Meta-Analysis of Rare Events

Dankmar Böhning and Antonello Maruotti

Southampton Statistical Sciences Research Institute
University of Southampton

5th of September 2014
Contents

What are rare events data?

What are the problems with rare events trials?

Effect estimation using Mantel-Haenszel

Poisson with fixed and random effects

Sensitivity analysis: the effect of excluding zero-studies

Zero-inflation models

Logistic regression modelling

Conditional logistic regression modelling
Contents

What are rare events data?

What are the problems with rare events trials?

Effect estimation using Mantel-Haenszel

Poisson with fixed and random effects

Sensitivity analysis: the effect of excluding zero-studies

Zero-inflation models

Logistic regression modelling

Conditional logistic regression modelling
What are rare events data?

motivation

- recent debate on the safety of the diabetes drug *rosiglitazone*
- meta-analysis (MA) by Nissen and Wolski (2007, 2010)
- Böhning, Mylona, Kimber (2014) focus on existing methodology to adapt to MA of rare event trials
Table: Study data of meta–analysis on rare events in Rosiglitazone and control arm; MI refers to the myocardial infarction deaths, CV to cardiovascular deaths, $n$ is the size of the respective study arm and ‘duration’ refers to the study period at risk (in weeks)

<table>
<thead>
<tr>
<th>ID</th>
<th>study label</th>
<th>treatment arm</th>
<th>control arm</th>
<th>treatment arm</th>
<th>control arm</th>
<th>treatment arm</th>
<th>control arm</th>
<th>treatment arm</th>
<th>control arm</th>
<th>treatment arm</th>
<th>control arm</th>
<th>treatment arm</th>
<th>control arm</th>
<th>treatment arm</th>
<th>control arm</th>
<th>treatment arm</th>
<th>control arm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$n$</td>
<td>MI</td>
<td>CV</td>
<td>$n$</td>
<td>MI</td>
<td>CV</td>
<td>duration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>49653/011</td>
<td>357</td>
<td>2</td>
<td>1</td>
<td>176</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>49653/020</td>
<td>391</td>
<td>2</td>
<td>0</td>
<td>207</td>
<td>1</td>
<td>0</td>
<td>52</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>49653/024</td>
<td>774</td>
<td>1</td>
<td>0</td>
<td>185</td>
<td>1</td>
<td>0</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>49653/093</td>
<td>213</td>
<td>0</td>
<td>0</td>
<td>109</td>
<td>1</td>
<td>0</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>49653/452</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td>0</td>
<td>0</td>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>DREAM</td>
<td>2635</td>
<td>15</td>
<td>12</td>
<td>2634</td>
<td>9</td>
<td>10</td>
<td>156</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>ADOPT19</td>
<td>1456</td>
<td>27</td>
<td>2</td>
<td>2895</td>
<td>41</td>
<td>5</td>
<td>208</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>RECORD</td>
<td>2220</td>
<td>64</td>
<td>60</td>
<td>2227</td>
<td>56</td>
<td>71</td>
<td>260</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
a second example

- Niël-Weise et al. (2007) did a MA on the effect of anti-infective-treated central venous catheters on catheter-related bloodstream infection (CRBSI) in the acute care setting.
- meta-analysis involved 18 clinical trials.
- control group is standard catheter.
Table: Meta-analysis on rare evidence data on the effect of anti-infective-treated catheter in comparison to standard catheter; CRBSI refers to catheter-related bloodstream infection events, $n$ is the size of the respective study arm

<table>
<thead>
<tr>
<th>study ID</th>
<th>control arm</th>
<th>treatment arm</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CRBSI</td>
<td>$n$</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>117</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>35</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>195</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>136</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>157</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>139</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>177</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>39</td>
</tr>
</tbody>
</table>
## What are rare events data?

<table>
<thead>
<tr>
<th>study ID</th>
<th>control arm</th>
<th></th>
<th>treatment arm</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CRBSI</td>
<td>n</td>
<td>CRBSI</td>
<td>n</td>
</tr>
<tr>
<td>9</td>
<td>19</td>
<td>103</td>
<td>1</td>
<td>97</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>122</td>
<td>1</td>
<td>113</td>
</tr>
<tr>
<td>11</td>
<td>7</td>
<td>64</td>
<td>0</td>
<td>66</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>58</td>
<td>0</td>
<td>70</td>
</tr>
<tr>
<td>13</td>
<td>5</td>
<td>175</td>
<td>3</td>
<td>188</td>
</tr>
<tr>
<td>14</td>
<td>11</td>
<td>180</td>
<td>6</td>
<td>187</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>105</td>
<td>0</td>
<td>118</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>262</td>
<td>0</td>
<td>252</td>
</tr>
<tr>
<td>17</td>
<td>3</td>
<td>362</td>
<td>1</td>
<td>345</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>69</td>
<td>4</td>
<td>64</td>
</tr>
</tbody>
</table>
a definition

MA of rare events trials deals with MA of trials which includes single-zero or double-zero trials.

A single-zero trial is a trial in which at least one arm are has no events. A double-zero trial is a trial in which both arms have no events.
Meta-Analysis of Rare Events

What are the problems with rare events trials?

Contents

What are rare events data?

What are the problems with rare events trials?

Effect estimation using Mantel-Haenszel

Poisson with fixed and random effects

Sensitivity analysis: the effect of excluding zero-studies

Zero-inflation models

Logistic regression modelling

Conditional logistic regression modelling
popular effect measures

- **risk difference** $RD$: risk in treatment arm – risk in control arm estimated by ($x$ number of events and $P$ is person-time)

  $$\hat{RD} = \frac{x_T}{P_T} - \frac{x_C}{P_C}$$

- **risk ratio** $RR$: risk in treatment arm / risk in control arm estimated by

  $$\hat{RR} = \frac{\frac{x_T}{P_T}}{\frac{x_C}{P_C}}$$

- **odds ratio** $OR$: odds in treatment arm / odds in control arm estimated by

  $$\hat{OR} = \frac{\frac{x_T}{(P_T - x_T)}}{\frac{x_C}{(P_C - x_C)}} = \frac{x_T (P_C - x_C)}{x_C (P_T - x_T)}$$
problems can occur on two levels with zero-studies

effect measure itself

▶ no problem for the risk difference

\[
\hat{RD} = \frac{x_T}{P_T} - \frac{x_C}{P_C}
\]

▶ risk ratio and odds ratio: it might be useless (0), infinite (∞), or undefined (0/0)

\[
\hat{RR} = \frac{x_T}{x_C} \frac{P_T}{P_C} \quad \text{and} \quad \hat{OR} = \frac{x_T}{(P_T - x_T)} \frac{x_C}{(P_C - x_C)}
\]
What are the problems with rare events trials?

Problems can occur on two levels with zero-studies:

Uncertainty assessment

- Risk difference
  \[ \text{var}(\hat{RD}) \approx \frac{x_T}{PT^2} + \frac{x_C}{PC^2} \]

- Risk ratio
  \[ \text{var}(\hat{log\,RR}) \approx \frac{1}{x_T} + \frac{1}{x_C} \]

- Odds ratio
  \[ \text{var}(\hat{log\,OR}) \approx \frac{1}{x_T} + \frac{1}{PT - x_T} + \frac{1}{x_C} + \frac{1}{PC - x_C} \]
problems can occur on two levels with zero-studies

disturbances with weighted average computation:

\[
\log RR = \frac{\sum_i w_i \log RR_i}{\sum_i w_i}
\]

where

\[
w_i = \frac{1}{\text{var}(\log RR)}
\]

in a similar way for RD and OR
Contents

What are rare events data?

What are the problems with rare events trials?

**Effect estimation using Mantel-Haenszel**

Poisson with fixed and random effects

Sensitivity analysis: the effect of excluding zero-studies

Zero-inflation models

Logistic regression modelling

Conditional logistic regression modelling
Strategies to cope with zero-studies

Pooling all studies:

\[ \hat{RR}_{\text{crude}} = \frac{\left( \sum_i x_i^T \right) / \left( \sum_i P_i^T \right)}{\left( \sum_i x_i^C \right) / \left( \sum_i P_i^C \right)} \]

Disadvantage: potentially strong confounding effect by ignoring study factor.
Mantel-Haenszel

\[ \hat{RR}_{MH} = \frac{\sum_i x_i^T P_i^C / P_i}{\sum_i x_i^C P_i^T / P_i}, \]

where \( P_i = P_i^C + P_i^T \)

- **advantage:** estimator is not sensitive to zero-studies
- **is also a** weighted estimator

\[ \frac{\sum_i w_i \hat{RR}_i}{\sum_i w_i} \]

using the weights \( w_i = x_i^C P_i^T / P_i \)
**Table:** Mantel-Haenszel estimate in the rare events meta–analysis of Rosiglitazone

<table>
<thead>
<tr>
<th>method</th>
<th>estimate</th>
<th>confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>MI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>crude</td>
<td>1.2561</td>
<td>0.9928 – 1.5911</td>
</tr>
<tr>
<td>MH</td>
<td>1.2782</td>
<td>1.0125 – 1.6137</td>
</tr>
<tr>
<td>CV</td>
<td></td>
<td></td>
</tr>
<tr>
<td>crude</td>
<td>1.1281</td>
<td>0.8496 – 1.4987</td>
</tr>
<tr>
<td>MH</td>
<td>1.0257</td>
<td>0.7760 – 1.3557</td>
</tr>
</tbody>
</table>
Meta-Analysis of Rare Events

Effect estimation using Mantel-Haenszel
Meta-Analysis of Rare Events

Effect estimation using Mantel-Haenszel
Meta-Analysis of Rare Events

Effect estimation using Mantel-Haenszel
Meta-Analysis of Rare Events

Effect estimation using Mantel-Haenszel

**Mantel-Haenszel for OR**

\[
\hat{OR}_{MH} = \frac{\sum_i x_i^T (P_i^C - x_i^C)/P_i}{\sum_i x_i^C (P_i^T - x_i^T)/P_i},
\]

where \( P_i = P_i^C + P_i^T \)

is also a **weighted estimator**

\[
\frac{\sum_i w_i \hat{OR}_i}{\sum_i w_i}
\]

using the weights \( w_i = x_i^C (P_i^T - x_i^T)/P_i \)
Mantel-Haenszel for \( RD \)

\[
\hat{RD}_{\text{MH}} = \frac{\sum_i (x_i^T P_i^C - x_i^C P_i^T)/P_i}{\sum_i (P_i^T P_i^C / P_i)},
\]

where \( P_i = P_i^C + P_i^T \)

\( \hat{RD}_{\text{MH}} \) is also a weighted estimator

\[
\sum_i w_i \hat{RD}_i / \sum_i w_i
\]

using the weights \( w_i = (P_i^T P_i^C)/P_i \)
testing homogeneity of effect

major difficulties with Mantel-Haenszel lies in establishing homogeneity of effect

\[ \chi^2_{k-1} = \sum_i \left( \hat{\log RR}_i - \hat{\log RR}_{MH} \right)^2 \frac{\text{var}(\hat{\log RR}_i)}{\text{var}(\hat{\log RR}_{MH})} \]

where \( k \) is the number of studies

- this statistic **will not work** in the case of zero-studies
- this question needs to be approached in a modelling framework
Contents

What are rare events data?

What are the problems with rare events trials?

Effect estimation using Mantel-Haenszel

**Poisson with fixed and random effects**

Sensitivity analysis: the effect of excluding zero-studies

Zero-inflation models

Logistic regression modelling

Conditional logistic regression modelling
Poisson regression

- consider number of events $X$ as a Poisson count with mean
  \[ E(X) = \mu P \]

- clearly, $\mu = E(X)/P$ is the \textbf{incidence risk}

- write in study $i$
  \[ E(X_{ij}) = \mu_j P_{ij} \]

  for $j = 1$ (treatment) and $j = 0$ (control)

- so that again $RR = \mu_1/\mu_0$
Poisson regression

- in study $i$
  \[ E(X_{ij}) = \mu_j P_{ij} \]

- take logarithms on both sides
  \[ \log E(X_{ij}) = \log P_{ij} + \log \mu_j = \log P_{ij} + \alpha + \beta \times j \]

- so that $\beta = \log(\mu_1/\mu_0)$ is the log-risk ratio

- $\log P_{ij}$ enters as a covariate with known coefficient into the model: an offset
Meta-Analysis of Rare Events

- Poisson with fixed and random effects
Poisson regression with random study effect

taking into account the study effect:

- the effect **homogeneity model**

\[
\log E(X_{ij}) = \log P_{ij} + \alpha_i + \beta \times j
\]

- the effect **heterogeneity model**

\[
\log E(X_{ij}) = \log P_{ij} + \alpha_i + \beta_i \times j
\]
Poisson regression with random study effect

two options:

▶ **fixed effects model**: \(\alpha_i\) and \(\beta_i\) are treated as fixed parameters
▶ disadvantage: many studies \(\rightarrow\) many parameters
▶ **Neyman-Scott problem** (sample size and number of parameters connected)
▶ **random effects model**: \(\alpha_i\) and \(\beta_i\) are treated as random quantities:

\[
\alpha_i \sim N(\alpha, \sigma^2_\alpha) \quad \text{and} \quad \beta_i \sim N(\beta, \sigma^2_\beta)
\]
Poisson regression with random study effect

this leads to the following different likelihoods (in the example of the homogeneity model)

- **fixed effects model:**
  \[
  \prod_i [\text{Po}(x_{i0} | P_{i0} \exp(\alpha_i)) \times \text{Po}(x_{i1} | P_{i1} \exp(\alpha_i + \beta))], \quad (1)
  \]

- **random effects model:**
  \[
  \prod_i \int [\text{Po}(x_{i0} | P_{i0} \exp(\alpha_i))
  
  \times \text{Po}(x_{i1} | P_{i1} \exp(\alpha_i + \beta))] \phi(\alpha_i | \alpha, \sigma^2_\alpha) d\alpha_i.
  \]
Poisson regression with random study effect

likelihood in the example of the **heterogeneity model**

- random effects model:

\[
\prod_i \int Po(x_{i0} | P_{i0} \exp(\alpha_i)) \times \\
\left[ \int Po(x_{i1} | P_{i1} \exp(\alpha_i + \beta_i)) \phi(\beta_i | 0, \sigma_\beta^2) d\beta_i \right] \phi(\alpha_i | \alpha, \sigma_\alpha^2) d\alpha_i.
\]
Poisson regression with random study effect

integrals have no closed form solution:

- Laplace approximation
- Gauss-Hermite quadrature
Meta-Analysis of Rare Events

- Poisson with fixed and random effects
Testing homogeneity with the likelihood ratio test

- **random effects model** $M_1$:

$$L_1 = \prod_i \int Po(x_{i0}|P_{i0}\exp(\alpha_i))Po(x_{i1}|P_{i1}\exp(\alpha_i+\beta))\phi(\alpha_i|0, \sigma^2_\alpha)d\alpha_i$$

- **NO random effects** $M_0$:

$$L_0 = \prod_i [Po(x_{i0}|P_{i0}\exp(\alpha)) \times Po(x_{i1}|P_{i1}\exp(\alpha + \beta))]$$

**likelihood ratio**

$$\log \lambda = 2 \log L_1/L_0$$

is $\chi^2$ with 1 df under the $M_0$
Testing homogeneity with the likelihood ratio test

- variance estimates cannot be negative:

\[ \hat{\sigma}^2_\alpha \geq 0 \]

- hence: distribution of \( \hat{\sigma}^2_\alpha \) cannot be normal
Meta-Analysis of Rare Events

- Poisson with fixed and random effects

![Graph showing normal density and variance estimator of random effect.](image-url)
Testing homogeneity with the likelihood ratio test

- asymptotic distribution:

\[ P\left( \frac{\hat{\sigma}^2}{\text{s.e.}(\hat{\sigma}^2)} < x \right) = 0.5 + 0.5\Phi(x) \]

where \( \Phi(x) \) is the CDF of a standard normal distribution

- similarly for the asymptotic distribution of the likelihood ratio

\[ \log \lambda = 2 \log \frac{L_1}{L_0} \sim 0.5 + 0.5\chi^2_1 \]

- in practice, conventionally computed P-values need only be divided by 2 since:

\[ P(\log \lambda > \log \lambda_{\text{obs}}) = 1 - \left[0.5 + 0.5\{1-P(\log \lambda > \log \lambda_{\text{obs,old}})\}\right] \]

\[ = 0.5P(\log \lambda > \log \lambda_{\text{obs,old}}) \]
Table: Poisson regression estimates in the rare events meta–analysis of Rosiglitazone; Log-L stands for the maximised log-likelihood

<table>
<thead>
<tr>
<th>Poisson model</th>
<th>estimate</th>
<th>confidence interval</th>
<th>Log-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>treatment MI</td>
<td>1.2561</td>
<td>0.9991 – 1.5793</td>
<td>-174.2054</td>
</tr>
<tr>
<td>treatment</td>
<td>1.2634</td>
<td>1.0006 – 1.5952</td>
<td>-137.9558</td>
</tr>
<tr>
<td>$\sigma_{\alpha}^2$</td>
<td>0.6346</td>
<td></td>
<td></td>
</tr>
<tr>
<td>treatment</td>
<td>1.2634</td>
<td>1.0006 – 1.5952</td>
<td>-137.9558</td>
</tr>
<tr>
<td>$\sigma_{\alpha}^2$</td>
<td>0.6346</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_{\beta}^2$</td>
<td>0.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table: Poisson regression estimates in the rare events meta–analysis of Rosiglitazone; Log-L stands for the maximised log-likelihood

<table>
<thead>
<tr>
<th>Poisson model</th>
<th>estimate</th>
<th>confidence interval</th>
<th>Log-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>CV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>treatment</td>
<td>1.1281</td>
<td>0.8579 – 1.4835</td>
<td>-172.0216</td>
</tr>
<tr>
<td>treatment</td>
<td>1.0192</td>
<td>0.7737 – 1.3426</td>
<td>-100.3095</td>
</tr>
<tr>
<td>( \sigma_\alpha )</td>
<td>1.2294</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sigma_\beta )</td>
<td>1.12294</td>
<td>0.7737 – 1.3426</td>
<td>-100.3095</td>
</tr>
</tbody>
</table>
model evaluation

▶ for model assessment we will use criteria that compromise between model fit and model complexity

▶ Akaike information criterion

\[ AIC = -2 \log L + 2p \]

▶ Bayesian Information criterion

\[ BIC = -2 \log L + p \log k \]

▶ where \( p \) is the number of parameters in the model

▶ and \( k \) is the number of trials in the meta-analysis

▶ we seek a model for which AIC and/or BIC are small
Contents

What are rare events data?

What are the problems with rare events trials?

Effect estimation using Mantel-Haenszel

Poisson with fixed and random effects

Sensitivity analysis: the effect of excluding zero-studies

Zero-inflation models

Logistic regression modelling

Conditional logistic regression modelling
sensitivity analysis:

how does the effect estimate of the risk ratio depend on the exclusion/inclusion of

- double-zero (DZ)
- single-zero (SZ)

studies?
Meta-Analysis of Rare Events

Sensitivity analysis: the effect of excluding zero-studies

Table: Poisson random effects regression estimates of the risk ratio: the effect of excluding double-zero (DZ) and single-zero (SZ) studies and none excluded (NONE); number of studies included is given in brackets in the first column

<table>
<thead>
<tr>
<th>excluding (k)</th>
<th>RR</th>
<th>SE</th>
<th>Z</th>
<th>P-value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NONE(56)</td>
<td>1.2633</td>
<td>0.1503</td>
<td>1.96</td>
<td>0.049</td>
<td>1.0006 – 1.5952</td>
</tr>
<tr>
<td>DZ(41)</td>
<td>1.2634</td>
<td>0.1503</td>
<td>1.97</td>
<td>0.049</td>
<td>1.0008 – 1.5955</td>
</tr>
<tr>
<td>SZ(15)</td>
<td>1.2101</td>
<td>0.1512</td>
<td>1.53</td>
<td>0.127</td>
<td>0.9473 – 1.5458</td>
</tr>
<tr>
<td><strong>CV</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NONE(56)</td>
<td>1.0193</td>
<td>0.1433</td>
<td>0.14</td>
<td>0.892</td>
<td>0.7738 – 1.3426</td>
</tr>
<tr>
<td>DZ(27)</td>
<td>1.0246</td>
<td>0.1441</td>
<td>0.17</td>
<td>0.863</td>
<td>0.7778 – 1.3497</td>
</tr>
<tr>
<td>SZ(8)</td>
<td>0.9427</td>
<td>0.1395</td>
<td>-0.40</td>
<td>0.690</td>
<td>0.7054 – 1.2599</td>
</tr>
</tbody>
</table>
Contents

What are rare events data?

What are the problems with rare events trials?

Effect estimation using Mantel-Haenszel

Poisson with fixed and random effects

Sensitivity analysis: the effect of excluding zero-studies

Zero-inflation models

Logistic regression modelling

Conditional logistic regression modelling
Zero-inflation models

count data with many zeros lead to the question:
- is there an excess of zero counts relative to the Poisson model
- an excess in zero-counts is called **zero-inflation**

\[
Pr[X = 0] = \pi + (1 - \pi)Po(0|\mu) \quad (2)
\]

\[
Pr[X = x] = (1 - \pi)Po(x|\mu) \text{ for } x = 1, 2, ...
\quad (3)
\]
Meta-Analysis of Rare Events

Zero-inflation models
Zero-inflation models

Lambert (1992) extended the simple ZIP-model to covariates:

\[
\log \mu_{ij} = \log P_{ij} + \alpha + \beta \times j
\]  
\[
\text{logit } \pi_{ij} = \log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha' + \beta' \times j.
\]
Meta-Analysis of Rare Events

- Zero-inflation models

---

### Zero-inflated Poisson Regression

**Fitting full model:**

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Log Likelihood</th>
<th>(not concave)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>-173.39408</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>-172.39408</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-172.39408</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-172.39408</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>-172.39408</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>-172.39408</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>-172.39408</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>-172.39408</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>-172.39408</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>-172.39408</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>-172.39408</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>-172.39408</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>-172.39408</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>-172.39408</td>
<td></td>
</tr>
</tbody>
</table>

**Inflation model = logit**

Log likelihood = -171.9951

<table>
<thead>
<tr>
<th></th>
<th>IRR</th>
<th>Std. Err.</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td>cv</td>
<td>1.131013</td>
<td>0.1588599</td>
<td>0.88</td>
</tr>
<tr>
<td>treatment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>exposure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>inflate</td>
<td>-3.522325</td>
<td>4.462084</td>
<td>-0.79</td>
</tr>
<tr>
<td>_cons</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.430</td>
<td>12.26785</td>
<td>5.223199</td>
</tr>
</tbody>
</table>
Contents

What are rare events data?

What are the problems with rare events trials?

Effect estimation using Mantel-Haenszel

Poisson with fixed and random effects

Sensitivity analysis: the effect of excluding zero-studies

Zero-inflation models

Logistic regression modelling

Conditional logistic regression modelling
Logistic regression

let $Y_{ij}$ denote the binary outcome for an event ($Y_{ij} = 1$) in study $i$ and treatment arm $j$ ($j = 0, 1$)

$\pi_{ij} = P(Y_{ij} = 1)$ probability of an event

logistic transformation

$$\log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha + \beta \times j$$

so that $\beta$ is the \textbf{log-odds ratio}
Logistic regression model

- each trial arm within each study contributes a binomial likelihood

\[
\begin{align*}
\binom{n_{ij}}{x_{ij}} \pi_{ij}^{x_{ij}} (1 - \pi_{ij})^{n_{ij} - x_{ij}}
\end{align*}
\]

- where

\[
\log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha + \beta \times j
\]
Logistic likelihood

\[ L = \prod_i \prod_j \left( \frac{n_{ij}}{x_{ij}} \right)^{\pi_{ij} x_{ij}} (1 - \pi_{ij})^{n_{ij} - x_{ij}} \]

where

\[ \log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha + \beta \times j \]
Logistic regression with random intercept effect for study

\[ \log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha_i + \beta \times j \]

\[ \alpha_i \sim N(\alpha, \sigma^2_\alpha) \]

where \( \alpha_i \) is a random intercept effect.
Mixed Logistic Likelihood

\[ L = \prod_i \int_{\alpha_i} \prod_j \left( \frac{n_{ij}}{x_{ij}} \right) \pi_{ij}^{x_{ij}} (1 - \pi_{ij})^{n_{ij} - x_{ij}} \phi(\alpha_i) d\alpha_i \]

where \( \phi(\alpha_i) \) is a normal density with mean \( \alpha \) and variance \( \sigma^2_\alpha \)

and

\[ \log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha_i + \beta \times j \]
Logistic regression with random intercept effect for study

\[ \log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha_i + \beta_i \times j \]

where \( \alpha_i \sim N(\alpha, \sigma_{\alpha}^2) \) is a random intercept effect

and \( \beta_i \sim N(\beta, \sigma_{\beta}^2) \) is a random slope (treatment) effect
Mixed Logistic Likelihood

\[ L = \prod_i \int_{\alpha_i} \left( \int_{\beta_i} \prod_j \left( \frac{n_{ij}}{x_{ij}} \right) \pi_{ij}^{x_{ij}} (1 - \pi_{ij})^{n_{ij} - x_{ij}} \phi(\beta_i) \, d\beta_i \right) \phi(\alpha_i) \, d\alpha_i \]

- where \( \phi(\alpha_i) \) is a normal density with mean \( \alpha \) and variance \( \sigma_{\alpha}^2 \)
- where \( \phi(\beta_i) \) is a normal density with mean \( \beta \) and variance \( \sigma_{\beta}^2 \)
- and

\[ \log \frac{\pi_{ij}}{1 - \pi_{ij}} = \alpha_i + \beta_i \times j \]
Meta-Analysis of Rare Events

Logistic regression modelling

Mixed-effects logistic regression

| x               | Odds Ratio | Std. Err. | z   | P>|z| | [95% Conf. Interval] |
|-----------------|------------|-----------|-----|-----|----------------------|
| Treat_bin_cons  | .2458528   | .0857953  | -4.02 | 0.000 | .1240598 .4872135   |
|                 | .0310596   | .0077302  | -13.95 | 0.000 | .0190698 .0505877   |
| study_02 var(Treat~n) | .6187426 | .5829695 |               |       | .0976161 3.921919    |
|                 | .7717062   | .3642633  |               |       | .3059603 1.946431    |

LR test vs. logistic regression: chi2(2) = 50.11 Prob > chi2 = 0.0000
Table: Logistic regression estimates in the rare evidence meta–analysis of CRBSI; Log-L stands for the maximised log-likelihood

<table>
<thead>
<tr>
<th>logistic model</th>
<th>estimate</th>
<th>confidence interval</th>
<th>Log-L</th>
</tr>
</thead>
<tbody>
<tr>
<td>treatment</td>
<td>0.30</td>
<td>0.20 – 0.47</td>
<td>-103.27</td>
</tr>
<tr>
<td>treatment</td>
<td>0.29</td>
<td>0.19 – 0.46</td>
<td>-79.70</td>
</tr>
<tr>
<td>$\sigma^2_\alpha$</td>
<td>0.74</td>
<td>0.30 – 1.87</td>
<td></td>
</tr>
<tr>
<td>treatment</td>
<td>0.25</td>
<td>0.12 – 0.49</td>
<td>-78.22</td>
</tr>
<tr>
<td>$\sigma^2_\alpha$</td>
<td>0.77</td>
<td>0.31 – 1.95</td>
<td></td>
</tr>
<tr>
<td>$\sigma^2_\beta$</td>
<td>0.62</td>
<td>0.10 – 3.92</td>
<td></td>
</tr>
</tbody>
</table>
Contents

What are rare events data?

What are the problems with rare events trials?

Effect estimation using Mantel-Haenszel

Poisson with fixed and random effects

Sensitivity analysis: the effect of excluding zero-studies

Zero-inflation models

Logistic regression modelling

Conditional logistic regression modelling
**recall:**

let $RR = \frac{\mu_1}{\mu_0}$ and $X_i = X_{i1} + X_{i0}$

- in study $i$, for treatment

$$E(X_{i1}) = \mu_1 P_{i1}$$

for control

$$E(X_{i0}) = \mu_0 P_{i0}$$
it follows:

- then \( E(X_{i1} + X_{i0}) = \mu_1 P_{i1} + \mu_0 P_{i0} \) so that

\[
E(X_{i1}|X_i) = X_i \frac{\mu_1 P_{i1}}{\mu_1 P_{i1} + \mu_0 P_{i0}} = X_i \frac{RR \frac{P_{i1}}{P_{i0}}}{1 + RR \frac{P_{i1}}{P_{i0}}}
\]

depends only on \( RR \), the parameter of interest
Meta-Analysis of Rare Events

Conditional logistic regression modelling

Table: Layout for conditional logistic regression in study $i$

<table>
<thead>
<tr>
<th></th>
<th>treatment</th>
<th>control</th>
<th>margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>events</td>
<td>$X_{i1}$</td>
<td>$X_{i0}$</td>
<td>$X_i$</td>
</tr>
<tr>
<td>person time</td>
<td>$P_{i1}$</td>
<td>$P_{i0}$</td>
<td>$P_i$</td>
</tr>
</tbody>
</table>

$X_{i1}|X_i \sim Bin(q_i,X_i)$ with $q_i = \frac{\mu_1 P_{i1}}{\mu_1 P_{i1} + \mu_0 P_{i0}} = \frac{RR \frac{P_{i1}}{P_{i0}}}{1 + RR \frac{P_{i1}}{P_{i0}}}$
Meta-Analysis of Rare Events

Conditional logistic regression modelling

**furthermore:**

- let $RR = \exp(\beta)$

\[
q_i = \frac{RR \frac{P_{i1}}{P_{i0}}}{1 + RR \frac{P_{i1}}{P_{i0}}}
\]

\[
= \frac{\exp[\beta + \log(\frac{P_{i1}}{P_{i0}})]}{1 + \exp[\beta + \log(\frac{P_{i1}}{P_{i0}})]}
\]

\[
\frac{q_i}{1 - q_i} = \exp[\beta + \log(\frac{P_{i1}}{P_{i0}})]
\]

\[
\log \left( \frac{q_i}{1 - q_i} \right) = \beta + \log(\frac{P_{i1}}{P_{i0}})
\]
hence:

\[
\log \left( \frac{q_i}{1 - q_i} \right) = \beta + \log \left( \frac{P_{i1}}{P_{i0}} \right)
\]

we find \( \hat{RR} \) as logistic regression with intercept only and offset \( \log \left( \frac{P_{i1}}{P_{i0}} \right) \)

note that \( \beta \) is a log-risk ratio
**Meta-Analysis of Rare Events**

Conditional logistic regression modelling

---

**Table:** Meta-analysis on rare evidence data on the effect of anti-infective-treated catheter in comparison to standard catheter; CRBSI \((X_{i1}, X_{i0})\) refers to catheter-related bloodstream infection events, \(n_{i1}, n_{i0}\) is the size of the respective study arm

<table>
<thead>
<tr>
<th>study ID</th>
<th>control arm</th>
<th>treatment arm</th>
<th>conditional</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(X_{i0})</td>
<td>(X_{i1})</td>
<td>(X_i)</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>
. melogit xt, offset(log_ratio) binomial(xsum) or

Iteration 0:  log likelihood = -26.454133
Iteration 1:  log likelihood = -26.199518
Iteration 2:  log likelihood = -26.199183
Iteration 3:  log likelihood = -26.199183

Logistic regression  Number of obs  =  18
Binomial variable:  xsum

Log likelihood = -26.199183  Wald chi2(0)  =   .
Prob > chi2     =   .

|           | Odds Ratio | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|-----------|------------|-----------|-------|------|---------------------|
| _cons     | .3072359   | .0678268  | -5.35 | 0.000 | .1993228             | .4735732 |
| log_ratio | 1 (offset) |           |       |      |                     |          |

Meta-Analysis of Rare Events
Conditional logistic regression modelling
can be easily extend to random effects model

\[
\log \left( \frac{q_i}{1 - q_i} \right) = \beta_i + \log \left( \frac{P_{i1}}{P_{i0}} \right)
\]

with \( \beta_i \sim N(\beta, \sigma^2_\beta) \)

Mixed-effects logistic regression

| Estimate | Std. Err. | z    | P>|z| | [95% Conf. Interval] |
|----------|-----------|------|------|---------------------|
| _cons    | 0.2718073 | 0.0920478 | -3.85 | 0.000               |
|          |           |       |      | 0.1399591 - 0.527863 |
| log_ratio| 1 (offset)|      |      |                     |
| study_short | 0.6007426 | 0.5915841 | 0.0871894 | 4.139168 |

LR test vs. logistic regression: \( \text{chibar2(01)} = 2.81 \) Prob>\( \text{chibar2} = 0.0469 \)