

## Appendix D. Diagnostic Test Simulation

## 4 Appendix D: Diagnostic test simulation

We simulated data from both linear mixed models and logistic mixed models with the specific parameter values given below. Thinking of time as representing an approximately weekly measurement and a study of consisting of 4 years of follow-up, we simulated data with 50 time units per “year.” We considered situations in which all the visits were “irregular” as well as a mix of “irregular” and “regular” visits; in the latter situation, up to 5 regular visits were possible at time 0 (baseline) and “yearly” thereafter (at times 50, 100, 150 and 200). It is common in clinical data collection to have a planned and “regular” visit schedule from which it is simple to classify observations as “irregular” or “regular.” In addition, this fact is sometimes recorded as part of the data collection.

We considered three forms of outcome dependence. The first two allowed dependence on the random effects, either the random effects portion of the linear predictor ( $z_{ij}^T b_i$ ) or the random intercept,  $b_{0i}$ . The third, perhaps the most realistic, was to allow the probability of observing the outcome,  $P\{R_{ij} = 1\}$ , to depend on a lagged (but typically unobserved) value of the outcome process. This reflects the situation in which patients with poorer values of the outcome schedule a visit in the future that triggers measurement of the outcome. In each of the three cases we standardized the values to have standard deviation 1 (to allow comparability) and used a logistic regression model. For example, in the lagged outcome scenario we constructed

$P\{R_{ij} = 1\}$  to depend on the outcome lagged by 5 time units,  $Y_{i,j-5}$ . Let  $Y_{i,j-5}^*$  denote the standardized value of  $Y_{i,j-5}$ . Our outcome dependent visit model was then

$$\text{logit}(P\{R_{ij} = 1\}) = \delta_{0,j} + \delta_Y Y_{i,j-5}^*. \quad (9)$$

In the all irregular visit scenario  $\delta_{0,j}$  was not dependent on  $j$ ; in the mixed regular and irregular visit scenario,  $\delta_{0,j}$  was given much larger values for  $j = 0, 50, 100, 150,$  and  $200$ . The specific values of the  $\delta_{0,j}$  (given below) were chosen to achieve different average sample sizes. We simulated a range of values of  $\delta_Y$  ranging from 0 (no outcome dependence) to 0.4, which reflects moderately strong outcome dependence [1, ].

Our outcome process was a simple version of the typical longitudinal data analysis model. The linear predictor had a binary group effect,  $G$ , with a 30% of the subjects having  $G = 1$ , a linear time effect ( $t$ ) and a group by time interaction. In addition we included random intercepts  $b_{0i}$  and random slopes with time  $b_{1i}$ :

$$\eta_{ij} = (\beta_0 + b_{0i}) + \beta_G G + (\beta_T + b_{1i})t + \beta_I G \times t. \quad (10)$$

## 4.1 Diagnostic test methods

Our general strategy will be as follows. First fit a generalized linear mixed model ignoring the potentially outcome-dependent visit process. Using the

estimated best predicted values of the random effects, form a test of association of some function of the estimated best predicted values and the visit process.

We assessed the following test statistics in the simulation studies described below:

- Test association of  $V_{ij}$  (=inter-visit time) with  $z_{ij}^T \tilde{b}_i$  (denoted IVTzb below) or both  $\tilde{b}_{0i}$  and  $\tilde{b}_{1i}$  (denoted IVT $b_0b_1$  in the tables below). The test was based on linear regression of the log transformation (due to skewness) of  $V_{ij}$  on the specified function of the best predicted values, adjusting for the covariates and using robust standard errors to account for the multiple inter-visit times per person. Because the time to the first visit and the time from the last visit are censored, they were omitted. Also, when the simulation included regular visits they were omitted from the calculation.
- Test association of  $n_i$  (=total number of visits for person  $i$ ) with the average value of  $z_{ij}^T \tilde{b}_i$  (denoted  $n_{izb}$  in the tables below) or both  $\tilde{b}_{0i}$  and  $\tilde{b}_{1i}$  (denoted  $N_i b_0 b_1$  in the tables below). The test was based on linear regression of  $n_i$  on the specified function of the best predicted values, adjusting for the covariates. A log transformation (after adding 0.5) was used because of the highly skewed distribution of the  $n_i$ . Again, when the simulation included regular visits they were omitted from the calculation.

- Test association of  $t_{i,j}$  (=visit time  $j$  for person  $i$ ) with the time-varying value of  $z_{ij}^T \tilde{b}_i$  (denoted Coxzb in the tables below) or both  $\tilde{b}_{0i}$  and  $\tilde{b}_{1i}$  (denoted Cox $b_0b_1$  in the tables below). The test was based on Cox regression of  $t_{ij}$  on the specified function of the best predicted values, adjusting for the covariates (excluding time because that is incorporated in the baseline hazard) and using robust standard errors to account for the multiple times per person. The time scale was used rather than the inter-visit time to accommodate regular visits as part of the baseline hazard. In this way, a high-intensity of visits at, say, the one year mark, can be accommodated. On an intervisit time scale, one year visits would not “line up”.
- While not motivated by the score statistic developed here, it has been suggested to simply include the cluster size ( $n_i$  – denoted GEE $n_i$  in the tables below) or the cumulative sample size up to time  $t_{i,j}$  as a covariate in the fitted model. We will denote the sample size up to time  $j$  for person  $i$  as  $n_{ij}^*$  and denote the diagnostic method in the tables below as GEE $n_i^*$ . This does not serve to reduce bias in the other predictors; in fact it can introduce bias. However, it might suffice as a diagnostic method. Preliminary assessment showed that the tests based on GEE fits had higher power so we report the assessment of including  $n_i$  or  $n_{ij}^*$  as a covariate in a GEE independence fit and using Wald tests based on robust standard errors.

## 4.2 Stata code for diagnostic test methods

Below we give the Stata code that illustrates reading in a dataset and performing the diagnostic tests described above.

```
clear all
set more off

*Import data for example
use kim_data.dta

*Variables are
*exam_date = date of exam
*subjid = ID variable for patient
*sex = sex of patient, coded as 0=male, 1=female
*MRSoutcome = modified Rankin scale, the outcome
*reg_visit = indicator of a regular visit coded as 0=no, 1=yes

*Calculate years since first visit
bysort subjid: egen first_date=min(exam_date)
bysort subjid: gen days_since=exam_date-first_date
gen year=days_since/365

*Fit model ignoring selection
mixed MRSoutcome c.year##i.sex c.year#c.year || subjid: year, cov(uns) iterate(300) tech(bfgs)

predict pred_xb
predict pred_bp, fitted
```

```

predict bp*, ref

*Diagnostic tests
gen failure=1

*Check association between interarrival times and random effects
*Diff between fixed and fitted predictions
gen diff=pred_bp-pred_xb

stset year, failure(failure) id(subjid) exit(time .)

*Analyses adjusted for covariates (not for year effects for Cox)
*Combined effect of best predicted values
stcox bp*, vce(robust) strata(sex)
testparm bp*

*Random slope only
stcox bp1, vce(robust) strata(sex)
testparm bp*

*Random intercept only
stcox bp2, vce(robust) strata(sex)
testparm bp*

*Random effects portion of linear predictor
stcox diff, vce(robust) strata(sex)
testparm diff

```

```

*Tests based on intervisit times
*First calculate all intervisit times
*Only consider times between or number of (below) irregular visits
drop if reg_visit
*Sort data
sort subjid year
*Exclude first by looking ahead (censored)
bysort subjid: gen ivta=year[_n+1]-year[_n]
*Exclude last visit time since censored
bysort subjid: replace ivta=. if _n==_N

*Intervisit times are highly skewed. Log transform
gen log_ivta=log(ivta)

*Adjusted for covariates
regress log_ivta diff c.year##i.sex, cluster(subjid)
testparm diff

regress log_ivta bp1 c.year##i.sex, cluster(subjid)
testparm bp1

regress log_ivta bp2 c.year##i.sex, cluster(subjid)
testparm bp2

regress log_ivta bp1 bp2 c.year##i.sex, cluster(subjid)
testparm bp*

*Tests based on n_i (number of observations per person)

```



```

*Save total length of follow-up to adjust n_i
bysort subjid: gen last_date=exam_date[_N]
gen fu_time=last_date-first_date
gen fu_years=fu_time/365

*Calculate cluster size and cumulative sample size for tests
bysort subjid: gen cumssize=_n-1
bysort subjid: gen clussize=_N

*Fit model using GEE but including cumulative sample size
xtgee y c.year##i.sex c.year#c.year cumssize, i(subjid) corr(inde) robust

*Fit model using GEE but including cluster sample size
xtgee y c.year##i.sex c.year#c.year clussize, i(subjid) corr(inde) robust

collapse (count) n_i=y (mean) pred_xb pred_bp diff ivta log_ivta (first) fu_years bp1 bp2 sex, b
*Regression of n_i on both b0 and b1
*Adjusted for covariates

*Calculate sample size per follow-up time
gen n_i_per=n_i/fu_years

*Skewed so log transform after adding 0.5
gen log_n_i_per=log(n_i_per+0.5)
*Use robust standard errors in case heteroscedastic

regress log_n_i_per diff i.sex, cluster(subjid)
testparm diff

```

```
regress log_n_i_per bp1 i.sex, cluster(subjid)
testparm bp1
```

```
regress log_n_i_per bp2 i.sex, cluster(subjid)
testparm bp2
```

```
regress log_n_i_per bp1 bp2 i.sex, cluster(subjid)
testparm bp*
```

### **4.3 Simulation results for diagnostic test methods in linear mixed models**

The scenarios below vary by sample sizes and types of outcome dependent visit processes. For each scenario there is a set of three tables. The first gives the power for the above described tests. The second and third tables give the estimated mean values of the MLE and GEE-independence estimators, respectively.

Tests based on the GEE fits did not do a good job of controlling the type I error rate. At a nominal level of 0.05 the achieved error rate ranged from 0.056 to 0.094 across the simulations reported here. The tests based on the cluster size were well-controlled; no achieved error rates exceeded 0.07 across all the simulations. The tests based on intervisit times and the Cox model fits were intermediate with a few of the simulations generating achieved error rates as high as 0.075.

Table 29: Power of various tests for an outcome dependent visit process when the outcome follows a linear mixed model with  $m = 100$  subjects and an average sample size of 5. Outcome dependence is on a lagged value of the outcome. Results are presented for the case of all irregular visits (top panel) or a mix of regular and irregular visits (bottom panel). The tests look for dependence of the actual visit times (Cox), intervisit times (IVT) or the number of visits ( $n_i$  or  $n_{ij}^*$ ) on the best predicted values of the random intercept and slope (b0b1) or the random effects portion of the linear predictor (zb).

Informative Visit Process	Power								
	$\delta_Y$	Coxb <sub>0</sub> b <sub>1</sub>	Coxzb	$N_i$ b <sub>0</sub> b <sub>1</sub>	$n_i$ zb	IVTb <sub>0</sub> b <sub>1</sub>	IVTzb	GEEn <sub>i</sub> <sup>*</sup>	GEEn <sub>i</sub>
Irregular visits									
	0.00	0.064	0.049	0.048	0.052	0.075	0.065	0.075	0.076
	0.10	0.144	0.116	0.160	0.206	0.146	0.130	0.152	0.216
	0.20	0.328	0.284	0.546	0.624	0.286	0.298	0.424	0.608
	0.25	0.412	0.380	0.652	0.750	0.332	0.404	0.574	0.752
	0.30	0.592	0.618	0.868	0.908	0.442	0.530	0.742	0.902
	0.35	0.706	0.706	0.948	0.964	0.612	0.698	0.854	0.966
	0.40	0.842	0.844	0.982	0.990	0.726	0.840	0.944	0.986
Mixed visits									
	0.00	0.074	0.058	0.056	0.058	0.098	0.064	0.089	0.093
	0.10	0.136	0.146	0.098	0.108	0.076	0.060	0.118	0.184
	0.20	0.356	0.392	0.298	0.362	0.114	0.082	0.230	0.346
	0.25	0.528	0.576	0.438	0.516	0.120	0.096	0.338	0.506
	0.30	0.676	0.738	0.578	0.642	0.172	0.152	0.458	0.648
	0.35	0.828	0.868	0.702	0.788	0.162	0.198	0.572	0.764
	0.40	0.904	0.936	0.832	0.892	0.202	0.252	0.694	0.882

Table 30: Mean values of parameter estimates from a maximum likelihood linear mixed model fit that ignores outcome dependence when the outcome follows a linear mixed model with  $m = 100$  subjects and an average sample size of 5. Outcome dependence is on a lagged value of the outcome. Results are presented for the case of all irregular visits (top) or a mix of regular and irregular visits (bottom) and a range of outcome dependence,  $\delta_Y$ .

Informative Visit Process	Simulated mean parameter estimates (SEs as subscripts)				
	$\delta_Y$	$\beta_0$ (true=0)	$\beta_T$ (true=2)	$\beta_G$ (true=1)	$\beta_I$ (true=1.5)
Irregular visits					
	0.00	0.003 <sub>0.005</sub>	1.986 <sub>0.009</sub>	1.001 <sub>0.007</sub>	1.500 <sub>0.007</sub>
	0.10	0.002 <sub>0.010</sub>	1.984 <sub>0.017</sub>	0.999 <sub>0.014</sub>	1.494 <sub>0.013</sub>
	0.20	0.023 <sub>0.010</sub>	1.969 <sub>0.018</sub>	1.012 <sub>0.014</sub>	1.449 <sub>0.013</sub>
	0.25	0.006 <sub>0.010</sub>	1.995 <sub>0.018</sub>	1.019 <sub>0.014</sub>	1.488 <sub>0.012</sub>
	0.30	0.019 <sub>0.010</sub>	1.997 <sub>0.018</sub>	1.036 <sub>0.014</sub>	1.479 <sub>0.013</sub>
	0.35	0.028 <sub>0.010</sub>	2.008 <sub>0.018</sub>	1.002 <sub>0.014</sub>	1.461 <sub>0.013</sub>
	0.40	0.034 <sub>0.011</sub>	2.000 <sub>0.017</sub>	1.020 <sub>0.014</sub>	1.470 <sub>0.013</sub>
Mixed visits					
	0.00	0.003 <sub>0.004</sub>	1.997 <sub>0.006</sub>	1.000 <sub>0.006</sub>	1.491 <sub>0.006</sub>
	0.10	0.000 <sub>0.007</sub>	2.025 <sub>0.011</sub>	0.996 <sub>0.012</sub>	1.497 <sub>0.011</sub>
	0.20	0.021 <sub>0.008</sub>	2.009 <sub>0.012</sub>	0.986 <sub>0.013</sub>	1.521 <sub>0.011</sub>
	0.25	-0.001 <sub>0.008</sub>	2.023 <sub>0.012</sub>	0.999 <sub>0.013</sub>	1.483 <sub>0.011</sub>
	0.30	0.004 <sub>0.008</sub>	2.035 <sub>0.011</sub>	0.990 <sub>0.013</sub>	1.488 <sub>0.011</sub>
	0.35	0.004 <sub>0.008</sub>	2.014 <sub>0.011</sub>	1.003 <sub>0.013</sub>	1.499 <sub>0.012</sub>
	0.40	0.003 <sub>0.008</sub>	2.017 <sub>0.012</sub>	1.017 <sub>0.012</sub>	1.491 <sub>0.011</sub>

Table 31: Mean values of parameter estimates from a GEE independence linear model fit that ignores outcome dependence when the outcome follows a linear mixed model with  $m = 100$  subjects and an average sample size of 5. Outcome dependence is on a lagged value of the outcome. Results are presented for the case of all irregular visits (top) or a mix of regular and irregular visits (bottom) and a range of outcome dependence,  $\delta_Y$ .

Informative Visit Process	Simulated mean parameter estimates (SEs as subscripts)				
	$\delta_Y$	$\beta_0$ (true=0)	$\beta_T$ (true=2)	$\beta_G$ (true=1)	$\beta_I$ (true=1.5)
Irregular visits					
	0.00	-0.008 <sub>0.006</sub>	2.014 <sub>0.015</sub>	1.011 <sub>0.009</sub>	1.489 <sub>0.009</sub>
	0.10	0.035 <sub>0.012</sub>	2.090 <sub>0.028</sub>	0.981 <sub>0.018</sub>	1.484 <sub>0.017</sub>
	0.20	0.042 <sub>0.013</sub>	2.085 <sub>0.029</sub>	1.032 <sub>0.019</sub>	1.481 <sub>0.019</sub>
	0.25	0.066 <sub>0.013</sub>	2.094 <sub>0.030</sub>	1.020 <sub>0.018</sub>	1.480 <sub>0.018</sub>
	0.30	0.107 <sub>0.013</sub>	2.075 <sub>0.028</sub>	1.010 <sub>0.017</sub>	1.473 <sub>0.019</sub>
	0.35	0.084 <sub>0.013</sub>	2.154 <sub>0.027</sub>	1.004 <sub>0.020</sub>	1.477 <sub>0.019</sub>
	0.40	0.140 <sub>0.014</sub>	2.127 <sub>0.030</sub>	0.967 <sub>0.018</sub>	1.486 <sub>0.019</sub>
Mixed visits					
	0.00	-0.010 <sub>0.008</sub>	2.004 <sub>0.010</sub>	1.003 <sub>0.011</sub>	1.503 <sub>0.008</sub>
	0.10	0.033 <sub>0.017</sub>	2.016 <sub>0.040</sub>	1.015 <sub>0.024</sub>	1.495 <sub>0.024</sub>
	0.20	0.049 <sub>0.019</sub>	2.115 <sub>0.042</sub>	1.016 <sub>0.023</sub>	1.499 <sub>0.025</sub>
	0.25	0.068 <sub>0.018</sub>	2.097 <sub>0.042</sub>	1.024 <sub>0.024</sub>	1.506 <sub>0.024</sub>
	0.30	0.083 <sub>0.018</sub>	2.161 <sub>0.040</sub>	0.982 <sub>0.023</sub>	1.469 <sub>0.024</sub>
	0.35	0.124 <sub>0.019</sub>	2.136 <sub>0.042</sub>	0.981 <sub>0.024</sub>	1.494 <sub>0.024</sub>
	0.40	0.130 <sub>0.019</sub>	2.175 <sub>0.041</sub>	0.987 <sub>0.026</sub>	1.480 <sub>0.024</sub>

Table 32: Power of various tests for an outcome dependent visit process when the outcome follows a linear mixed model with  $m = 100$  subjects and an average sample size of 9. Outcome dependence is on a lagged value of the outcome. Results are presented for the case of all irregular visits (top panel) or a mix of regular and irregular visits (bottom panel). The tests look for dependence of the actual visit times (Cox), intervisit times (IVT) or the number of visits ( $n_i$  or  $n_{ij}^*$ ) on the best predicted values of the random intercept and slope (b0b1) or the random effects portion of the linear predictor (zb).

Informative Visit Process	Power								
	$\delta_Y$	Coxb <sub>0</sub> b <sub>1</sub>	Coxzb	$N_i$ b <sub>0</sub> b <sub>1</sub>	$n_i$ zb	IVTb <sub>0</sub> b <sub>1</sub>	IVTzb	GEEn <sub>i</sub> <sup>*</sup>	GEEn <sub>i</sub>
Irregular visits									
	0.00	0.079	0.043	0.050	0.056	0.072	0.066	0.071	0.082
	0.10	0.192	0.204	0.268	0.328	0.162	0.232	0.218	0.322
	0.20	0.604	0.604	0.790	0.858	0.512	0.650	0.644	0.852
	0.25	0.798	0.838	0.950	0.970	0.740	0.846	0.802	0.964
	0.30	0.940	0.958	0.988	1.000	0.870	0.956	0.916	0.990
	0.35	0.980	0.986	0.998	0.998	0.966	0.990	0.978	1.000
	0.40	0.998	0.998	1.000	1.000	0.994	0.998	0.994	1.000
Mixed visits									
	0.00	0.075	0.045	0.059	0.061	0.085	0.075	0.090	0.094
	0.10	0.192	0.238	0.164	0.204	0.064	0.076	0.120	0.198
	0.20	0.596	0.718	0.546	0.576	0.194	0.232	0.394	0.590
	0.25	0.802	0.870	0.718	0.804	0.262	0.320	0.506	0.736
	0.30	0.916	0.968	0.854	0.890	0.342	0.452	0.698	0.886
	0.35	0.976	0.994	0.946	0.970	0.456	0.616	0.810	0.960
	0.40	0.994	0.998	0.984	0.994	0.562	0.730	0.922	0.992

Table 33: Mean values of parameter estimates from a maximum likelihood linear mixed model fit that ignores outcome dependence when the outcome follows a linear mixed model with  $m = 100$  subjects and an average sample size of 9. Outcome dependence is on a lagged value of the outcome. Results are presented for the case of all irregular visits (top) or a mix of regular and irregular visits (bottom) and a range of outcome dependence,  $\delta_Y$ .

Informative Visit Process	Simulated mean parameter estimates (SEs as subscripts)				
	$\delta_Y$	$\beta_0$ (true=0)	$\beta_T$ (true=2)	$\beta_G$ (true=1)	$\beta_I$ (true=1.5)
Irregular visits					
	0.00	-0.001 <sub>0.004</sub>	2.000 <sub>0.004</sub>	0.997 <sub>0.006</sub>	1.494 <sub>0.005</sub>
	0.10	0.006 <sub>0.008</sub>	1.997 <sub>0.009</sub>	1.013 <sub>0.012</sub>	1.488 <sub>0.010</sub>
	0.20	0.014 <sub>0.008</sub>	1.990 <sub>0.008</sub>	0.997 <sub>0.012</sub>	1.496 <sub>0.010</sub>
	0.25	0.022 <sub>0.008</sub>	2.012 <sub>0.008</sub>	0.985 <sub>0.012</sub>	1.478 <sub>0.010</sub>
	0.30	0.034 <sub>0.009</sub>	1.977 <sub>0.008</sub>	0.975 <sub>0.012</sub>	1.494 <sub>0.010</sub>
	0.35	0.028 <sub>0.009</sub>	1.990 <sub>0.008</sub>	0.992 <sub>0.012</sub>	1.490 <sub>0.010</sub>
	0.40	0.020 <sub>0.009</sub>	1.992 <sub>0.008</sub>	1.014 <sub>0.012</sub>	1.493 <sub>0.009</sub>
Mixed visits					
	0.00	-0.004 <sub>0.003</sub>	2.001 <sub>0.004</sub>	1.005 <sub>0.006</sub>	1.500 <sub>0.005</sub>
	0.10	0.001 <sub>0.007</sub>	1.995 <sub>0.007</sub>	0.992 <sub>0.011</sub>	1.507 <sub>0.010</sub>
	0.20	0.002 <sub>0.007</sub>	2.002 <sub>0.007</sub>	1.002 <sub>0.011</sub>	1.499 <sub>0.010</sub>
	0.25	-0.011 <sub>0.007</sub>	1.994 <sub>0.007</sub>	1.018 <sub>0.012</sub>	1.498 <sub>0.010</sub>
	0.30	0.006 <sub>0.007</sub>	2.003 <sub>0.007</sub>	1.005 <sub>0.011</sub>	1.508 <sub>0.010</sub>
	0.35	-0.000 <sub>0.007</sub>	1.999 <sub>0.007</sub>	1.008 <sub>0.011</sub>	1.513 <sub>0.010</sub>
	0.40	0.008 <sub>0.007</sub>	1.995 <sub>0.007</sub>	1.003 <sub>0.012</sub>	1.497 <sub>0.010</sub>

Table 34: Mean values of parameter estimates from a GEE independence linear model fit that ignores outcome dependence when the outcome follows a linear mixed model with  $m = 100$  subjects and an average sample size of 9. Outcome dependence is on a lagged value of the outcome. Results are presented for the case of all irregular visits (top) or a mix of regular and irregular visits (bottom) and a range of outcome dependence,  $\delta_Y$ .

Informative Visit Process	Simulated mean parameter estimates (SEs as subscripts)				
	$\delta_Y$	$\beta_0$ (true=0)	$\beta_T$ (true=2)	$\beta_G$ (true=1)	$\beta_I$ (true=1.5)
Irregular visits					
	0.00	0.003 <sub>0.006</sub>	1.998 <sub>0.008</sub>	0.985 <sub>0.009</sub>	1.501 <sub>0.006</sub>
	0.10	0.028 <sub>0.012</sub>	2.015 <sub>0.016</sub>	0.995 <sub>0.018</sub>	1.494 <sub>0.013</sub>
	0.20	0.034 <sub>0.012</sub>	2.039 <sub>0.017</sub>	1.004 <sub>0.018</sub>	1.500 <sub>0.013</sub>
	0.25	0.031 <sub>0.013</sub>	2.108 <sub>0.018</sub>	1.000 <sub>0.018</sub>	1.459 <sub>0.013</sub>
	0.30	0.071 <sub>0.013</sub>	2.052 <sub>0.016</sub>	0.965 <sub>0.018</sub>	1.491 <sub>0.013</sub>
	0.35	0.059 <sub>0.012</sub>	2.094 <sub>0.015</sub>	1.018 <sub>0.018</sub>	1.473 <sub>0.014</sub>
	0.40	0.052 <sub>0.014</sub>	2.117 <sub>0.017</sub>	1.034 <sub>0.019</sub>	1.479 <sub>0.013</sub>
Mixed visits					
	0.00	-0.009 <sub>0.008</sub>	2.001 <sub>0.011</sub>	1.000 <sub>0.011</sub>	1.505 <sub>0.008</sub>
	0.10	0.016 <sub>0.016</sub>	2.017 <sub>0.020</sub>	0.965 <sub>0.024</sub>	1.513 <sub>0.015</sub>
	0.20	0.052 <sub>0.017</sub>	2.039 <sub>0.022</sub>	0.965 <sub>0.024</sub>	1.522 <sub>0.017</sub>
	0.25	0.029 <sub>0.017</sub>	2.039 <sub>0.023</sub>	1.016 <sub>0.024</sub>	1.505 <sub>0.015</sub>
	0.30	0.048 <sub>0.017</sub>	2.095 <sub>0.021</sub>	0.990 <sub>0.023</sub>	1.501 <sub>0.016</sub>
	0.35	0.055 <sub>0.017</sub>	2.068 <sub>0.023</sub>	1.049 <sub>0.024</sub>	1.486 <sub>0.016</sub>
	0.40	0.092 <sub>0.018</sub>	2.046 <sub>0.022</sub>	1.042 <sub>0.025</sub>	1.463 <sub>0.016</sub>



Table 35: Power of various tests for an outcome dependent visit process when the outcome follows a linear mixed model with  $m = 100$  subjects and an average sample size of 5. Outcome dependence is on the conditional linear predictor. Results are presented for the case of all irregular visits (top panel) or a mix of regular and irregular visits (bottom panel). The tests look for dependence of the actual visit times (Cox), intervisit times (IVT) or the number of visits ( $n_i$  or  $n_{ij}^*$ ) on the best predicted values of the random intercept and slope (b0b1) or the random effects portion of the linear predictor (zb).

Informative Visit Process	Power								
	$\delta_Y$	Coxb <sub>0</sub> b <sub>1</sub>	Coxzb	$N_i$ b <sub>0</sub> b <sub>1</sub>	$n_i$ zb	IVTb <sub>0</sub> b <sub>1</sub>	IVTzb	GEEn <sub>i</sub> <sup>*</sup>	GEEn <sub>i</sub>
Irregular visits									
	0.00	0.064	0.049	0.048	0.052	0.075	0.065	0.075	0.076
	0.10	0.154	0.126	0.184	0.230	0.136	0.140	0.164	0.238
	0.20	0.368	0.334	0.646	0.688	0.310	0.342	0.476	0.658
	0.25	0.474	0.454	0.734	0.816	0.368	0.438	0.660	0.838
	0.30	0.680	0.680	0.902	0.946	0.528	0.634	0.832	0.944
	0.35	0.798	0.818	0.978	0.986	0.702	0.802	0.904	0.982
	0.40	0.918	0.920	0.996	0.998	0.844	0.916	0.970	0.994
Mixed visits									
	0.00	0.074	0.058	0.056	0.058	0.098	0.064	0.089	0.093
	0.10	0.150	0.158	0.110	0.130	0.084	0.064	0.140	0.202
	0.20	0.438	0.460	0.352	0.406	0.100	0.106	0.272	0.398
	0.25	0.572	0.674	0.500	0.568	0.140	0.146	0.410	0.546
	0.30	0.760	0.794	0.666	0.734	0.186	0.178	0.526	0.706
	0.35	0.894	0.928	0.806	0.858	0.196	0.224	0.648	0.800
	0.40	0.946	0.964	0.906	0.932	0.244	0.324	0.774	0.918

Table 36: Mean values of parameter estimates from a maximum likelihood linear mixed model fit that ignores outcome dependence when the outcome follows a linear mixed model with  $m = 100$  subjects and an average sample size of 5. Outcome dependence is on the conditional linear predictor of the outcome. Results are presented for the case of all irregular visits (top) or a mix of regular and irregular visits (bottom) and a range of outcome dependence,  $\delta_Y$ .

Informative Visit Process	Simulated mean parameter estimates (SEs as subscripts)				
	$\delta_Y$	$\beta_0$ (true=0)	$\beta_T$ (true=2)	$\beta_G$ (true=1)	$\beta_I$ (true=1.5)
Irregular visits					
	0.00	0.003 <sub>0.005</sub>	1.986 <sub>0.009</sub>	1.001 <sub>0.007</sub>	1.500 <sub>0.007</sub>
	0.10	0.002 <sub>0.010</sub>	1.989 <sub>0.016</sub>	1.001 <sub>0.014</sub>	1.498 <sub>0.013</sub>
	0.20	0.026 <sub>0.010</sub>	1.968 <sub>0.018</sub>	1.007 <sub>0.014</sub>	1.493 <sub>0.012</sub>
	0.25	0.004 <sub>0.010</sub>	1.993 <sub>0.018</sub>	1.020 <sub>0.014</sub>	1.488 <sub>0.013</sub>
	0.30	0.017 <sub>0.010</sub>	2.007 <sub>0.018</sub>	1.034 <sub>0.014</sub>	1.478 <sub>0.013</sub>
	0.35	0.025 <sub>0.010</sub>	2.020 <sub>0.018</sub>	0.998 <sub>0.014</sub>	1.458 <sub>0.013</sub>
	0.40	0.034 <sub>0.010</sub>	2.003 <sub>0.017</sub>	1.021 <sub>0.014</sub>	1.462 <sub>0.013</sub>
Mixed visits					
	0.00	0.003 <sub>0.004</sub>	1.997 <sub>0.006</sub>	1.000 <sub>0.006</sub>	1.491 <sub>0.006</sub>
	0.10	-0.001 <sub>0.007</sub>	2.025 <sub>0.011</sub>	0.998 <sub>0.012</sub>	1.497 <sub>0.011</sub>
	0.20	0.020 <sub>0.008</sub>	2.010 <sub>0.012</sub>	0.988 <sub>0.012</sub>	1.520 <sub>0.011</sub>
	0.25	-0.002 <sub>0.008</sub>	2.027 <sub>0.012</sub>	1.000 <sub>0.013</sub>	1.481 <sub>0.011</sub>
	0.30	0.004 <sub>0.008</sub>	2.040 <sub>0.011</sub>	0.989 <sub>0.013</sub>	1.488 <sub>0.011</sub>
	0.35	0.003 <sub>0.008</sub>	2.016 <sub>0.011</sub>	1.005 <sub>0.013</sub>	1.498 <sub>0.012</sub>
	0.40	0.002 <sub>0.008</sub>	2.024 <sub>0.011</sub>	1.019 <sub>0.012</sub>	1.489 <sub>0.011</sub>

Table 37: Mean values of parameter estimates from a GEE independence linear model fit that ignores outcome dependence when the outcome follows a linear mixed model with  $m = 100$  subjects and an average sample size of 5. Outcome dependence is on the conditional linear predictor of the outcome. Results are presented for the case of all irregular visits (top) or a mix of regular and irregular visits (bottom) and a range of outcome dependence,  $\delta_Y$ .

Informative Visit Process	Simulated mean parameter estimates (SEs as subscripts)				
	$\delta_Y$	$\beta_0$ (true=0)	$\beta_T$ (true=2)	$\beta_G$ (true=1)	$\beta_I$ (true=1.5)
Irregular visits					
	0.00	-0.008 <sub>0.006</sub>	2.014 <sub>0.015</sub>	1.011 <sub>0.009</sub>	1.489 <sub>0.009</sub>
	0.10	0.038 <sub>0.012</sub>	2.094 <sub>0.028</sub>	0.982 <sub>0.018</sub>	1.482 <sub>0.017</sub>
	0.20	0.047 <sub>0.013</sub>	2.094 <sub>0.029</sub>	1.033 <sub>0.019</sub>	1.476 <sub>0.019</sub>
	0.25	0.075 <sub>0.013</sub>	2.105 <sub>0.030</sub>	1.022 <sub>0.018</sub>	1.479 <sub>0.018</sub>
	0.30	0.115 <sub>0.013</sub>	2.096 <sub>0.029</sub>	1.014 <sub>0.017</sub>	1.468 <sub>0.019</sub>
	0.35	0.099 <sub>0.013</sub>	2.166 <sub>0.027</sub>	1.000 <sub>0.019</sub>	1.478 <sub>0.018</sub>
	0.40	0.147 <sub>0.014</sub>	2.160 <sub>0.029</sub>	0.972 <sub>0.018</sub>	1.482 <sub>0.019</sub>
Mixed visits					
	0.00	-0.010 <sub>0.008</sub>	2.004 <sub>0.010</sub>	1.003 <sub>0.011</sub>	1.503 <sub>0.008</sub>
	0.10	0.040 <sub>0.017</sub>	2.010 <sub>0.040</sub>	1.014 <sub>0.024</sub>	1.496 <sub>0.024</sub>
	0.20	0.055 <sub>0.019</sub>	2.132 <sub>0.042</sub>	1.011 <sub>0.023</sub>	1.510 <sub>0.025</sub>
	0.25	0.071 <sub>0.018</sub>	2.123 <sub>0.042</sub>	1.025 <sub>0.024</sub>	1.504 <sub>0.024</sub>
	0.30	0.092 <sub>0.018</sub>	2.166 <sub>0.040</sub>	0.984 <sub>0.024</sub>	1.463 <sub>0.024</sub>
	0.35	0.131 <sub>0.019</sub>	2.159 <sub>0.042</sub>	0.988 <sub>0.023</sub>	1.495 <sub>0.024</sub>
	0.40	0.142 <sub>0.019</sub>	2.186 <sub>0.041</sub>	0.993 <sub>0.025</sub>	1.476 <sub>0.023</sub>

Table 38: Power of various tests for an outcome dependent visit process when the outcome follows a linear mixed model with  $m = 100$  subjects and an average sample size of 5. Outcome dependence is on the random intercept. Results are presented for the case of all irregular visits (top panel) or a mix of regular and irregular visits (bottom panel). The tests look for dependence of the actual visit times (Cox), intervisit times (IVT) or the number of visits ( $n_i$  or  $n_{ij}^*$ ) on the best predicted values of the random intercept and slope (b0b1) or the random effects portion of the linear predictor (zb).

Informative Visit Process	Power								
	$\delta_Y$	Coxb <sub>0</sub> b <sub>1</sub>	Coxzb	N <sub>i</sub> b <sub>0</sub> b <sub>1</sub>	n <sub>i</sub> zb	IVTb <sub>0</sub> b <sub>1</sub>	IVTzb	GEE n <sub>i</sub> <sup>*</sup>	GEE n <sub>i</sub>
Irregular visits									
	0.00	0.064	0.049	0.048	0.052	0.075	0.065	0.075	0.076
	0.10	0.246	0.190	0.274	0.332	0.212	0.194	0.292	0.370
	0.20	0.638	0.496	0.838	0.842	0.514	0.498	0.718	0.842
	0.25	0.750	0.650	0.898	0.910	0.634	0.636	0.876	0.954
	0.30	0.910	0.820	0.980	0.978	0.836	0.798	0.940	0.970
	0.35	0.972	0.908	0.988	0.992	0.916	0.902	0.984	0.992
	0.40	0.992	0.952	0.998	0.998	0.964	0.952	0.988	0.996
Mixed visits									
	0.00	0.074	0.058	0.056	0.058	0.098	0.064	0.089	0.093
	0.10	0.218	0.192	0.186	0.178	0.090	0.068	0.180	0.262
	0.20	0.622	0.574	0.526	0.530	0.184	0.150	0.440	0.532
	0.25	0.804	0.736	0.678	0.710	0.216	0.218	0.580	0.690
	0.30	0.916	0.844	0.832	0.848	0.314	0.280	0.724	0.830
	0.35	0.968	0.924	0.940	0.910	0.400	0.384	0.816	0.900
	0.40	0.992	0.962	0.974	0.966	0.480	0.450	0.894	0.960

Table 39: Mean values of parameter estimates from a maximum likelihood linear mixed model fit that ignores outcome dependence when the outcome follows a linear mixed model with  $m = 100$  subjects and an average sample size of 5. Outcome dependence is on the random intercept. Results are presented for the case of all irregular visits (top) or a mix of regular and irregular visits (bottom) and a range of outcome dependence,  $\delta_Y$ .

Informative Visit Process	Simulated mean parameter estimates (SEs as subscripts)				
	$\delta_Y$	$\beta_0$ (true=0)	$\beta_T$ (true=2)	$\beta_G$ (true=1)	$\beta_I$ (true=1.5)
Irregular visits					
	0.00	0.003 <sub>0.005</sub>	1.986 <sub>0.009</sub>	1.001 <sub>0.007</sub>	1.500 <sub>0.007</sub>
	0.10	0.019 <sub>0.010</sub>	1.989 <sub>0.016</sub>	1.001 <sub>0.014</sub>	1.500 <sub>0.013</sub>
	0.20	0.058 <sub>0.010</sub>	1.972 <sub>0.018</sub>	1.012 <sub>0.014</sub>	1.491 <sub>0.013</sub>
	0.25	0.058 <sub>0.010</sub>	1.972 <sub>0.017</sub>	1.015 <sub>0.014</sub>	1.509 <sub>0.013</sub>
	0.30	0.073 <sub>0.009</sub>	1.986 <sub>0.017</sub>	1.046 <sub>0.013</sub>	1.485 <sub>0.013</sub>
	0.35	0.099 <sub>0.010</sub>	1.985 <sub>0.017</sub>	1.002 <sub>0.014</sub>	1.485 <sub>0.013</sub>
	0.40	0.113 <sub>0.009</sub>	1.980 <sub>0.016</sub>	1.022 <sub>0.014</sub>	1.499 <sub>0.014</sub>
Mixed visits					
	0.00	0.003 <sub>0.004</sub>	1.997 <sub>0.006</sub>	1.000 <sub>0.006</sub>	1.491 <sub>0.006</sub>
	0.10	0.006 <sub>0.007</sub>	2.025 <sub>0.011</sub>	0.998 <sub>0.012</sub>	1.499 <sub>0.011</sub>
	0.20	0.035 <sub>0.007</sub>	2.011 <sub>0.011</sub>	0.989 <sub>0.012</sub>	1.521 <sub>0.011</sub>
	0.25	0.021 <sub>0.008</sub>	2.022 <sub>0.012</sub>	0.997 <sub>0.013</sub>	1.488 <sub>0.012</sub>
	0.30	0.026 <sub>0.008</sub>	2.040 <sub>0.011</sub>	0.990 <sub>0.013</sub>	1.493 <sub>0.011</sub>
	0.35	0.029 <sub>0.008</sub>	2.016 <sub>0.011</sub>	1.006 <sub>0.013</sub>	1.506 <sub>0.012</sub>
	0.40	0.038 <sub>0.007</sub>	2.023 <sub>0.011</sub>	1.012 <sub>0.012</sub>	1.501 <sub>0.012</sub>

Table 40: Mean values of parameter estimates from a GEE independence linear model fit that ignores outcome dependence when the outcome follows a linear mixed model with  $m = 100$  subjects and an average sample size of 5. Outcome dependence is on the random intercept. Results are presented for the case of all irregular visits (top) or a mix of regular and irregular visits (bottom) and a range of outcome dependence,  $\delta_Y$ .

Informative Visit Process	Simulated mean parameter estimates (SEs as subscripts)				
	$\delta_Y$	$\beta_0$ (true=0)	$\beta_T$ (true=2)	$\beta_G$ (true=1)	$\beta_I$ (true=1.5)
Irregular visits					
	0.00	-0.008 <sub>0.006</sub>	2.014 <sub>0.015</sub>	1.011 <sub>0.009</sub>	1.489 <sub>0.009</sub>
	0.10	0.089 <sub>0.012</sub>	2.112 <sub>0.027</sub>	0.992 <sub>0.018</sub>	1.477 <sub>0.017</sub>
	0.20	0.174 <sub>0.013</sub>	2.099 <sub>0.028</sub>	1.014 <sub>0.018</sub>	1.500 <sub>0.019</sub>
	0.25	0.222 <sub>0.012</sub>	2.132 <sub>0.029</sub>	1.002 <sub>0.017</sub>	1.496 <sub>0.018</sub>
	0.30	0.296 <sub>0.012</sub>	2.096 <sub>0.027</sub>	0.991 <sub>0.017</sub>	1.498 <sub>0.019</sub>
	0.35	0.313 <sub>0.012</sub>	2.169 <sub>0.027</sub>	0.985 <sub>0.019</sub>	1.514 <sub>0.018</sub>
	0.40	0.380 <sub>0.013</sub>	2.175 <sub>0.027</sub>	0.971 <sub>0.018</sub>	1.503 <sub>0.019</sub>
Mixed visits					
	0.00	-0.010 <sub>0.008</sub>	2.004 <sub>0.010</sub>	1.003 <sub>0.011</sub>	1.503 <sub>0.008</sub>
	0.10	0.093 <sub>0.016</sub>	2.034 <sub>0.040</sub>	1.016 <sub>0.024</sub>	1.493 <sub>0.024</sub>
	0.20	0.187 <sub>0.018</sub>	2.118 <sub>0.041</sub>	1.011 <sub>0.022</sub>	1.506 <sub>0.026</sub>
	0.25	0.214 <sub>0.016</sub>	2.154 <sub>0.040</sub>	1.022 <sub>0.024</sub>	1.513 <sub>0.025</sub>
	0.30	0.279 <sub>0.017</sub>	2.192 <sub>0.039</sub>	0.970 <sub>0.023</sub>	1.478 <sub>0.024</sub>
	0.35	0.342 <sub>0.016</sub>	2.192 <sub>0.040</sub>	0.970 <sub>0.022</sub>	1.512 <sub>0.024</sub>
	0.40	0.371 <sub>0.016</sub>	2.248 <sub>0.038</sub>	0.984 <sub>0.025</sub>	1.476 <sub>0.025</sub>

## 4.4 Simulation results for diagnostic test methods in logistic mixed models

Table 41: Power of various tests for an outcome dependent visit process when the outcome follows a logistic mixed model with  $m = 200$  subjects and an average sample size of 9. Outcome dependence is on a lagged value of the outcome. Results are presented for the case of all irregular visits (top panel) or a mix of regular and irregular visits (bottom panel). The tests look for dependence of the actual visit times (Cox), intervisit times (IVT) or the number of visits ( $n_i$  or  $n_{ij}^*$ ) on the best predicted values of the random intercept and slope (b0b1) or the random effects portion of the linear predictor (zb).

Informative Visit Process	Power								
	$\delta_Y$	Coxb <sub>0</sub> b <sub>1</sub>	Coxzb	$N_i$ b <sub>0</sub> b <sub>1</sub>	$n_i$ zb	IVTb <sub>0</sub> b <sub>1</sub>	IVTzb	GEEn <sub>i</sub> <sup>*</sup>	GEEn <sub>i</sub>
Irregular visits									
	0.00	0.045	0.039	0.040	0.038	0.047	0.047	0.056	0.058
	0.10	0.198	0.220	0.246	0.290	0.192	0.252	0.278	0.386
	0.20	0.652	0.692	0.752	0.812	0.594	0.668	0.652	0.884
	0.25	0.822	0.838	0.884	0.894	0.754	0.836	0.868	0.966
	0.30	0.906	0.904	0.918	0.922	0.880	0.904	0.952	0.996
	0.35	0.932	0.938	0.940	0.940	0.926	0.934	0.984	1.000
	0.40	0.914	0.914	0.916	0.916	0.914	0.914	0.998	1.000
Mixed visits									
	0.00	0.052	0.052	0.051	0.054	0.064	0.058	0.069	0.073
	0.10	0.220	0.270	0.210	0.240	0.084	0.110	0.152	0.202
	0.20	0.596	0.684	0.530	0.608	0.228	0.256	0.366	0.506
	0.25	0.822	0.856	0.730	0.788	0.340	0.400	0.518	0.734
	0.30	0.860	0.892	0.796	0.846	0.414	0.456	0.594	0.826
	0.35	0.888	0.892	0.876	0.882	0.570	0.620	0.802	0.922
	0.40	0.934	0.936	0.934	0.932	0.682	0.722	0.866	0.984



Table 42: Median values of parameter estimates from a maximum likelihood logistic mixed model fit that ignores outcome dependence when the outcome follows a logistic mixed model with  $m = 200$  subjects and an average sample size of 9. Outcome dependence is on a lagged value of the outcome. Results are presented for the case of all irregular visits (top) or a mix of regular and irregular visits (bottom) and a range of outcome dependence,  $\delta_Y$ .

Informative Visit Process	Simulated mean parameter estimates (SEs as subscripts)				
	$\delta_Y$	$\beta_0$ (true=-2)	$\beta_T$ (true=1)	$\beta_G$ (true=-1)	$\beta_I$ (true=0.5)
Irregular visits					
	0.00	-2.011 <sub>0.005</sub>	1.007 <sub>0.002</sub>	-1.023 <sub>0.008</sub>	0.509 <sub>0.004</sub>
	0.10	-2.002 <sub>0.010</sub>	1.018 <sub>0.004</sub>	-1.063 <sub>0.017</sub>	0.517 <sub>0.008</sub>
	0.20	-1.999 <sub>0.010</sub>	1.022 <sub>0.005</sub>	-1.011 <sub>0.018</sub>	0.504 <sub>0.008</sub>
	0.25	-1.985 <sub>0.010</sub>	1.022 <sub>0.005</sub>	-1.017 <sub>0.018</sub>	0.508 <sub>0.008</sub>
	0.30	-1.970 <sub>0.010</sub>	1.021 <sub>0.005</sub>	-1.063 <sub>0.018</sub>	0.522 <sub>0.009</sub>
	0.35	-1.976 <sub>0.009</sub>	1.023 <sub>0.005</sub>	-1.018 <sub>0.017</sub>	0.516 <sub>0.008</sub>
	0.40	-1.968 <sub>0.010</sub>	1.025 <sub>0.004</sub>	-1.006 <sub>0.019</sub>	0.510 <sub>0.008</sub>
Mixed visits					
	0.00	-2.013 <sub>0.005</sub>	1.005 <sub>0.002</sub>	-1.011 <sub>0.008</sub>	0.509 <sub>0.004</sub>
	0.10	-2.003 <sub>0.009</sub>	1.011 <sub>0.004</sub>	-1.031 <sub>0.016</sub>	0.511 <sub>0.008</sub>
	0.20	-2.005 <sub>0.009</sub>	1.017 <sub>0.004</sub>	-1.036 <sub>0.016</sub>	0.506 <sub>0.007</sub>
	0.25	-1.997 <sub>0.009</sub>	1.012 <sub>0.004</sub>	-1.048 <sub>0.017</sub>	0.517 <sub>0.007</sub>
	0.30	-1.990 <sub>0.009</sub>	1.015 <sub>0.004</sub>	-1.023 <sub>0.017</sub>	0.511 <sub>0.008</sub>
	0.35	-1.984 <sub>0.009</sub>	1.014 <sub>0.004</sub>	-1.036 <sub>0.017</sub>	0.515 <sub>0.008</sub>
	0.40	-1.979 <sub>0.009</sub>	1.013 <sub>0.004</sub>	-1.002 <sub>0.016</sub>	0.505 <sub>0.007</sub>

Table 43: Median values of parameter estimates from a GEE independence logistic model fit that ignores outcome dependence when the outcome follows a logistic mixed model with  $m = 200$  subjects and an average sample size of 9. Outcome dependence is on a lagged value of the outcome. Results are presented for the case of all irregular visits (top) or a mix of regular and irregular visits (bottom) and a range of outcome dependence,  $\delta_Y$ .

Informative Visit Process	Simulated mean parameter estimates (SEs as subscripts)				
	$\delta_Y$	$\beta_0$ (true=-2)	$\beta_T$ (true=1)	$\beta_G$ (true=-1)	$\beta_I$ (true=0.5)
Irregular visits					
	0.00	-1.732 <sub>0.005</sub>	1.062 <sub>0.005</sub>	-0.741 <sub>0.007</sub>	0.348 <sub>0.003</sub>
	0.10	-1.704 <sub>0.010</sub>	1.094 <sub>0.009</sub>	-0.762 <sub>0.014</sub>	0.356 <sub>0.006</sub>
	0.20	-1.702 <sub>0.010</sub>	1.126 <sub>0.009</sub>	-0.769 <sub>0.015</sub>	0.363 <sub>0.006</sub>
	0.25	-1.700 <sub>0.010</sub>	1.161 <sub>0.010</sub>	-0.738 <sub>0.015</sub>	0.346 <sub>0.006</sub>
	0.30	-1.649 <sub>0.010</sub>	1.129 <sub>0.009</sub>	-0.730 <sub>0.014</sub>	0.342 <sub>0.005</sub>
	0.35	-1.645 <sub>0.010</sub>	1.149 <sub>0.010</sub>	-0.739 <sub>0.014</sub>	0.335 <sub>0.006</sub>
	0.40	-1.622 <sub>0.011</sub>	1.157 <sub>0.010</sub>	-0.764 <sub>0.015</sub>	0.348 <sub>0.006</sub>
Mixed visits					
	0.00	-1.731 <sub>0.007</sub>	1.063 <sub>0.007</sub>	-0.754 <sub>0.009</sub>	0.353 <sub>0.004</sub>
	0.10	-1.695 <sub>0.013</sub>	1.094 <sub>0.013</sub>	-0.793 <sub>0.018</sub>	0.364 <sub>0.008</sub>
	0.20	-1.698 <sub>0.014</sub>	1.139 <sub>0.013</sub>	-0.792 <sub>0.019</sub>	0.360 <sub>0.008</sub>
	0.25	-1.691 <sub>0.014</sub>	1.148 <sub>0.014</sub>	-0.765 <sub>0.019</sub>	0.355 <sub>0.008</sub>
	0.30	-1.662 <sub>0.014</sub>	1.145 <sub>0.013</sub>	-0.739 <sub>0.018</sub>	0.339 <sub>0.008</sub>
	0.35	-1.655 <sub>0.014</sub>	1.175 <sub>0.013</sub>	-0.776 <sub>0.019</sub>	0.352 <sub>0.008</sub>
	0.40	-1.628 <sub>0.014</sub>	1.166 <sub>0.014</sub>	-0.752 <sub>0.018</sub>	0.350 <sub>0.007</sub>

Table 44: Power of various tests for an outcome dependent visit process when the outcome follows a logistic mixed model with  $m = 200$  subjects and an average sample size of 9. Outcome dependence is on the random intercept. Results are presented for the case of all irregular visits (top panel) or a mix of regular and irregular visits (bottom panel). The tests look for dependence of the actual visit times (Cox), intervisit times (IVT) or the number of visits ( $n_i$  or  $n_{ij}^*$ ) on the best predicted values of the random intercept and slope (b0b1) or the random effects portion of the linear predictor (zb).

Informative Visit Process	Power								
	$\delta_Y$	Coxb <sub>0</sub> b <sub>1</sub>	Coxzb	N <sub>i</sub> b <sub>0</sub> b <sub>1</sub>	n <sub>i</sub> zb	IVTb <sub>0</sub> b <sub>1</sub>	IVTzb	GEE n <sub>i</sub> <sup>*</sup>	GEE n <sub>i</sub>
Irregular visits									
	0.00	0.048	0.045	0.045	0.044	0.049	0.045	0.054	0.060
	0.10	0.564	0.526	0.678	0.710	0.480	0.506	0.628	0.788
	0.20	0.904	0.906	0.920	0.920	0.896	0.906	0.974	0.996
	0.25	0.930	0.928	0.932	0.932	0.932	0.930	1.000	1.000
	0.30	0.906	0.910	0.910	0.910	0.908	0.910	1.000	1.000
	0.35	0.932	0.934	0.934	0.934	0.934	0.934	1.000	1.000
	0.40	0.898	0.898	0.898	0.898	0.898	0.898	1.000	1.000
Mixed visits									
	0.00	0.053	0.046	0.047	0.041	0.062	0.057	0.054	0.060
	0.10	0.456	0.476	0.408	0.448	0.180	0.206	0.324	0.460
	0.20	0.918	0.908	0.894	0.916	0.580	0.570	0.790	0.898
	0.25	0.934	0.936	0.930	0.934	0.770	0.756	0.912	0.976
	0.30	0.946	0.948	0.948	0.948	0.856	0.846	0.978	0.998
	0.35	0.922	0.924	0.924	0.924	0.898	0.886	0.992	1.000
	0.40	0.932	0.934	0.934	0.934	0.926	0.932	0.998	1.000

Table 45: Median values of parameter estimates from a maximum likelihood logistic mixed model fit that ignores outcome dependence when the outcome follows a logistic mixed model with  $m = 200$  subjects and an average sample size of 9. Outcome dependence is on the random intercept of the outcome. Results are presented for the case of all irregular visits (top) or a mix of regular and irregular visits (bottom) and a range of outcome dependence,  $\delta_Y$ .

Informative Visit Process	Simulated mean parameter estimates (SEs as subscripts)				
	$\delta_Y$	$\beta_0$ (true=-2)	$\beta_T$ (true=1)	$\beta_G$ (true=-1)	$\beta_I$ (true=0.5)
Irregular visits					
	0.00	-2.002 <sub>0.005</sub>	1.005 <sub>0.002</sub>	-1.027 <sub>0.008</sub>	0.508 <sub>0.004</sub>
	0.10	-1.978 <sub>0.010</sub>	1.007 <sub>0.004</sub>	-1.010 <sub>0.016</sub>	0.505 <sub>0.008</sub>
	0.20	-1.930 <sub>0.009</sub>	1.003 <sub>0.005</sub>	-1.014 <sub>0.017</sub>	0.515 <sub>0.008</sub>
	0.25	-1.908 <sub>0.009</sub>	1.003 <sub>0.004</sub>	-1.008 <sub>0.016</sub>	0.501 <sub>0.008</sub>
	0.30	-1.878 <sub>0.009</sub>	0.995 <sub>0.004</sub>	-1.023 <sub>0.016</sub>	0.514 <sub>0.008</sub>
	0.35	-1.852 <sub>0.009</sub>	1.006 <sub>0.005</sub>	-1.026 <sub>0.015</sub>	0.513 <sub>0.008</sub>
	0.40	-1.849 <sub>0.009</sub>	1.004 <sub>0.005</sub>	-1.017 <sub>0.016</sub>	0.510 <sub>0.008</sub>
Mixed visits					
	0.00	-2.021 <sub>0.004</sub>	1.006 <sub>0.002</sub>	-1.007 <sub>0.008</sub>	0.504 <sub>0.004</sub>
	0.10	-2.013 <sub>0.009</sub>	1.004 <sub>0.004</sub>	-0.986 <sub>0.015</sub>	0.500 <sub>0.007</sub>
	0.20	-1.978 <sub>0.009</sub>	1.004 <sub>0.004</sub>	-0.992 <sub>0.015</sub>	0.498 <sub>0.007</sub>
	0.25	-1.945 <sub>0.009</sub>	1.002 <sub>0.004</sub>	-1.030 <sub>0.016</sub>	0.514 <sub>0.007</sub>
	0.30	-1.960 <sub>0.009</sub>	1.005 <sub>0.004</sub>	-0.994 <sub>0.016</sub>	0.497 <sub>0.008</sub>
	0.35	-1.929 <sub>0.009</sub>	0.999 <sub>0.004</sub>	-1.017 <sub>0.015</sub>	0.513 <sub>0.008</sub>
	0.40	-1.926 <sub>0.009</sub>	1.001 <sub>0.004</sub>	-1.021 <sub>0.016</sub>	0.518 <sub>0.008</sub>

Table 46: Median values of parameter estimates from a GEE independence logistic model fit that ignores outcome dependence when the outcome follows a logistic mixed model with  $m = 200$  subjects and an average sample size of 9. Outcome dependence is on the random intercept of the outcome. Results are presented for the case of all irregular visits (top) or a mix of regular and irregular visits (bottom) and a range of outcome dependence,  $\delta_Y$ .

Informative Visit Process	Simulated mean parameter estimates (SEs as subscripts)				
	$\delta_Y$	$\beta_0$ (true=-2)	$\beta_T$ (true=1)	$\beta_G$ (true=-1)	$\beta_I$ (true=0.5)
Irregular visits					
	0.00	-1.731 <sub>0.005</sub>	1.069 <sub>0.005</sub>	-0.753 <sub>0.007</sub>	0.348 <sub>0.003</sub>
	0.10	-1.627 <sub>0.010</sub>	1.057 <sub>0.009</sub>	-0.762 <sub>0.014</sub>	0.357 <sub>0.006</sub>
	0.20	-1.544 <sub>0.009</sub>	1.077 <sub>0.009</sub>	-0.771 <sub>0.014</sub>	0.357 <sub>0.006</sub>
	0.25	-1.516 <sub>0.009</sub>	1.076 <sub>0.009</sub>	-0.758 <sub>0.014</sub>	0.354 <sub>0.006</sub>
	0.30	-1.474 <sub>0.009</sub>	1.072 <sub>0.009</sub>	-0.749 <sub>0.013</sub>	0.347 <sub>0.005</sub>
	0.35	-1.428 <sub>0.009</sub>	1.085 <sub>0.009</sub>	-0.778 <sub>0.014</sub>	0.360 <sub>0.006</sub>
	0.40	-1.418 <sub>0.010</sub>	1.091 <sub>0.009</sub>	-0.741 <sub>0.014</sub>	0.357 <sub>0.006</sub>
Mixed visits					
	0.00	-1.745 <sub>0.007</sub>	1.073 <sub>0.007</sub>	-0.768 <sub>0.009</sub>	0.360 <sub>0.004</sub>
	0.10	-1.641 <sub>0.013</sub>	1.057 <sub>0.013</sub>	-0.758 <sub>0.019</sub>	0.361 <sub>0.008</sub>
	0.20	-1.566 <sub>0.012</sub>	1.082 <sub>0.013</sub>	-0.758 <sub>0.017</sub>	0.358 <sub>0.008</sub>
	0.25	-1.517 <sub>0.013</sub>	1.082 <sub>0.013</sub>	-0.769 <sub>0.019</sub>	0.357 <sub>0.008</sub>
	0.30	-1.471 <sub>0.012</sub>	1.089 <sub>0.013</sub>	-0.804 <sub>0.018</sub>	0.370 <sub>0.008</sub>
	0.35	-1.438 <sub>0.013</sub>	1.100 <sub>0.013</sub>	-0.759 <sub>0.019</sub>	0.357 <sub>0.008</sub>
	0.40	-1.389 <sub>0.012</sub>	1.092 <sub>0.011</sub>	-0.784 <sub>0.017</sub>	0.371 <sub>0.007</sub>

## References

- [1] McCulloch Charles E., Neuhaus John M., Olin Rebecca L.. Biased and unbiased estimation in longitudinal studies with informative visit processes *Biometrics*. 2016:1315–1324.