

#### Blinded Sample Size Recalculation in Longitudinal Clinical Trials Using Generalized Estimating Equations

Daniel Wachtlin 2013/10/29

Virtual Journal Club DIA Stat Community & PSI



### Background

- Uncertainty related to initial parameter estimates in the planning stage of clinical trials
- Increased complexity of sample size calculation in longitudinal clinical trials (intra subject correlation)
- <u>One</u> analysis approach: Generalized Estimating Equations (GEE)
- Chance to correct for initial misspecifications would be helpful
- <u>One</u> Solution: Blinded sample size recalculation with Internal Pilot Study (IPS) Design
- Primary question: Is type I/II error preserved within recalculation procedure?



#### **Recalculation procedure (generic)**





# Simulation study (1)

Investigated setting

• Model: 
$$y_{ij} = \beta_1 + \beta_2 r_i + \beta_3 j + \beta_4 r_i j + \varepsilon_{ij}$$

- y<sub>ii</sub> Continuous outcome for subject *i* at time *j*
- β Model coefficients
- r<sub>i</sub> Treatment indicator of subject *i*, balanced designs investigated
- j Measurement time j of a subject,  $j=\{1,2,3,4,5,6\}$
- ε<sub>ii</sub> Model error term related to measurement *j* of subject *i*
- Two randomized treatments
- Parallel group design
- Balanced treatment allocation
- Parameter of interest:  $\beta_4$
- Generalized estimating equations: 
   *E*<sub>ii</sub> correlated within subjects

### Simulation study (2)

- Sample size calculation
  - Formula of Jung and Ahn (2003)

$$n = \frac{\sigma^2 s_t^2 (z_{1-\alpha/2} + z_{1-\gamma})^2}{\beta_{40}^2 \mu_0^2 \sigma_r^2 \sigma_t^4}$$

$$\sigma^{2} = Var(\varepsilon_{ij}).$$

$$s_{t}^{2} = \sum_{j=1}^{K} \sum_{j'=1}^{K} p_{jj'} \rho_{jj'}(t_{j} - \mu_{1})(t_{j'} - \mu_{1}).$$

$$\mu_{k} = \mu_{0}^{-1} \sum_{j=1}^{K} p_{j} t_{j}^{k}, k = 1, 2.$$

$$\mu_{0} = \sum_{j=1}^{K} p_{j} - \sigma_{t}^{2} = \mu_{2} - \mu_{1}^{2}.$$

- Constant risk of dropout per period, only permanent dropouts
- Correlation according to the damped exponential family of correlation structures (Munoz (1992)):  $Corr(y_{ij}, y_{ij+t}) = \rho^{t^{\theta}}, 0 \le \rho, \theta \le 1$
- Significance level: 5%, Power: 80%

# Simulation study (3)

- Reestimation of selected parameters
  - Size of internal pilot study:  $n_{IPS} = \pi \cdot n, 0 \le \pi \le 1$
  - Reestimating covariance matrix
    - Simultaneously estimate  $\rho$ ,  $\theta$ ,  $\sigma^2$
    - Damped exponential family of correlation structures (Munoz (1992))
    - Variability of error term identical for treatment groups, subjects and measurement occasions,  $Var(\varepsilon_{ij}) = Var(\varepsilon)$
    - Non-linear model
    - Starting values  $\rho$ ,  $\theta$ ,  $\sigma^2 = 1$
    - Minimize unweighted sum of deviations between model based and empirical covariance matrix
  - Reestimating constant risk of dropout
    - Derive from rate of observed values at last measurement time
    - Assume identical risk of dropout for both treatment groups
    - $\hat{h}_{IPS} = 1 \sqrt[5]{\hat{p}_6}$

# Simulation study (4)

- Sample size recalculation
  - Use updated parameter estimates
  - Blinded procedure
  - Unrestricted design (Birkett/Day (1994))
  - $n_{total} = MAX(\pi \cdot n_{ini}, n_{recalc})$
- Analysis
  - (Weighted) GEE (Robins/Rotnitzky (1995))
  - Inverse probability weighting, weights only differ between measurement times
  - Working correlation matrix = Independent (Mancl/Leroux (1996), Ziegler/Vens (2010))
  - Test  $\beta_4 = 0$ , one-sided significance level 2.5%
- Simulations based on 10.000 samples under H<sub>0</sub> and H<sub>a</sub>

### **Scenarios**

Scen.	ρ	θ	$\sigma^2$	h	π	n	n <sub>IPS</sub>
1	0.25	0.5	1	0.00	0.5	203	102
2	0.25	0.5	1	0.00	0.2	203	41
3	0.25	0.5	1	0.00	0.8	203	162
4	0.7	0.1	1	0.00	0.5	67	34
5	0.7	1	1	0.00	0.5	223	112
6	0.25	0.5	5	0.00	0.5	1015	508
7	0.1	0.1	1	0.00	0.5	172	86
8	0.1	0.9	1	0.00	0.5	197	98
9	0.25	0.5	1	0.01	0.2	208	42
10	0.25	0.5	1	0.05	0.2	231	46
11	0.25	0.5	1	0.01	0.8	208	166
12	0.25	0.5	1	0.05	0.8	231	185

## Results (1)



## Results (2)



#### Illustration of results of $\theta$ estimates

• 
$$\rho_{ij} = \begin{pmatrix}
1 & 0.5 & 0.5 & 0.5 \\
0.5 & 1 & 0.5 & 0.5 \\
0.5 & 0.5 & 1 & 0.5 \\
0.5 & 0.5 & 0.5 & 1
\end{pmatrix}, \rho = 0.5, \theta = 0$$

• 
$$\rho_{ij} = \begin{pmatrix}
1 & 0.5 & 0.375 & 0.301 \\
0.5 & 1 & 0.5 & 0.375 \\
0.375 & 0.5 & 1 & 0.5 \\
0.301 & 0.375 & 0.5 & 1
\end{pmatrix}, \rho = 0.5, \theta = 0.5$$

• 
$$\rho_{ij} = \begin{pmatrix}
1 & 0.5 & 0.25 & 0.125 \\
0.5 & 1 & 0.5 & 0.25 \\
0.25 & 0.5 & 1 & 0.5 \\
0.125 & 0.25 & 0.5 & 1
\end{pmatrix}, \rho = 0.5, \theta = 1$$

11

### **Results (3)**

![](_page_11_Figure_1.jpeg)

### **Results (4)**

![](_page_12_Figure_1.jpeg)

# **Results (5)**

Scenario	n	n <sub>IPS</sub>	$\alpha_{FIXED}$	(1-γ) <sub>FIXED</sub>	$\alpha_{IPS}$	(1-γ) <sub>IPS</sub>
1	203	102	2.37	80.34	2.70	80.61
2	203	41	2.87	80.42	2.56	79.20
3	203	162	2.51	80.20	2.49	80.69
4	67	34	2.73	79.93	3.19	82.66
5	223	112	2.67	79.90	2.78	79.90
6	1015	508	2.54	80.36	2.38	79.88
7	172	86	2.54	80.68	2.53	81.23
8	197	98	2.55	80.15	2.54	79.49
9	208	42	2.86	79.88	2.64	79.10
10	231	46	2.57	80.09	2.34	79.37
11	208	166	2.60	80.50	2.66	81.03
12	231	185	2.86	80.01	2.75	80.52

# Summary

- Mostly unbiased estimation of  $\rho$ ,  $\sigma^2$
- Estimation of  $\theta$  associated with high variability and risk of bias
- Sample size on average near (slightly above) the fixed sample size results
- Results confirmed in the presence of missing data
- Impact of IPS size on estimates and resulting sample size distribution
- Type I error mostly very near to nominal value
- Robust power results
- Few limitations
  - Starting values for reestimating covariance parameters
  - Bound effects / biased estimates can be anticipated by simulating extreme scenarios
  - Simplified assumptions for investigated setting