Validity of summary statistics-based mixed-effects group fMRI
Camille Maumet

To cite this version:
Camille Maumet. Validity of summary statistics-based mixed-effects group fMRI. OHBM 2018, Jun 2018, Singapore, Singapore. hal-04091271
Validity of summary statistics-based mixed-effects group fMRI

Camille Maumet and Thomas Nichols

1. University of Rennes, Inria, CNRS, Inserm, IRISA, Rennes, France.
What is the link between meta-analysis and group fMRI?

**Subject level**
- Pre-processed data
- Model fitting and estimation
- Contrast and std. err. maps

**Group level**
- Model fitting and estimation
- Contrast and std. err. maps

**Meta-analytic level**
- Model fitting and estimation
- Contrast and std. err. maps
- Inference
- Detections (meta-analysis)
Methods
GLM for group fMRI

\[ Y = X\beta + \varepsilon \]

where \( \varepsilon \sim N(0, \tau^2 + \sigma^2_i) \)
Solving the GLM: OLS vs. FGLS

\[ Y = X\beta + \varepsilon, \quad \text{where } \varepsilon \sim N(0, \tau^2 + \sigma_i^2) \]

Assuming \( \tau^2 + \sigma_i^2 \) constant, **Ordinary Least Squares (OLS)** gives the following group statistic estimate:

\[
T = \left( \sum_{i=1}^{N} \frac{Y_i}{\sqrt{N}} \right) / \hat{\sigma}_C^2
\]

where \( \hat{\sigma}_C^2 \) is the usual one-sample variance.

Under \( H_0 \):
\[ T \sim \tau_{N-1} \]
Solving the GLM: OLS vs. FGLS

\[ Y = X\beta + \varepsilon, \quad \text{where } \varepsilon \sim \text{N}(0, \tau^2 + \sigma^2_i) \]

Feasible Generalised Least Squares (FGLS) gives the following group statistic estimate:

\[
T = \left( \sum \kappa_i Y_i \right) / \sqrt{\sum_{i=1}^{N} \kappa_i} \quad \text{where } \kappa_i = 1 / (\hat{\tau}^2 + s_i^2)
\]

where \( s_i^2 \) is the study i’s sampling variance.

In large samples, under \( H_0 \): \( T \sim \mathcal{T}_{N - 1} \)
Solving the GLM: OLS vs. FGLS

<table>
<thead>
<tr>
<th></th>
<th>Statistic</th>
<th>Assumptions</th>
<th>Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS</td>
<td>( T = \left( \sum_{i=1}^{N} \frac{Y_i}{\sqrt{N}} \right) / \frac{\sigma_i^2}{\sigma_C^2} )</td>
<td>( \tau^2 + \sigma_i^2 ) constant</td>
<td>'One-sample t-test' (default)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>'Mixed effects: Simple OLS'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3dttest++</td>
</tr>
<tr>
<td>FGLS</td>
<td>( T = \left( \sum \kappa_i Y_i \right) / \sqrt{\sum_{i=1}^{N} \kappa_i} )</td>
<td>Large samples</td>
<td>'Mixed-effects analysis'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>'Mixed effects: FLAME1' (default)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3dMEMA</td>
</tr>
</tbody>
</table>
**Goal:** assess the validity of OLS and FGLS under violation of their assumption

**Method:** Monte Carlo simulations

Difference between sorted observed P-values (computed by OLS or FGLS) and sorted expected -log10 P-values.

"True" -log10 P-values computed empirically by MC simulations

If the method is exact then the difference will lie in the 95% CI (grey ribbon) around zero (black line).

Conservative method

Invalid method
And in meta-analyses?

Question: Does this issue also affects group fMRI studies?
Simulations
Between- and within-subject variances in real fMRI data

21 studies investigating pain

<table>
<thead>
<tr>
<th>Study</th>
<th># subjects</th>
<th># outliers</th>
<th>Mean $\tau^2$</th>
<th>Mean $\sigma^2_{\text{in}}$</th>
<th>Max/Min $\sigma^2_{\text{in}}$ ratio</th>
<th>Mean $\sigma^2_{\text{out}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>25</td>
<td>2</td>
<td>354</td>
<td>100</td>
<td>4</td>
<td>282</td>
</tr>
<tr>
<td>02</td>
<td>25</td>
<td>1</td>
<td>680</td>
<td>100</td>
<td>5</td>
<td>317</td>
</tr>
<tr>
<td>03</td>
<td>20</td>
<td>2</td>
<td>800</td>
<td>100</td>
<td>5</td>
<td>281</td>
</tr>
<tr>
<td>04</td>
<td>20</td>
<td>3</td>
<td>886</td>
<td>100</td>
<td>9</td>
<td>347</td>
</tr>
<tr>
<td>05</td>
<td>9</td>
<td>0</td>
<td>1802</td>
<td>3</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td><strong>0-15%</strong></td>
<td><strong>18-170</strong></td>
<td><strong>100</strong></td>
<td><strong>2-9</strong></td>
<td></td>
<td><strong>164-300</strong></td>
</tr>
</tbody>
</table>
Simulation setup

N ∈ \{ 25, 50 \} subjects

τ^2 = 1 and mean(\sigma^2_{in}) = 1

Two settings for the within-subject variance:

- max(\sigma^2_{in}) / min(\sigma^2_{in}) ∈ \{ 1, 2, 4, 6, 8, 10 \}
- 16% of outliers with \sigma^2_{out} ∈ \{ 0.25, 0.5, 2, 4 \}

10^6 realisations
Results
Varying levels of heteroscedasticity

\[
\frac{\max(\sigma^2_{in})}{\min(\sigma^2_{in})}
\]
In the presence of outliers

\[ \sigma_{out}^2 / \text{mean}(\sigma_{in}^2) \]
Real data
Conclusions

FGLS can be invalid in small samples, especially in the presence of strong heteroscedasticity or ‘low variance’ outliers.

RFX remained valid in all studied settings.

Acknowledgments: This work was supported by the Wellcome Trust. We gratefully acknowledge the use of the pain dataset from the Tracey pain group, FMRIB, Oxford. Part of the work was conducted while TEN and CM were at the University of Warwick and used the High Performance Computing cluster of the Department of Statistics, University of Warwick.