pmatrix} = \left(\sum_{i=1}^{N} \boldsymbol{D}_i^T \hat{\boldsymbol{V}}_i^{-1} \boldsymbol{D}_i\right)^{-1} \sum_{i=1}^{N} \boldsymbol{D}_i^T \hat{\boldsymbol{V}}_i^{-1} \boldsymbol{Y}_i, \quad (13)$$

which will be asymptotically the same as the estimators from Equation (11). This follows by the asymptotic theory developed by Zeger and Liang (1986). As long as $\left(\sum_{i=1}^{N} \boldsymbol{D}_i^T \boldsymbol{V}_i^{-1} \boldsymbol{D}_i\right)^{-1}$ exists, these estimators will be consistent

$$\begin{pmatrix} \hat{\beta}_e \\ \hat{\mu}_2 \\ \hat{\mu}_1 \end{pmatrix} \xrightarrow{p} \begin{pmatrix} \beta \\ \mu_2 \\ \mu_1 \end{pmatrix} \quad \text{as} \quad N \to \infty,$$

and have an asymptotic distribution given by

$$N^{\frac{1}{2}} \begin{pmatrix} \hat{\beta}_e - \beta \\ \hat{\mu}_2 - \mu_2 \\ \hat{\mu}_1 - \mu_1 \end{pmatrix} \xrightarrow{D} N(0, \boldsymbol{\Sigma}_e) \quad \text{as} \quad N \to \infty,$$

where, $\boldsymbol{\Sigma}_e = \left(E(\boldsymbol{D}_i^T \boldsymbol{V}_i^{-1} \boldsymbol{D}_i)\right)^{-1}$. After some algebra, we derive the asymptotic variance of $\hat{\beta}_e$ to be

$$\Sigma_{e1} = \frac{1}{(1-\delta)}\sigma_{22}^{(0)} + \frac{1}{\delta}\sigma_{22}^{(1)} - \frac{1}{\delta(1-\delta)\sigma_{11}}\left((1-\delta)\sigma_{12}^{(1)} + \delta\sigma_{12}^{(0)}\right)^2. \quad (14)$$

Upon comparing (14) to (10), we note the equality of the asymptotic variance of $\hat{\beta}_e$ and $\hat{\beta}_2$.

3. ASYMPTOTIC RELATIVE EFFICIENCY COMPARISONS

We have now shown that all five estimators proposed in the previous sections are either unbiased or consistent under the semiparametric model. In our experience, we have found that some statisticians are reluctant to use the ANCOVA models because of the belief that the underlying linear relationships that these models are based on would have to hold. We have demonstrated that the estimate of treatment effect using OLS from these models are robust and give consistent asymptotically normal estimators under very general assumptions. The question of interest is then to study the relative efficiencies of these estimators under different scenarios.

Since the primary focus of pretest–posttest trials is to assess the treatment difference β in the mean follow-up response, we will consider the relative efficiency of the five estimators for β. We denote the various estimators by $\hat{\beta}_t$ from the two-sample t test, $\hat{\beta}_p$ from the paired t test, $\hat{\beta}_1$ from the analysis covariance I, $\hat{\beta}_2$ from the analysis of covariance II, and $\hat{\beta}_e$ the estimator using GEE's.

As expected, $\hat{\beta}_e$ is the most efficient estimator for β. This would have to follow since the efficient GEE estimator is also semiparametric efficient under these general moment conditions. By semiparametric efficient we mean that this estimator has the smallest variance among all consistent asymptotically normal estimators for β under the semiparametric assumptions. The general result for semiparametric efficiency of this estimator for restricted moment models was given by Chamberlain (1987) and Newey (1990). The analytic calculations for the asymptotic variance given in Section 2 verify these results. We note that the estimator $\hat{\beta}_2$, from the analysis of covariance II model, has an asymptotic variance which is equal to that of the efficient GEE estimator. Therefore, this estimator, which is based on an assumed linear relationship of the follow-up response to the baseline response, treatment indicator, and an interaction term of treatment by baseline response, is not only consistent but efficient as well.

The ANCOVA I estimator, $\hat{\beta}_1$ is based on a model which assumes a linear relationship between the follow-up response and the baseline response and treatment indicator with no interaction term. Again, the estimator $\hat{\beta}_1$ is consistent under the weaker semiparametric assumption but is not necessarily efficient. The difference between the asymptotic variance of $\hat{\beta}_1$ and that of $\hat{\beta}_e$ or $\hat{\beta}_2$ is $\Sigma_1 - \Sigma_{e1} = \frac{(2\delta-1)^2}{\delta(1-\delta)\sigma_{11}}\left\{\sigma_{12}^{(1)} - \sigma_{12}^{(0)}\right\}^2 \geq 0$. Equality of the asymptotic variance holds if and only if the covariance between the baseline and follow-up responses is the same for both the treatment and control groups or the assigned probabil-

Table 1. Comparison of the Five Estimators With Large Correlation Between Baseline and Follow-Up Responses

Estimators for β	Average	Sample variance	Variance estimate	OLS variance	Sample size
ANCOVA I	0.557	0.712	0.575	0.655	20
ANCOVA II	0.552	0.724	0.567	0.720	20
GEE method	0.482	0.609	0.594		20
Two-sample *t* test	0.506	0.710	0.731		20
Paired *t* test	0.507	0.610	0.632		20
ANCOVA I	0.502	0.386	0.375	0.382	50
ANCOVA II	0.501	0.387	0.370	0.392	50
GEE method	0.487	0.375	0.371		50
Two-sample *t* test	0.499	0.445	0.447		50
Paired *t* test	0.499	0.380	0.385		50
ANCOVA I	0.509	0.275	0.267	0.268	100
ANCOVA II	0.509	0.275	0.263	0.271	100
GEE method	0.495	0.266	0.263		100
Two-sample *t* test	0.501	0.314	0.312		100
Paired *t* test	0.500	0.269	0.269		100

ity $\delta = 0.5$; otherwise, $\hat{\beta}_1$ is less efficient than $\hat{\beta}_e$ and $\hat{\beta}_2$. Thus, for the many studies where treatment is assigned with equal probability, the ANCOVA I estimator of treatment effect is fully efficient.

The asymptotic variance of $\hat{\beta}_p$, the estimator based on the paired t test, is generally larger than that of $\hat{\beta}_e$ and $\hat{\beta}_2$, where the difference between the asymptotic variances of $\hat{\beta}_p$ and $\hat{\beta}_e$ is $\Sigma_p - \Sigma_{e1} = \frac{1}{\delta(1-\delta)\sigma_{11}}\{\sigma_{11} - (1-\delta)\sigma_{12}^{(1)} - \delta\sigma_{12}^{(0)}\}^2 \geq 0$. This difference, however, is identically equal to zero if and only if the variance of the response at baseline equals the average of the treatment-specific covariances of baseline and follow-up responses; namely, $\sigma_{11} = (1-\delta)\sigma_{12}^{(1)} + \delta\sigma_{12}^{(0)}$. When this is the case, the estimator of treatment difference based on the paired t test is fully efficient. This would occur, for example, if the difference between follow-up and baseline responses $D_i = Y_{2i} - Y_{1i}$ and the baseline response Y_{1i} are conditionally uncorrelated given Z_i, in which case $\sigma_{12}^{(1)} = \sigma_{12}^{(0)} = \sigma_{11}$.

Finally, the asymptotic variance of $\hat{\beta}_t$ used in the two-sample t test is larger than that of $\hat{\beta}_e$ and $\hat{\beta}_2$, except when the baseline and follow-up responses are uncorrelated.

4. SIMULATION EXPERIMENTS

The results given earlier are all based on the asymptotic properties of the five estimators. To study the small sample properties of these estimators, we conducted a series of Monte Carlo simulation experiments with sample sizes (20, 50, and 100). Four different scenarios were considered, each designed to test different aspects of the theory that was developed throughout this article. In each of the four scenarios, estimates were obtained for the treatment difference β using the five methods described in this article. The results, which are presented in Tables 1–4, are all based on 5,000 simulations. The tables present the average of the estimates across the 5,000 simulations as a basis for evaluating the bias. The Monte Carlo sample variance was used to evaluate the relative efficiency of the different estimators. In addition, we computed the average of the estimators of the asymptotic variance which can be compared to the empirical sample variance to evaluate the accuracy of the asymptotic theory. For the ANCOVA I and ANCOVA II models, we estimated the variance of $\hat{\beta}_1$ and $\hat{\beta}_2$ by substituting empirical estimates for the variance and covariance terms appearing in formulas (7) and (10), respectively. In practice, if OLS is used to estimate these parameters, then the variance for the estimator of treat-

Table 2. Comparison of the Five Estimators With Low Correlation Between Baseline and Follow-Up Responses

Estimators for β	Average	Sample variance	Variance estimate	OLS variance	Sample size
ANCOVA I	0.586	0.568	0.462	0.525	20
ANCOVA II	0.586	0.575	0.464	0.550	20
GEE method	0.501	0.481	0.465		20
Two-sample *t* test	0.511	0.475	0.490		20
Paired *t* test	0.513	0.610	0.630		20
ANCOVA I	0.516	0.307	0.291	0.300	50
ANCOVA II	0.516	0.308	0.292	0.304	50
GEE method	0.500	0.298	0.292		50
Two-sample *t* test	0.503	0.299	0.299		50
Paired *t* test	0.503	0.380	0.384		50
ANCOVA I	0.515	0.212	0.205	0.210	100
ANCOVA II	0.515	0.212	0.206	0.211	100
GEE method	0.500	0.206	0.206		100
Two-sample *t* test	0.502	0.208	0.209		100
Paired *t* test	0.501	0.265	0.269		100

Table 3. Comparison of the Five Estimators When the Assigned Probability $\delta = 0.7$

Estimators for β	Average	Sample variance	Variance estimate	OLS variance	Sample size
ANCOVA I	0.504	0.425	0.408	0.439	50
ANCOVA II	0.503	0.428	0.385	0.452	50
GEE method	0.522	0.397	0.386		50
Two-sample *t* test	0.499	0.453	0.456		50
Paired *t* test	0.494	0.407	0.415		50
ANCOVA I	0.502	0.291	0.287	0.303	100
ANCOVA II	0.502	0.291	0.274	0.307	100
GEE method	0.512	0.278	0.274		100
Two-sample *t* test	0.500	0.316	0.317		100
Paired *t* test	0.498	0.287	0.288		100

ment effect is generally estimated from standard formulas using the residuals from the OLS fit and the design matrix. Although this may not yield consistent estimators for the asymptotic variance under the general semiparametric models considered here, these were included for comparison and referred to as "OLS variance" in the tables.

The mean parameters are set to be $\beta = 0.5$, $\mu_1 = 0.0$, and $\mu_2 = -0.25$. In the first two simulation scenarios we used a sample size of 20, 50, and 100. The baseline covariates were generated independently from a standard normal distribution. Treatment was assigned into the two groups with probability $\delta = 0.5$. To demonstrate the robustness of the different estimators to deviations from the standard assumption of linearity between baseline and follow-up responses, we generated follow-up data assuming a quadratic relationship. Namely,

$$Y_{2i} = (\mu_2 + \beta Z_i) + (\beta_1 + \beta_2 Z_i)(Y_{1i} - \mu_1) + (\beta_3 + \beta_4 Z_i)\{(Y_{1i} - \mu_1)^2 - \sigma_{11}\} + \epsilon_i, \quad (15)$$

where ϵ_i was generated from a standard normal distribution. The relationship between the baseline and follow-up outcomes are determined by the coefficients $\beta_1, \beta_2, \beta_3$, and β_4. In the first simulation experiment, they are set to be 0.5, 0.6, 0.4, and 0.3, respectively. These coefficients were chosen so that the condition for the paired t test to have good efficiency is satisfied. The correlations between baseline and follow-up outcomes are 0.40 in the control group and 0.62 in the treatment group.

In the second experiment, the coefficients $\beta_1, \beta_2, \beta_3$, and β_4 are set to be 0.1, 0.1, 0.1, and 0.1, respectively. These were chosen so that the correlations between baseline and follow-up outcomes are smaller: 0.10 in the control group and 0.19 in the treatment group.

The results of these two simulations are depicted in Tables 1 and 2. For sample sizes 50 and 100, all five of these methods gave estimators for β with almost no bias. The variance estimators were also very accurate. When the sample size is 20, the GEE, the ANCOVA I, and ANCOVA II estimators exhibited larger biases as compared to the other two methods. In the first scenario, the two-sample t test gave the least efficient result with the other four methods yielding similar empirical variances. In contrast, for the second scenario, the paired t test gave the least efficient results with the other four being similar. These results confirm the asymptotic theory. With sample sizes 50 and 100, the empirical results suggest that the GEE estimator was slightly more efficient than either of the two ANCOVA estimators. In general, we found that the OLS variance estimator for the ANCOVA models behaved well. For small samples, the OLS variance estimators were actually less biased than our estimated variances. As the sample size increased, however, the OLS variance estimators showed some slight bias, whereas, the bias of our estimated variances disappeared.

Table 4. Comparison of the Five Estimators Using t Distribution

Estimators for β	Average	Sample variance	Variance estimate	OLS variance	Sample size
ANCOVA I	0.511	0.983	0.743	0.897	20
ANCOVA II	0.489	1.129	0.758	1.143	20
GEE method	0.458	0.876	0.868		20
Two-sample *t* test	0.499	1.022	1.036		20
Paired *t* test	0.508	0.933	0.949		20
ANCOVA I	0.481	0.559	0.502	0.545	50
ANCOVA II	0.479	0.570	0.519	0.622	50
GEE method	0.470	0.556	0.549		50
Two-sample *t* test	0.500	0.652	0.646		50
Paired *t* test	0.501	0.587	0.593		50
ANCOVA I	0.506	0.402	0.378	0.393	100
ANCOVA II	0.506	0.401	0.382	0.412	100
GEE method	0.490	0.394	0.388		100
Two-sample *t* test	0.502	0.452	0.448		100
Paired *t* test	0.503	0.410	0.411		100

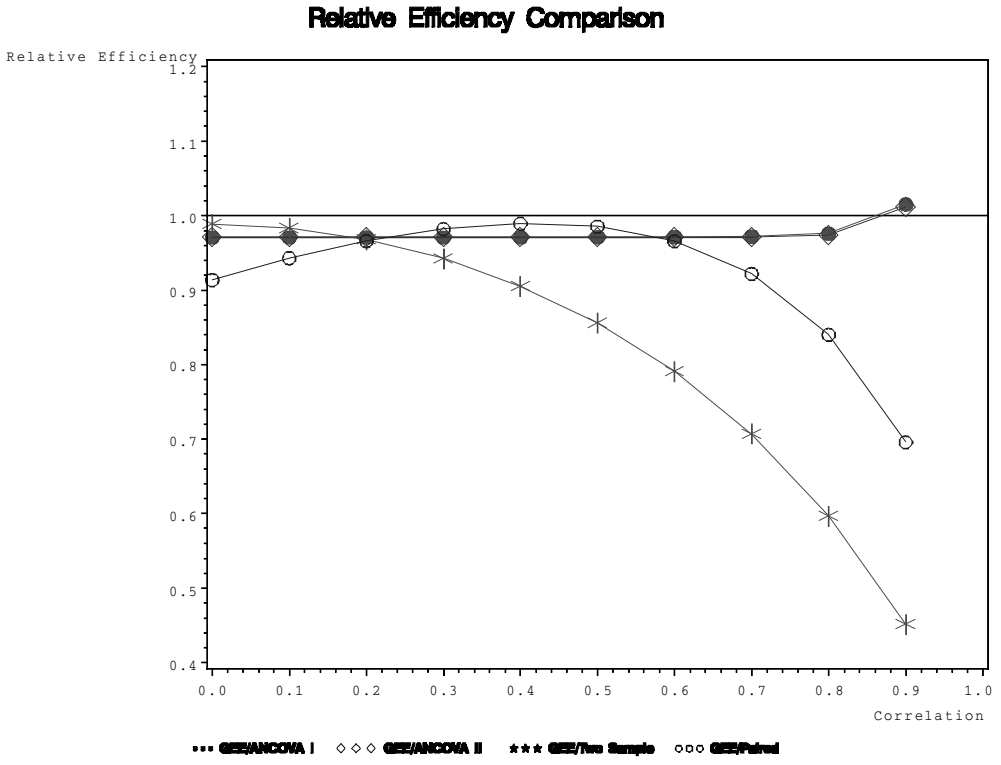

Figure 1. Histogram for $\psi = (-.7, -1.2)$.

The third simulation scenario was similar to the first except that we chose the randomization probability to be 0.7 rather than 0.5. For this scenario, the theory predicts that the ANCOVA I estimator may be less efficient than the GEE and ANCOVA II estimators. Because of the very small number of patients that would be randomized to the treatment arm with the smaller randomization probability, we did not include sample size 20. The results of the simulations for this scenario are given in Table 3. The empirical results were very similar to those in Table 1 suggesting that the GEE estimator is slightly more efficient with no difference in effciency between the two ANCOVA estimators.

In the fourth simulation scenario, we used sample sizes of 20, 50, and 100, and considered the robustness of these methods to deviations from normality. Thus, for this scenario, the baseline response Y_{1i} and the error distribution ϵ_i from Equation (15) are generated from a t distribution with seven degrees of freedom instead of a normal distribution. Everything else was identical to the first scenario. The results are depicted in Table 4. For this scenario, with sample sizes 20 and 50, the GEE estimator and the two ANCOVA estimators had some bias with the GEE estimator exhibiting the largest bias. With sample size 100, however, all the estimators behaved well with the GEE estimator being the most efficient.

To further investigate the effect of correlation of baseline and follow-up responses on the relative efficiency of our five estimators, we conducted several more simulations with a sample size of 100 and 5,000 simulations. We used a scenario similar to the first except that we chose the coefficients in (15) in such a way so that the correlation between baseline and follow-up responses were equal by treatment and varied from 0.0 to 0.9. Since all the estimators are nearly unbiased with this sample size, we evaluated efficiency based on the ratio of the empirical variances from the 5,000 simulations. The GEE estimator was used as a basis for comparison and the results are presented in Figure 1. As in our previous experiments, the GEE estimator had the best efficiency across the range of correlations with the two ANCOVA estimators giving similar but slightly less efficient results. The estimators based on the two-sample t test and the paired t test were generally less efficient than the other three, with the paired t test being more efficient than the two-sample t test when the correlation was greater than 0.2 and the reverse when the correlation was less than 0.2.

5. CONCLUSION AND DISCUSSION

This article considers five competing estimators that have been proposed for estimating treatment difference in studies which use a randomized pretest–posttest design under general distribution-free semiparametric assumptions. All five of these estimators were shown to be consistent and asymptotically normal. In terms of their large sample properties, we found the estimators $\hat{\beta}_e$ from the GEE method and $\hat{\beta}_2$ from the ANCOVA II model to have the same asymptotic variance and are asymptotically the most efficient. The estimator $\hat{\beta}_1$ from the ANCOVA I model was also fully efficient if the randomization probability to the treatment arm was $\delta = 0.5$ or when the covariance between the baseline and follow-up responses is the same for both the treatment and control groups. Although this estimator may be less efficient than $\hat{\beta}_e$ or $\hat{\beta}_2$, we found the loss of efficiency to be trivial in all the cases we considered. The estimator $\hat{\beta}_t$ from the two-sample t test is fully efficient only when the baseline and follow-up responses were uncorrelated, otherwise, there could be a substantial loss of efficiency compared to the other three estimators above. The estimator $\hat{\beta}_p$ from the paired

t test, although generally preferable to $\hat{\beta}_t$ when there is moderate to strong correlation between the baseline and follow-up responses, can also result in a substantial loss of efficiency compared to the first three. We did, however, show that when the difference between follow-up and baseline responses were conditionally uncorrelated to baseline response by treatment, then $\hat{\beta}_p$ was fully efficient.

Simulation experiments were conducted to compare these five estimators under a wide variety of circumstances with small to moderate sample sizes. We found with small sample sizes, on the order of 20 individuals, that the three efficient estimators $\hat{\beta}_e$, $\hat{\beta}_2$, and $\hat{\beta}_1$ tended to give estimates of treatment difference that were biased. Moreover, the variance estimators performed poorly. For such sample sizes it may be preferable to use the estimator based on the paired t test unless the evidence suggests a small correlation between baseline and follow-up responses, in which case, the estimator based on the two-sample t test is better. However, we found that with sample sizes of 50 or greater, the estimator $\hat{\beta}_e$ from the GEE equations performed the best. We believe this should be the estimator of choice for moderate to large sample sizes.

[Received May 2000. Revised May 2001.]

REFERENCES

Chamberlain, G. (1987), "Asymptotic Efficiency in Estimation With Conditional Moment Restrictions," *Journal of Econometrics*, 34, 305–324.

Crager, M. R. (1987), "Analysis of Covariance in Parallel-Group Clinical Trials with Pretreatment Baseline," *Biometrics*, 43, 895–901.

Feldt, L. S. (1958), "A Comparison of the Precision of Three Experimental Designs Employing A Concomitant Variable," *Psychometrika,* 23, 335–353.

Follmann, D. A. (1991), "The Effect of Screening on some Pretest–Posttest Test Variances," *Biometrics*, 47, 763–771.

Laird, N. (1983), "Further Comparative Analyses of Pretest–Posttest Research Designs," *The American Statistician*, 37, 329–340.

Newey, W. K. (1990), "Semiparametric Efficiency Bounds," *Journal of Applied Econometrics*, 5, 99–135.

Pigrache, R. M., and Graham, B. R. (1976), "A Modification of the Johnson–Neyman Technique Comparing two Regressions, Applied to Treatment Effects Dependent on Baseline Levels," *Biological Psychology*, 4, 213–236.

Singer, J. M., and Andrade, D. F. (1997), "Regression Models for the Analysis of Pretest/Posttest Data," *Biometrics*, 53, 729–735.

Stein, R. A. (1989), "Adjusting Treatment Effects for Baseline and Other Predictor Variables," *ASA Bioparmaceutical Section*, 274–280.

Zeger, S. L., and Liang, K. Y. (1986), "Longitudinal Data Analysis for Discrete and Continuous Outcomes," *Biometrics*, 42, 121–130.