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)
- **One analysis approach:** Generalized Estimating Equations (GEE)
- Chance to correct for initial misspecifications would be helpful
- **One 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)

1. Sample size calculation
2. Recruitment of subjects needed for IPS
3. Reestimation of selected parameters
4. Sample size recalculation
5. Recruitment of further subjects (if needed)
6. Data analysis

Recruitment of subjects
Simulation study (1)

- Investigated setting
  - Model: \[ y_{ij} = \beta_1 + \beta_2 r_i + \beta_3 j + \beta_4 r_i j + \varepsilon_{ij} \]
    - \( y_{ij} \): Continuous outcome for subject \( i \) at time \( j \)
    - \( \beta \): Model coefficients
    - \( r_i \): Treatment indicator of subject \( i \), balanced designs investigated
    - \( j \): Measurement time \( j \) of a subject, \( j = \{1, 2, 3, 4, 5, 6\} \)
    - \( \varepsilon_{ij} \): 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: \( \varepsilon_{ij} \) correlated within subjects
Simulation study (2)

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

\[ n = \frac{\sigma^2 s_t^2 (z_{1-\alpha/2} + z_{1-\gamma})^2}{\beta^2_0 \mu_0^2 \sigma_r^2 \sigma_t^4} \]

\[ \sigma^2 = \text{Var}(\varepsilon_{ij}). \quad s_t^2 = \sum_{j=1}^{K} \sum_{j'=1}^{K} p_{jj'} \rho_{jj'} (t_j - \mu_1) (t_{j'} - \mu_1). \]

\[ \mu_k = \mu_0^{-1} \sum_{j=1}^{K} p_j t_j^k, \quad k = 1, 2. \quad \mu_0 = \sum_{j=1}^{K} p_j. \quad \sigma_t^2 = \mu_2 - \mu_1^2. \]

- Constant risk of dropout per period, only permanent dropouts
- Correlation according to the damped exponential family of correlation structures (Munoz (1992)): \( \text{Corr}(y_{ij}, y_{ij+t}) = \rho^t \theta, \quad 0 \leq \rho, \theta \leq 1 \)
- Significance level: 5%, Power: 80%
Simulation study (3)

- Reestimation of selected parameters
  - Size of internal pilot study: \( n_{IPS} = \pi \cdot n, 0 \leq \pi \leq 1 \)

- 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, \( Var(e_{ij}) = Var(e) \)
  - 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))
  - $n_{total} = \text{MAX}(\pi \cdot n_{ini}, n_{recalc})$

- 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 $H_0$ and $H_a$
## Scenarios

<table>
<thead>
<tr>
<th>Scen.</th>
<th>$\rho$</th>
<th>$\theta$</th>
<th>$\sigma^2$</th>
<th>$h$</th>
<th>$\pi$</th>
<th>$n$</th>
<th>$n_{\text{IPS}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.25</td>
<td>0.5</td>
<td>1</td>
<td>0.00</td>
<td>0.5</td>
<td>203</td>
<td>102</td>
</tr>
<tr>
<td>2</td>
<td>0.25</td>
<td>0.5</td>
<td>1</td>
<td>0.00</td>
<td>0.2</td>
<td>203</td>
<td>41</td>
</tr>
<tr>
<td>3</td>
<td>0.25</td>
<td>0.5</td>
<td>1</td>
<td>0.00</td>
<td>0.8</td>
<td>203</td>
<td>162</td>
</tr>
<tr>
<td>4</td>
<td>0.7</td>
<td>0.1</td>
<td>1</td>
<td>0.00</td>
<td>0.5</td>
<td>67</td>
<td>34</td>
</tr>
<tr>
<td>5</td>
<td>0.7</td>
<td>1</td>
<td>1</td>
<td>0.00</td>
<td>0.5</td>
<td>223</td>
<td>112</td>
</tr>
<tr>
<td>6</td>
<td>0.25</td>
<td>0.5</td>
<td>5</td>
<td>0.00</td>
<td>0.5</td>
<td>1015</td>
<td>508</td>
</tr>
<tr>
<td>7</td>
<td>0.1</td>
<td>0.1</td>
<td>1</td>
<td>0.00</td>
<td>0.5</td>
<td>172</td>
<td>86</td>
</tr>
<tr>
<td>8</td>
<td>0.1</td>
<td>0.9</td>
<td>1</td>
<td>0.00</td>
<td>0.5</td>
<td>197</td>
<td>98</td>
</tr>
<tr>
<td>9</td>
<td>0.25</td>
<td>0.5</td>
<td>1</td>
<td>0.01</td>
<td>0.2</td>
<td>208</td>
<td>42</td>
</tr>
<tr>
<td>10</td>
<td>0.25</td>
<td>0.5</td>
<td>1</td>
<td>0.05</td>
<td>0.2</td>
<td>231</td>
<td>46</td>
</tr>
<tr>
<td>11</td>
<td>0.25</td>
<td>0.5</td>
<td>1</td>
<td>0.01</td>
<td>0.8</td>
<td>208</td>
<td>166</td>
</tr>
<tr>
<td>12</td>
<td>0.25</td>
<td>0.5</td>
<td>1</td>
<td>0.05</td>
<td>0.8</td>
<td>231</td>
<td>185</td>
</tr>
</tbody>
</table>
Results (1)
Results (2)
Illustration of results of $\theta$ estimates

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

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

\[ \rho_{ij} = \begin{pmatrix} 1 & 0.5 & 0.25 & 0.125 \\ 0.5 & 1 & 0.5 & 0.25 \\ 0.25 & 0.5 & 1 & 0.5 \\ 0.125 & 0.25 & 0.5 & 1 \end{pmatrix} , \rho = 0.5, \theta = 1 \]
Results (3)
Results (4)
## Results (5)

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