STATISTICS IN MEDICINE
Statist. Med. 2007; **26**:5100–5115
Published online 3 May 2007 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sim.2909

Baseline patient characteristics and mortality associated with longitudinal intervention compliance

Julia Y. Lin[1,*,†], Thomas R. Ten Have[2], Hillary R. Bogner[3] and Michael R. Elliott[4]

[1]*Center for Multicultural Mental Health Research, Somerville, MA 02143, U.S.A.*
[2]*Department of Biostatistics and Epidemiology, School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, U.S.A.*
[3]*Department of Family Practice and Community Medicine, School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, U.S.A.*
[4]*Department of Biostatistics, School of Public Health, and Survey Methodology Program, Institute for Social Research, University of Michigan, Ann Arbor, MI 48109, U.S.A.*

SUMMARY

Lin *et al.* (http://www.biostatsresearch.com/upennbiostat/papers/, 2006) proposed a nested Markov compliance class model in the Imbens and Rubin compliance class model framework to account for time-varying subject noncompliance in longitudinal randomized intervention studies. We use superclasses, or latent compliance class principal strata, to describe longitudinal compliance patterns, and time-varying compliance classes are assumed to depend on the history of compliance. In this paper, we search for good subject-level baseline predictors of these superclasses and also examine the relationship between these superclasses and all-cause mortality. Since the superclasses are completely latent in all subjects, we utilize multiple imputation techniques to draw inferences. We apply this approach to a randomized intervention study for elderly primary care patients with depression. Copyright © 2007 John Wiley & Sons, Ltd.

KEY WORDS: longitudinal compliance class model; noncompliance; principal stratification; latent class model; multiple imputation; geriatric depression

1. INTRODUCTION

In longitudinal randomization studies where subjects are randomized at baseline and interventions are administered repeatedly over time, subject adherence to intervention may vary over time.

*Correspondence to: Julia Y. Lin, Center for Multicultural Mental Health Research, 120 Beacon Street, Somerville, MA 02143, U.S.A.
†Email: jlin@charesearch.org

Contract/grant sponsor: NIH; contract/grant numbers: T32MH065218-1A1, R01-MH-61892-01A2, P30-MH066270

Received 17 August 2006
Accepted 20 March 2007

Copyright © 2007 John Wiley & Sons, Ltd.

In [1] (Lin, Ten Have, and Elliott, LTE), we examined the relationship between longitudinal intervention compliance and depression outcomes in the Prevention of Suicide in Primary Care Elderly: Collaborative Trial (PROSPECT) [2]. In this paper, we are interested in the subject-level baseline predictors of longitudinal intervention compliance, and the relationship between longitudinal intervention compliance and mortality.

The PROSPECT study was a randomized intervention study targeting elderly patients in primary care clinics with depression. There were two intervention arms: usual care and the intervention. In the usual care arm, patients received standard care. In the intervention arm, patients were assigned to meet with health specialists who educated patients, their families, and physicians about depression, treatment, and monitored adherence to the intervention. Primary care clinics were randomized to the intervention arms rather than individual patients to prevent contamination of interventions between patients within the same clinic and for practicality. Subject adherence to meeting with health specialists were measured at five time points over a two-year period, and adherence could vary over time.

Angrist *et al.* [3] and Imbens and Rubin [4] proposed using latent compliance classes to account for subject noncompliance in the context of cross-sectional randomized treatment studies. In studies where there are treatment noncompliance, the relationship between treatment randomization and outcomes may be confounded by post-randomization treatment adherence. Compliance classes are considered baseline factors, and utilizing them allows us to examine the relationship between post-randomization treatment adherence and outcomes.

In a two-arm randomized intervention study with an usual care arm and an intervention arm there are four possible compliance classes: complier, always-taker, never-taker, and defier. Compliers are those that would adhere to the intervention to which they are randomly assigned; always-takers are those that would choose to receive the intervention regardless of their randomization assignment; never-takers are those that would choose to receive the usual care regardless of their randomization assignment; and defiers are those that would follow the regime opposite the arm to which they are assigned. These compliance classes are latent because those randomized to and who follow the usual care regime could be compliers or never-takers; those randomized to and who follow the intervention regime could be compliers or always-takers; those randomized to the usual care but who follow the intervention regime could be always-takers or defiers; those randomized to the intervention but who refused the intervention, and instead follow the usual care regime, could be never-takers or defiers.

In studies such as PROSPECT where there are two intervention arms and subjects randomized to the usual care arm have no access to the intervention, there are only two possible compliance classes: compliers and never-takers. Since those randomized to the usual care group have no access to the intervention, there are no always-takers or defiers by definition. In this study design, the compliance classes are observed in those randomized to the intervention arm but are latent in those randomized to the usual care arm. Subjects randomized to the usual care arm can only be observed to follow the usual care regime; therefore, they could be compliers or never-takers. Subjects randomized to the intervention arm and observed to follow the intervention regime are compliers, and those observed to follow the usual care regime are never-takers.

In LTE [1] we extended the Imbens and Rubin [4] compliance class model to longitudinal studies where subjects are randomized at baseline and randomization status stays constant over time but adherence to the intervention could vary over time. With two possible compliance classes and five time points, we have a total of 32 (2^5) possible longitudinal compliance patterns. The proposed principal stratification strategy uses principal strata [5, 6] (PS), or 'compliance superclasses', to

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

summarize the longitudinal compliance patterns in the data, within which intent-to-treat contrasts are made to estimate the effect of the intervention on the depression outcome to account for intervention noncompliance.

In this paper, we are focusing on the compliance superclass and how they relate to baseline patient characteristics and all-cause mortality. We may make inferences on each of the longitudinal compliance patterns. However, such inferences may be sensitive to model assumptions when the number of subjects exhibiting each of the patterns are small, and inferences on such specific groups may be impractical and not clinically meaningful. It may be more useful to summarize the longitudinal compliance patterns with broader PS, or superclasses, and look at the relationship between the superclasses and baseline patient characteristics and mortality. Use of latent classes to summarize longitudinal patterns of adherence has been done on observed adherence to prescribed medication in the PROSPECT [7].

It is of clinical importance to identify baseline patient characteristics that relate to intervention adherence. If clinicians can identify patients who are likely to comply with particular interventions, it allows them to target those that are most likely to benefit from the interventions. If clinicians can identify patients who are less likely to comply with interventions, they can modify the interventions or offer extra incentive to increase patient compliance.

Previous research suggests that depression is associated with increased risk of death [8, 9], and there may be an association between treatment adherence and mortality [10–13]. We are interested in examining the relationship between longitudinal intervention compliance and all-cause mortality. We obtained subjects' vital status from the National Death Index, and continued to follow subjects after the PROSPECT study concluded.

For those subjects randomized to the intervention arm in the PROSPECT, their time-varying compliance classes are observed. Alternative to classifying subjects using latent compliance superclasses is to classify subjects into pre-specified observed PS based on observed time-varying compliance classes for those randomized to the intervention arm. Such PS could be subjects who complied with their randomization throughout the entire study period (perfect compliers) and ones who did not. In this paper, we compare baseline subject characteristics and mortality between latent compliance superclasses, then we compare the results to those based on observed PS among subjects randomized to the intervention arm.

We define notations in Section 2, and discuss analysis methods and details of the models in Section 3. We present the analysis results in Section 4, and finally make concluding remarks in Section 5.

2. NOTATION

Let Z_i denote the randomization status for subject i where $i = (1, \ldots, N)$, and $Z_i \in (0, 1)$ for randomized to the usual care and the intervention arm, respectively. Let D_{ij} denote the actual intervention received for subject i at time j where $j = (1, 2, 3, 4, 5)$ for 4-, 8-, 12-, 18-, and 24-month follow-up, respectively, and $D_{ij} \in (0, 1)$ for usual care and intervention, respectively. Note that subjects are only randomized at baseline; therefore, Z_i does not have a subscript j. Let Y_{ij} denote the outcome for subject i at time j. The partially latent compliance classes for subject i at time j is denoted by C_{ij} where $C_{ij} \in (c, n)$ for complier and never-taker, respectively. Let U_i denote membership of the latent compliance superclass for subject i, where $U_i = (1, \ldots, K)$ for assumed K numbers of superclasses. Let $\mathbf{Q}_i$ denote vector of baseline covariates for subject i

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

associated with intervention compliance, and $\mathbf{A}_i$ denote vector of baseline covariates associated with depression outcome.

3. ANALYSIS METHOD

In this analysis we implement a three-stage method. In the first stage, we fit the Markov nested compliance class model proposed in LTE [1] to describe the subject longitudinal intervention compliance behavior in the PROSPECT using latent compliance superclasses. In the second stage we treat the latent compliance superclasses as missing data, and utilize multiple imputation techniques with information from the Markov chain Monte Carlo (MCMC) chains in the first stage. In the third stage we relate the multiply imputed compliance superclass to baseline subject characteristics and all-cause mortality. We proceed to describe the latent compliance class and superclass model in stage 1, the incomplete data model in stage 2, and the complete data model in stage 3.

3.1. Stage 1: *Compliance principal stratification*

In the first stage we use the non-stationary Markov compliance class (MCC) model proposed in LTE [1] to describe time-varying compliance behavior. Latent compliance superclasses are used to summarize latent longitudinal compliance class patterns. We assume subject compliance superclass $(U_i = k) \sim \text{Multinomial}(1, p_k)$ where $\sum_k p_k = 1$. The MCC model assumes that subjects' compliance behavior is related to their history of compliance. More specifically, subject compliance is modeled as dependent on compliance in the previous follow-up period, latent compliance superclass, and baseline covariates

$$P(C_{i1}, \ldots, C_{i5}|U_i, \mathbf{Q}_i) = P(C_{i1}|U_i, \mathbf{Q}_i)P(C_{i2}|C_{i1}, U_i) \ldots P(C_{i5}|C_{i4}, U_i) \tag{1}$$

The time-varying compliance probabilities are estimated from the compliance probabilities at the first follow-up and the time-varying compliance class transitional probabilities. To account for the relationship between baseline covariates and subject intervention compliance, we include baseline covariates $\mathbf{Q}_i$ in modeling the compliance probability at the first follow-up using a multinomial logit model

$$P(C_{i1} = \eta|U_i = k, \mathbf{Q}_i) = \omega_{k\eta}(\mathbf{Q}_i)$$

$$\omega_{k\eta}(\mathbf{Q}_i) = \frac{\exp(\alpha_{0k\eta} + \boldsymbol{\alpha}_{1\eta}\mathbf{Q}_i)}{\left[\sum_{\eta'} \exp(\alpha_{0k\eta'} + \boldsymbol{\alpha}_{1\eta'}\mathbf{Q}_i)\right]}, \quad \sum_{\eta} \omega_{k\eta}(\mathbf{Q}_i) = 1 \ \ \forall k \tag{2}$$

The $\alpha_{0k\eta}$ and $\boldsymbol{\alpha}_{1\eta}$ for one of the compliance class η are constrained to 0 for identifiability. We assume the compliance class transition $(C_{ij} = \eta|C_{i,j-1} = \eta', U_i = k) \sim \text{Multinomial}(1, \pi_{kj\eta'\eta})$ where $\sum_{\eta} \pi_{kj\eta'\eta} = 1 \ \forall k, j, \eta'$.

The main outcome of the study is the Hamilton Depression Rating Scale (HAMD). It measures depression severity, with higher scores indicating more severe depression. It was measured at each of the follow-up visits for each patient. We assume outcomes within an individual are independent given randomization, time-varying compliance classes, baseline covariates, and subject-level

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

random effect

$$\begin{aligned}(Y_{ij}|C_{i1},\ldots,C_{ij},Z_i=z,\mathbf{A}_i,\mathbf{W}_i,\boldsymbol{\lambda},\boldsymbol{\gamma},\boldsymbol{\varphi}_i,\sigma^2) &\overset{\text{ind}}{\sim} \mathrm{N}(\mu_{ijz},\sigma^2)\\ \mu_{ijz}=\sum_{\eta'}[I(C_{ij}=\eta',Z_i=z)\lambda_{j\eta'z}]+\mathbf{A}_i^{\mathrm{T}}\boldsymbol{\gamma}+\mathbf{W}_i^{\mathrm{T}}\boldsymbol{\varphi}_i &\end{aligned} \tag{3}$$

The conditional mean of the outcome has three components: compliance class-specific effect of randomization, the effect of baseline covariates, and the subject-specific random effect to account for within-subject correlation in the repeatedly measured outcomes. The compliance class-specific effect of randomization on the outcome is represented by $\sum_{\eta'}[I(C_{ij}=\eta',Z_i=z)\lambda_{j\eta'z}]$, and $\boldsymbol{\lambda}$ denotes the vector of $\lambda_{j\eta'z}$ parameters. The effect of the baseline covariates on the outcome is represented by $\mathbf{A}_i^{\mathrm{T}}\boldsymbol{\gamma}$ where $\mathbf{A}_i$ denotes the vector of baseline covariates of subject i, and the column vector $\boldsymbol{\gamma}$ denotes the corresponding coefficients. The random effect $\boldsymbol{\varphi}_i$ is used to account for within-subject correlation in the outcomes, where $\mathbf{W}_i$ denotes the random-effect design matrix for subject i. We consider a random subject-level intercept model.

In fitting the compliance class model we assume two latent compliance superclasses ($U_i=(1,2)$). Let $\mathbf{Q}_i$ be the baseline HAMD score, and $\mathbf{A}_i$ be the baseline HAMD and baseline suicidal ideation. Model parameters are estimated using Bayesian MCMC method. Gibbs sampling [4, 14–16] is used to obtain draws from the posterior distributions of the parameters. The priors used in the estimation are described in the Appendix. The posterior distributions of the $\boldsymbol{\alpha}=(\alpha_{0k\eta},\boldsymbol{\alpha}_{1\eta})$ parameters are not of a known parametric form. Therefore, we use the Metropolis–Hasting algorithm [16, 17] to draw the $\boldsymbol{\alpha}$ parameters.

We plotted and assessed the posterior predictive mean residuals from the HAMD model (3) and found no strong relationship with the posterior predictive means of μ_{ijz}, baseline HAMD, and baseline suicide ideation, and no significant departure from normality among the residuals. However, there may be some weak evidence suggesting heteroscedasticity of the residual variance.

In the PROSPECT data, intervention received (D_{ij}) is missing for some patients at some time points. We assume these missing values are 0, or have not met with the health specialists. To our knowledge, all patients who met with the health specialists have outcomes recorded; therefore, it is reasonable to assume that if there were no outcomes recorded, then the patient did not meet with the health specialists. The outcome (Y_{ij}) is also missing for some patients at some time points. We assume a latent ignorable missing data mechanism [18] and assume that data are missing at random given latent compliance class and covariates. At each iteration of the MCMC procedure, we impute the missing outcomes conditional on compliance classes, randomization, baseline covariates, and subject-level random effect. We draw missing outcome Y_{ij}^{mis} for subject i at time j from its posterior predictive distribution given current values of parameters C_{ij}, $\lambda_{j\eta'z}$, $\boldsymbol{\gamma}$, $\mathbf{W}_i$, $\boldsymbol{\varphi}_i$, σ^2, and vector of observed outcomes $\mathbf{Y}^{\text{obs}}$

$$\begin{aligned}(Y_{ij}^{\text{mis}}|\mathbf{Y}^{\text{obs}},C_{ij},Z_i=z,\mathbf{A}_i,\mathbf{W}_i,\lambda_{j\eta'z},\boldsymbol{\gamma},\boldsymbol{\varphi}_i,\sigma^2) &\sim \mathrm{N}(\mu_{ijz}^*,\sigma^2)\\ \mu_{ijz}^*=\sum_{\eta'}[I(C_{ij}=\eta',Z_i=z)\lambda_{j\eta'z}]+\mathbf{A}_i^{\mathrm{T}}\boldsymbol{\gamma}+\mathbf{W}_i^{\mathrm{T}}\boldsymbol{\varphi}_i &\end{aligned} \tag{4}$$

Estimation in stage 1 was implemented using the statistical software R.

3.2. Stage 2: Incomplete data model

Longitudinal compliance superclass are latent for all subjects, so we essentially have a missing data problem where all data are missing. From stage 1 we have an estimate for each model

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

parameter at each iteration of the MCMC chain, from which we can compute posterior predictive distributions of the compliance superclasses. We use multiple imputation technique to impute the compliance superclasses based on their posterior predictive distributions. To improve computation speed, we only save the compliance superclasses for every 25th iteration of the MCMC chain. Some of the baseline subject characteristics were missing in the data. We use PROC MI in SAS to multiply impute the missing baseline covariates at every iteration of the MCMC chain using all other observed baseline covariates considered in the analysis.

3.3. Stage 3: Complete data models

Next, we use the multiple imputed compliance superclasses and baseline subject characteristics from stage 2 to examine the relationship between longitudinal compliance, baseline subject characteristics, and all-cause mortality. Specifically, we treat the compliance superclasses as the dependent variable in examining the relationship between subject-level baseline characteristics and longitudinal compliance, and treat the compliance superclasses as the independent variable in examining the relationship between longitudinal compliance and mortality.

To assess the within-clinic correlation in a preliminary analysis, we used PROC NLMIXED in SAS to run random-effects logistic regression models using observed perfect compliance in those randomized to the treatment group as the outcome variable, and each of their baseline covariates as the independent variable with a random intercept for each clinic. Then, we use the estimated variance of the random effect of clinics to calculate the within-clinic correlation, $\sigma^2/(\sigma^2+\pi/3)$, where σ^2 is the between-clinic variance from the random-effects models. The within-clinic correlation for the compliance superclass was notable (ranges from 0.25 to 0.26); therefore, we use random-effects logistic regression models to account for within-clinic correlations. Similarly, using PROC NLMIXED we ran frailty models with time-to-death as the outcome variable and observed perfect compliance as the independent variable. We calculate the between-subject variance the same way we calculated the between-clinic variance, then calculate the within-clinic correlation. The within-clinic correlation for mortality was close to 0; hence, we ignore it for the mortality analysis.

For the analysis relating each of the baseline covariates to the compliance superclass, we use random-effects logistic regression model with clinic-level random intercept to account for the within-clinic correlation of the compliance superclass

$$P(U_i=k|X_i, G_l)=\frac{\exp(\delta_{0k}+X_i\delta_{1k}+I(i\in l)G_l)}{\sum_{k=1}^{K}\exp(\delta_{0k}+X_i\delta_{1k}+I(i\in l)G_l)} \tag{5}$$

where X_i is the subject-level baseline covariate of interest, $l\in(1,\ldots,L)$ denotes the lth clinic, and G_l is the clinic-level random effect for the lth clinic. We assume δ_{0k} and δ_{1k} for one of the compliance superclass is 0 for identifiability, and assume that the random effect $G_l\sim \mathrm{N}(0,\sigma_G^2)$. The marginal likelihood integrating over the random effects is given by

$$\int\prod_{i=1}^{N}\left[\frac{\exp(\delta_{0k}+X_i\delta_{1k}+I(i\in l)G_l)}{\sum_{k=1}^{K}\exp(\delta_{0k}+X_i\delta_{1k}+I(i\in l)G_l)}\right]^{I(U_i=k)}\times\frac{1}{\sqrt{2\pi}\sigma_G}\exp\left[-\frac{(I(i\in l)G_l)^2}{2\sigma_G^2}\right]\mathrm{d}G_l \tag{6}$$

The parameters are estimated by maximizing the marginal likelihood and using Gaussian quadrature integration methods implemented with the SAS NLMIXED procedure. We use a cluster-specific approach here rather than a population average approach (e.g. generalized estimating equations

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

approach) because we are interested in describing the change in a subject's compliance superclass membership given change in the baseline covariates.

We use the survival function to assess the probability distribution for time to death. The survival function conditional on compliance superclass is defined as $S(t|U_i=k)=P(T_i\geqslant t|U_i=k)$, which is the probability that a subject survives beyond time t given that the subject is in the kth compliance superclass. The Kaplan–Meier estimate of the survival function is

$$\hat{S}(t|U_i=k)=\prod_{t^*\leqslant t}\left[1-\frac{d_{kt^*}}{N_{kt^*}}\right]$$

where d_{kt^*} denotes the number of deaths occurring at time t^* in compliance superclass k, and N_{kt^*} denotes the number of subjects remaining in the study in compliance superclass k who have not died prior to time t^*. The ratio d_{kt^*}/N_{kt^*} estimates the probability that death occurs between t^* and the following observed event time given an individual survives up to time t^* in compliance superclass k.

We also use the hazard function to describe the rate of death. When there are no covariates, the hazard rate can be defined as

$$h(t)=\lim_{\Delta t\to 0}\frac{P(t\leqslant t^*\leqslant t+\Delta t|t^*\geqslant t)}{\Delta t}$$

which is the instantaneous rate of death at time t. The cumulative hazard, $H(t)=\int_0^t h(u)\,\mathrm{d}u$, can be estimated nonparametrically by $\sum_{t^*\leqslant t} d_{kt^*}/N_{kt^*}$. Various models can be used to relate covariates to hazard rates. We use the Cox proportional hazards model to examine the relationship between compliance superclass and all-cause mortality. We assume the proportional hazards

$$h(t|U_i=k)=h_0(t)\exp(\beta_k I(U_i=k)) \tag{7}$$

where β_k for one of the compliance superclass (reference superclass) is assumed 0 for identifiability. Let $t_1,\ldots,t_N$ denote the ordered time of death, $U_{(i)}$ denotes the compliance superclass of the subject who died at time t_i, and $R_{(i)}$ denotes the set of individuals who are still alive just prior to time t_i. The partial likelihood is given by

$$\prod_{i=1}^{N}\frac{\sum_{k=1}^{K}\exp(\beta_k)I(U_{(i)}=k)}{\sum_{i^*\in R_{(i)}}\sum_{k=1}^{K}\exp(\beta_k)I(U_{i^*}=k)} \tag{8}$$

The baseline hazard $h_0(t)$ is the hazard rate for subjects in the reference compliance superclass and is estimated nonparametrically. The parameter β_k is estimated using maximum likelihood method and interpreted as the log-relative risk of death for the kth superclass *versus* the reference superclass.

Now we address how to pool parameter estimates across imputations. Let θ_m be the parameter of interest, which is δ_{1k} from the logistic regression model, Kaplan–Meier survival estimate $S(t|U_i=k)$, and β_k from the Cox model of the mth imputed data set, where $m=1,\ldots,M$. In this analysis $M=300$. Let $\hat{\theta}_m$ denote the estimate of θ_m. We can summarize $\hat{\theta}_m$ across multiple imputations by calculating the mean and associated variance. The combined estimates of the parameter of interest is $\hat{\theta}=(1/M)\sum_{m=1}^{M}\hat{\theta}_m$. There are two types of variability associated

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

with the above estimate: within-imputation variance, and between-imputation variance. The average within-imputation variance is $W=(1/M)\sum_{m=1}^{M}\mathrm{var}(\hat{\theta}_m)$. The between-imputation variance is $B=(1/(M-1))\sum_{m=1}^{M}(\hat{\theta}-\hat{\theta}_m)^2$. The total variance associated with $\hat{\theta}$ is $W+((M+1)/M)B$. The parameter $\hat{\theta}$ has a t-distribution with degrees of freedom

$$\left[(M-1)\left(1+\frac{M}{M+1}\frac{W}{B}\right)^2\right]$$

[19, p. 211]. However, when M is large (here, 300), the t-distribution is closely approximated by the normal distribution.

4. RESULTS

4.1. Compliance principal stratification

In LTE [1] we assumed three latent compliance superclasses and found a very small class whose behavior was hard to interpret clinically. We did not think analysis based on such small group would lead to any meaningful conclusions with regard to their survival probabilities or predictors of compliance. Thus, in this analysis we assume there are two latent compliance superclasses (i.e. $K=2$).

Table I shows the time-varying compliance probabilities at the sample average baseline HAMD score of 18.1. The first latent compliance superclass consists of subjects who are slightly more likely to comply with assigned intervention at the 4-month follow-up than not comply, and compliance decreases over time (increasing noncompliers). The second latent compliance superclass consists of subjects who are highly compliant during the first 18 months of the study with a small drop in compliance during the last 6 months (high compliers). The mean posterior probabilities of memberships in the increasing noncomplier and high complier superclasses and their associated 95 per cent credible intervals are 0.51 (0.36, 0.65), and 0.49 (0.35, 0.64), respectively.

The posterior means of the log odds of compliance at 4 months and their associated 95 per cent credible intervals within each of the latent compliance superclasses controlling for baseline HAMD are -0.48 (-1.63, 0.69) and 4.61 (1.48, 9.89) for the increasing noncomplier and the

Table I. Posterior means and 95 per cent credible intervals (in parentheses) for the time and latent compliance superclass-varying compliance probabilities assuming the average baseline HAMD of 18.1 and latent compliance superclass probabilities.

Time (months)	Increasing noncompliers	High compliers
4	0.65 (0.55, 0.77)	0.99 (0.92, 1.00)
8	0.35 (0.21, 0.53)	0.99 (0.90, 1.00)
12	0.16 (0.00, 0.36)	0.99 (0.90, 1.00)
18	0.07 (0.02, 0.26)	0.98 (0.76, 1.00)
24	0.03 (0.01, 0.08)	0.80 (0.61, 1.00)
$P(U_i=k)$	0.51 (0.36, 0.65)	0.49 (0.35, 0.64)

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

Table II. Posterior means and 95 per cent credible intervals (in parentheses) of the transitional probabilities.

U_i	j	$P(C_{i,j}=c\|C_{i,j-1}=c, U_i=k)$	$P(C_{i,j}=c\|C_{i,j-1}=n, U_i=k)$
Increasing noncomplier	2	0.53 (0.33, 0.71)	0.01 (0.00, 0.07)
	3	0.42 (0.00, 0.70)	0.00 (0.00, 0.02)
	4	0.12 (0.00, 1.00)	0.05 (0.02, 0.09)
	5	0.23 (0.03, 0.56)	0.01 (0.00, 0.04)
High complier	2	1.00 (1.00, 1.00)	0.06 (0.00, 1.00)
	3	1.00 (0.99, 1.00)	0.81 (0.00, 1.00)
	4	0.98 (0.83, 1.00)	0.44 (0.00, 1.00)
	5	0.82 (0.73, 1.00)	0.48 (0.00, 1.00)

high complier superclasses, respectively. This suggests that at the 4-month follow-up, those in the increasing noncomplier superclass are less likely to comply with their intervention assignment, and those in the high complier superclass are more likely to comply with their intervention assignment controlling for baseline HAMD. In our compliance class model we assume that the association between baseline HAMD and compliance probability at 4 months is the same across all latent compliance superclasses. The log odds of 4-month compliance for a unit increase in the baseline HAMD is 0.06 (0.01, 0.12) suggesting that those with more severe depression at baseline are more likely to comply with their intervention assignment.

The latent transitional probabilities of the time-varying compliance within each latent compliance superclass in Table II shows that both the increasing noncompliers and the high compliers are more likely to be in the complier compliance class if they are in the complier class in the previous follow-up period than if they are in the never-taker class and switch to complier class. Subjects in the high complier superclass are much more likely to transition into the complier class than subjects in the increasing noncomplier superclass. Few patients in the high complier superclass switched from the never-taker class to the complier class, causing large variability in the transitional probabilities; hence, we see the wide range in the credible intervals.

In the PROSPECT study, 320 of the 598 study subjects were randomized to the intervention arm. Forty-six per cent ($N = 147$) of those 320 subjects are observed to be perfect compliers who met with the health specialist throughout the entire study period.

4.2. Baseline patient characteristics

In this section, we explore the relationship between longitudinal compliance and selected baseline subject characteristics that may be related to longitudinal intervention compliance, examining one characteristic at a time.

We look at age, years of education, numbers of comorbidities measured by the Charlson Comorbidity Index, cognitive functioning measured by the Mini Mental State Examination, race in terms of white or ethnic minority as defined by the respondent, sex, marital status, diagnosis of diabetes, diagnosis of heart disease, smoking, diagnosis of major depression, and physical functioning measured by the physical component of the SF-36 health survey at baseline. We assume that those with missing information on diagnosis of diabetes, heart diseases, and smoking at baseline do not have those diagnoses or smoke.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

Table III. Summary of the baseline characteristics.

Covariate	Per cent or Mean (SD)	Min	Max	Missing *N*
Age (years)	70.2 (7.9)	60	94	0
Education (years)	12.7 (3.3)	0	25	0
Comorbidity (#)	3.0 (2.4)	0	12	0
Cognitive function	27.4 (2.7)	0	30	0
Minority race	30.0	0	1	0
Female	72.2	0	1	0
Married	37.1	0	1	0
Diabetes	21.0	0	1	0
Heart disease	20.0	0	1	0
Smoking	8.4	0	1	0
Major depression	66.0	0	1	0
SF-36 physical health	39.6 (13.1)	8.9	70	135

Table IV. Odds ratio of high complier superclass membership relative to increasing noncomplier superclass membership with a unit change in the baseline characteristics.

Predictor	OR (95 per cent CI)
Age (year)	0.99 (0.96, 1.02)
Education (year)	1.00 (0.94, 1.06)
Comorbidity (#)	0.94 (0.86, 1.03)
Cognitive function (1 patient)	1.05 (0.97, 1.14)
Minority race	0.81 (0.48, 1.37)
Female	0.98 (0.62, 1.55)
Married	0.98 (0.63, 1.52)
Diabetes	0.94 (0.58, 1.52)
Heart disease	0.90 (0.53, 1.53)
Smoking	1.08 (0.53, 2.18)
Major depression	0.84 (0.55, 1.28)
SF-36 physical health (10 patients)	1.18 (0.99, 1.41)

Eighteen subjects (3.0 per cent) with missing years of education, number of comorbidities, race, and marital status were excluded from the analysis resulting in a sample of 580 subjects. The baseline characteristics of the 580 analysis sample are described in Table III. Those included in the analysis did not differ from those excluded from the analysis on age, cognitive functioning, diagnosis of diabetes, diagnosis of cardiovascular diseases, smoking, and diagnosis of major depression at baseline. However, those included in the analysis are more likely to be females (72.2 per cent) than those excluded from the analysis (50.0 per cent; $\chi^2_1 = 4.24$, $p = 0.04$). Of the remaining 580 subjects, 135 had missing physical functioning scores but all other baseline covariates were observed. Only missing physical functioning scores were multiply imputed. Age, education, number of comorbidities, cognitive function, race, sex, marital status, diagnosis of diabetes, diagnosis of heart disease, smoking, and diagnosis of major depression were used in the imputation model for missing physical functioning scores.

Table IV shows the point estimates and the associated 95 per cent confidence intervals for the odds ratio of high complier superclass membership *versus* increasing noncomplier superclass

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

Table V. Odds ratio of observed perfect complier PS membership relative to observed nonperfect complier PS membership with a unit change in the baseline characteristics (intervention arm only).

Predictor	OR (95 per cent CI)
Age (year)	0.97 (0.94, 1.00)
Education (year)	1.03 (0.96, 1.11)
Comorbidity (#)	0.88 (0.80, 0.97)
Cognitive function (1 patient)	1.07 (0.98, 1.17)
Minority race	0.73 (0.38, 1.40)
Female	1.03 (0.63, 1.71)
Married	1.03 (0.61, 1.74)
Diabetes	0.91 (0.52, 1.59)
Heart disease	0.51 (0.28, 0.94)
Smoking	0.96 (0.45, 2.05)
Major depression	0.91 (0.56, 1.49)
SF-36 physical health (10 patients)	1.37 (1.09, 1.71)

membership given a unit increase in each of the baseline covariates. A unit is defined as one point in the original scales of the variables, except for the physical functioning scores where a more clinically relevant unit is 10 points on the original scale. We found a marginally significant association between physical functioning and latent compliance superclass. A higher score on the physical component of the SF-36 indicates better physical health. The odds of high complier superclass membership with a 10-point increase in the SF-36 physical score is 1.18 with 95 per cent confidence interval (0.99, 1.41). This suggests that healthier subjects may more likely be highly compliant. We adjusted for multiple comparisons using Holm's test, which has been shown to protect the family-wise error rate [20]. Physical functioning remained marginally significant after multiple comparisons adjustment.

Table V displays the odds ratio of being observed to be a perfect complier *versus* not a perfect complier given a unit change in each of the baseline subject characteristics. We find that better physical health is associated with observed compliance. Subjects with higher number of comorbidities are less likely to be perfect compliers than subjects with lower number of comorbidities (odds ratio 0.88 (0.80, 0.97)). Subjects with diagnoses of heart disease at baseline are less likely to be perfect compliers than subjects without diagnoses (odds ratio 0.51 (0.28, 0.94)). Subjects with higher SF-36 scores are more likely to be perfect compliers than subjects with lower scores (odds ratio 1.37 (1.09, 1.71)). The associations between numbers of comorbidities and physical functioning remained significant after a multiple comparisons adjustment using Holm's test.

The comparisons of baseline subject characteristics between latent compliance superclasses and observed compliance PS yielded evidence that those with better physical functioning may more likely be highly compliant than those with worse physical functioning. Previous research [21–24] has shown that patient adherence to treatment is associated with ease of complying. In this analysis, adherence to the intervention is defined by meeting with the health specialists at the clinics during each of the follow-up periods. It may be easier for those in good physical health to go to the clinics to meet with the health specialists than those in bad physical health. The comparisons based on the observed compliance PS also showed that higher number of comorbidities and diagnoses of heart disease at baseline are associated with lower compliance. Both these characteristics pertain

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

to physical health. It is consistent with our reasoning that it may be easier for healthier subjects to attend meetings with health specialists than subjects who are unhealthy.

4.3. Mortality

In this section we look at the relationship between all-cause mortality and longitudinal compliance. Mortality data were obtained from the National Death Index. The mean follow-up time for mortality status is 28.9 months with a standard deviation of 8.0 months. The minimum follow-up time is 0.8 months and maximum is 43.7 months.

Figure 1 displays the Kaplan–Meier survival curves stratified on the multiply imputed latent compliance superclass, and Figure 2 displays the survival curves stratified on the observed compliance PS. The bold lines in Figure 1 represent the mean Kaplan–Meier survival curves, calculated from taking the averages of the survival estimates at each time point over the MCMC iterations. The thinner bounding lines in both figures represent the 95 per cent confidence intervals of the survival curves. We see that those in high complier superclass and perfect complier PS have higher probabilities of survival than those in increasing noncomplier superclass or non-perfect complier PS.

We fitted the Cox proportional hazards model using the latent compliance superclass as the independent variable. We found that those in the high complier superclass are less likely to die than those in the increasing noncomplier superclass (hazard ratio: 0.32 (0.15, 0.68)). We applied

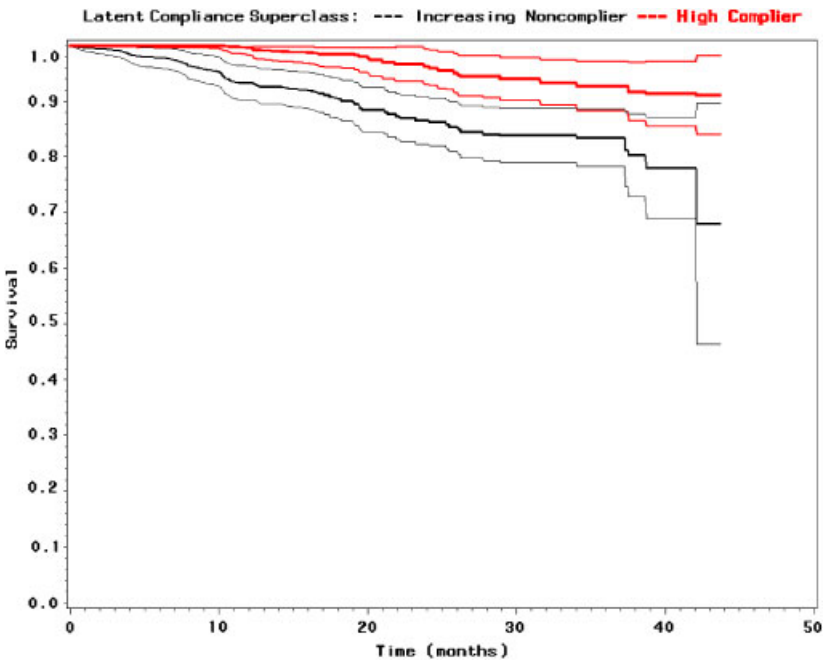

Figure 1. Multiple imputation Kaplan–Meier survival curves by latent compliance superclass with 95 per cent confidence intervals.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

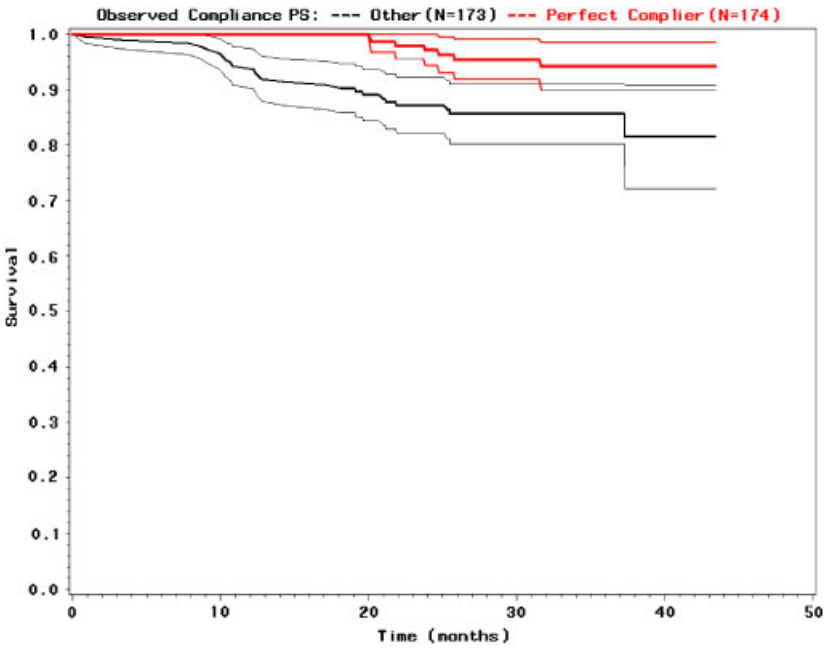

Figure 2. Multiple imputation Kaplan–Meier survival curves by observed compliance PS with 95 per cent confidence intervals (intervention arm only).

the same Cox model to observed compliance PS and found that perfect compliers are also less likely to die than non-perfect compliers (hazard ratio: 0.29 (0.13, 0.67)).

High compliers randomized to the intervention arm are more likely to receive the intervention than increasing noncompliers randomized to the intervention arm. A model with main effects of randomization and latent compliance superclass, and the interaction between randomization and superclass shows that the interaction is not a significant predictor of mortality. This suggests that mortality is associated with underlying patient compliance tendencies rather than the actual intervention received.

5. DISCUSSION

In this paper, we presented a MCC model in the principal stratification framework to characterize subject longitudinal compliance patterns. The method was discussed in the context of the PROSPECT where subjects were randomized once at baseline and randomization status remained the same throughout the study. Interventions were applied repeatedly over time and subject adherence to their randomized intervention arm was measured at five follow-up time points and may vary over time. In the PROSPECT, those randomized to the usual care arm have no access to the intervention; therefore, yields only two possible compliance classes: compliers and never-takers. Two compliance classes and five measurement times yield 32 (2^5) possible longitudinal compliance patterns. We use broader compliance PS, or 'compliance superclasses', to describe the longitudinal

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

compliance patterns present in the data. The time-varying compliance classes are modeled to be dependent on the history of compliance behavior, or more specifically, the compliance class in the previous follow-up.

The analyses utilized the principal stratification framework to relate post-randomization factors to outcomes. Specifically, we used latent compliance superclasses to describe longitudinal profiles of intervention adherence. An advantage of using the proposed latent compliance superclass compared to using observed adherence to the intervention is that superclasses are considered pre-randomization factors, thus allowing us to examine the relationship between post-randomization intervention adherence and outcomes.

PS also provide simple classification for summarizing different longitudinal compliance patterns. In PROSPECT, there are 32 possible longitudinal compliance patterns. It may not be clinically useful to examine subject characteristics and outcomes within each of the compliance patterns, and such inferences may be sensitive to model assumptions when the number of subjects exhibiting the patterns is small. We used two latent compliance superclasses to describe the compliance patterns in the data.

We examined the relationship between the latent compliance superclass and baseline patient characteristics. It may be useful for clinicians to be able to identify patients with different compliance tendencies. Targeting patients who are more likely to adhere to particular interventions allows clinicians to tailor interventions to patients and optimize patient adherence and outcomes. The ability to identify patients who are less likely to adhere to particular interventions will alleviate inefficient use of clinical staff time and medical resources. The intervention in the PROSPECT entails meeting with health specialists. In this analysis we find that those in better physical health, indicated by higher SF-36 physical component scores, are more likely to comply with the intervention arm to which they are randomized over time. It may be easier for those in good physical health to attend meetings with health specialists than those in bad physical health.

We also examined the relationship between the latent compliance superclass and all-cause mortality. The analysis based on the latent compliance superclass and based on the observed compliance PS both showed that higher compliance are associated with lower mortality. In addition, this relationship is unlikely to be explained by the receipt of the intervention. This association between subject compliance and mortality was also found in other studies [10–13], although these studies were based on observed post-randomization intervention adherence rather than baseline compliance PS. There may be some underlying factors associated with a person's tendency to comply with treatment and mortality. Simpson *et al.* [13] suggested that subject treatment adherence may be a surrogate marker for overall healthy behavior.

A reviewer raised the point that the association between compliance and mortality may be due to the association between higher compliance and better physical health. It is also possible that compliance is related to innate traits that promote healthy behaviors and vitality. While our current methodology allows us to examine the associations between compliance, patient characteristics, and mortality, it does not tell us why certain compliance patterns lead to lower mortality. However, the findings may still serve as an useful tool for characterizing groups of subjects for whom the intervention may be successful, and characterizing subjects who may be at risk for treatment noncompliance and mortality.

In studies where subjects randomized to the usual care arm have no access to the intervention, compliance classes are observed in those randomized to the intervention arm. We could compare baseline subject characteristics and outcomes in the subset of subjects who are randomized to the intervention arm. However, note that the confidence intervals associated with the odds ratios when

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

looking at the baseline covariates and the hazard ratios in the mortality analysis are narrower for the estimates based on the latent compliance superclass than the estimates based on the observed compliance PS. Utilizing latent principal stratification strategy in forming compliance classes allows the inclusion of usual care subjects in the analysis, thus increase efficiency in the estimates as demonstrated in our analyses and also potentially reducing bias.

The MCC model described in stage 1 of the analysis method section is computationally intensive. In this paper, we performed a three-stage *ad hoc* analysis comparing baseline subject characteristics and mortality between the latent compliance superclasses. We treat the superclass assignment as missing data and multiply impute superclasses from the subject-level posterior predictive distribution of belonging to a given superclass. This type of analysis provides a fast way to pinpoint possible predictors of longitudinal compliance and the effect of longitudinal compliance on subsequent outcomes such as mortality. Ideally, we want to incorporate these predictors into the MCC model. Future research will entail fitting the MCC model using physical health as a predictor of the compliance superclass, and using both depression severity and mortality as joint outcomes.

APPENDIX

We describe the priors used in the stage 1 MCMC estimation of the model parameters.

Let $\beta = [\lambda_{1c0}, \ldots, \lambda_{5n1}, \boldsymbol{\gamma}]^{\mathrm{T}}$ denote the fixed effects. We assume the conjugate priors $\beta \sim \mathrm{MVN}(\mu_\beta, \Sigma_\beta)$ and $\sigma^2 \sim \mathrm{Inv}-\chi^2(\mathrm{df} = v_\sigma, \psi)$. We assume $\boldsymbol{\varphi}_i \sim \mathrm{MVN}(\mathbf{0}, \Sigma_\varphi)$ for the subject-level random effects, and the hyperprior $\Sigma_\varphi \sim \mathrm{Inv}-\mathrm{Wishart}(\mathrm{df} = v_\varphi, \Gamma)$. For compliance superclass and compliance class probabilities, we assume the priors $(p_1, p_2) \sim \mathrm{Dirichlet}(a_1, a_2)$, $\boldsymbol{\alpha} \sim \mathrm{MVN}(0, \Sigma_\alpha)$, and $(\pi_{kj\eta' c}, \pi_{kj\eta' n}) \sim \mathrm{Dirichlet}(b_c, b_n)$ $\forall k, j, \eta'$ where k denotes the kth superclass.

ACKNOWLEDGEMENTS

The authors wish to thank the reviewers, Dr Joseph Gallo, and Prof. Paul Rosenbaum for their insightful comments that helped to improve this manuscript.

REFERENCES

1. Lin JY, Ten Have TR, Elliott MR. *Nested Markov Compliance Class Model in the Presence of Time-Varying Noncompliance*. http://www.biostatsresearch.com/upennbiostat/papers/, Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA, 2006.
2. Bruce ML, Ten Have TR, Reynolds CF, Katz IR, Schulberg HC, Mulstant BH, Brown GK, McAvay GJ, Pearson JL, Alexopoulos GS. Reducing suicidal ideation and depressive symptoms in depressed older primary care patients. *Journal of the American Medical Association* 2004; **291**:1081–1091.
3. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 1996; **91**:444–455.
4. Imbens GW, Rubin DB. Bayesian inference for causal effects in randomized experiments with noncompliance. *The Annals of Statistics* 1997; **25**:305–327.
5. Frangakis CE, Rubin DB. Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika* 1999; **86**:365–379.
6. Frangakis CE, Rubin DB. Principal stratification in causal inference. *Biometrics* 2002; **58**:21–29.
7. Bogner HR, Lin JY, Morales KH. Patterns of early adherence to the antidepressant citalopram among older primary care patients: the PROSPECT study. *The International Journal of Psychiatry in Medicine* 2006; **36**:103–119.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim

8. Gallo JJ, Bogner HR, Morales KH, Post EP, Ten Have TR, Bruce ML. Depression, cardiovascular disease, diabetes, and 2-year mortality among older primary care patients. *American Journal of Geriatric Psychiatry* 2005; **13**(9):748–755.
9. Schulz R, Drayer RA, Rollman BL. Depression as a risk factor for non-suicide morality in the elderly. *Biological Psychiatry* 2002; **52**:205–225.
10. Coronary Drug Project Investigators. Influence of adherence to treatment and response of cholesterol on mortality in the coronary drug project. *New England Journal of Medicine* 1980; **303**:1038–1041.
11. Horwitz RI, Viscoli CM, Donaldson RM, Murray CJ, Ransohoff DF, Horwitz RI, Berkman L, Horwitz SM, Ransohoff DF, Sindelar J. Treatment adherence and risk of death after a myocardial infarction. *Lancet* 1990; **336**:542–545.
12. Gump BB, Mathews KA. Special intervention reduces CVD mortality for adherent participants in the multiple risk factor intervention trial. *Annals of Behavioral Medicine* 2003; **26**:61–68.
13. Simpson SH, Eurich DT, Majumdar SR, Padwal RS, Tsuyuki RT, Varney J, Johnson JA. A meta-analysis of the association between adherence to drug therapy and mortality. *British Medical Journal* 2006; **333**:15–20.
14. Geman S, Geman D. Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1984; **6**:721–741.
15. Gelfand AE, Smith AFM. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* 1990; **85**:398–409.
16. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis* (2nd edn). Chapman & Hall: New York, 2004.
17. Hasting WK. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 1970; **57**:97–109.
18. Peng Y, Little RJA, Raghunathan TE. An extended general location model for causal inference from data subject to noncompliance and missing values. *Biometrics* 2004; **60**:598–607.
19. Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. Wiley: New Jersey, 2002.
20. Holm S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 1979; **6**:65–70.
21. Cohen SJ. *New Directions in Patient Compliance*. Lexington Books: Lexington, MA, 1979.
22. Lane SD. Compliance satisfaction and physician–patient communication. In *Communication Year Book 7*, Bostrom RN (ed.). Sage: Beverly Hills, CA, 1983; 772–799.
23. Perkins DO. Predictors of noncompliance in patients with schizophrenia. *Journal of Clinical Psychiatry* 2002; **63**:1121–1128.
24. van Servellen G, Chang B, Garcia L, Lombardi E. Individual and system level factors associated with treatment nonadherence in human immunodeficiency virus-infected men and women. *AIDS Patient Care STDS* 2002; **16**:269–281.

Copyright © 2007 John Wiley & Sons, Ltd.

DOI: 10.1002/sim