Meta-analysis techniques

Finn Årup Nielsen

CIMBI
Department of Informatics and Mathematical Modelling
Technical University of Denmark

September 19, 2012
Meta-analysis techniques

Why meta-analysis?

“Why Most Published Research Findings Are False” (Ioannidis, 2005):

“There is increasing concern that in modern research, false findings may be the majority or even the vast majority of published research claims.”

“The greater the flexibility in designs, definitions, outcomes, and analytical modes in a scientific field, the less likely the research findings are true.”

Is the last quote especially true for neuroimaging?
Why meta-analysis?

“The Difference Between ‘Significant’ and ‘Not Significant’ is not Itself Statistically Significant” (Gelman and Stern, 2006)

Two apparently conflicting studies—one significant, another not significant—may not necessarily be conflicting. One may simply not have enough power.
Meta-analysis techniques

Information increase

There are too much data for one person to grasp

The results across experiments are too conflicting

Need for tools that collect data across studies, bring order to data, make search easy and automate analyses to bring out consensus results: **meta-analysis and meta-analytic databases**

Figure 1: Increase in the number of articles in PubMed which are returned after searching on posterior cingulate and related brain areas.
Meta-analysis

The page-one definition (Hartung et al., 2008):

The statistical analysis of a large collection of analysis results from individual studies for the purpose of integrating the findings.
The effect size is the central measure in ordinary meta-analysis.

The mean effect size is (usually) independent of the number of subjects.

Effect sizes can be formed from other variables than those shown in the figure: correlation, proportions.
Meta-analysis techniques

Standardized mean difference (SMD)

For continuous data with (approximately) normal distribution. Example:

<table>
<thead>
<tr>
<th>Study</th>
<th>Bipolars</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
</tr>
<tr>
<td>Strakowski SM, 1999</td>
<td>24</td>
<td>7.1</td>
</tr>
<tr>
<td>Altshuler LL, 2000</td>
<td>24</td>
<td>3825.9</td>
</tr>
</tbody>
</table>
| ...                          | ...     | ...     | ...

Table 1: Data for meta-analysis with SMD. Amygdala volume from bipolar patients and controls.

Take the difference between the means of the two groups (experimentals $e$ and controls $c$) and divide by the pooled standard deviation

$$g_{smd} = \frac{\bar{x}_e - \bar{x}_c}{s_{pooled}}$$

$g_{smd}$ independent of unit of the original study, e.g., whether a brain volume was reported in cubic millimeters or cubic centimeters. It is also independent of the number of subjects ($n_e$ in experimental group)
Meta-analysis techniques

SDM — details & inference

With, e.g., \( s_e \) as the standard deviation for the experimental group:

\[
\text{s}_{\text{pooled}} = \sqrt{\frac{(n_e-1)s_e^2 + (n_c-1)s_c^2}{n_e + n_c - 2}}
\]  

(1)

\[
d_{\text{smd}} = \mathbb{E}[g_{\text{smd}}] \approx \left(1 - \frac{3}{4(n_e + n_c) - 9}\right) g_{\text{smd}}
\]  

(2)

\[
\sqrt{\text{Var}[g_{\text{smd}}]} \approx \frac{1}{\tilde{n}} + \frac{g_{\text{smd}}}{2(n_e + n_c - 3.94)} \quad \text{where} \quad \tilde{n} = \frac{n_e n_c}{n_e + n_c}.
\]  

(3)

If the effect size is small \((g_{\text{smd}} \to 0)\) and the two groups are of the same size \((n_e = n_c)\) then the variance becomes proportional to the number of subjects in the groups

\[
\sqrt{\text{Var}[g_{\text{smd}}]} \approx \frac{2}{n_e} = \frac{2}{n_c},
\]  

(4)

i.e., the more subjects the better the effect size is determined.
Odds ratio

For binary data we can construct a contingency table for the results:

<table>
<thead>
<tr>
<th></th>
<th>“Success”</th>
<th>“Failure”</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Experimentals”</td>
<td>$n_{es}$</td>
<td>$n_{ef}$</td>
<td>$n_e$</td>
</tr>
<tr>
<td>“Controls”</td>
<td>$n_{cs}$</td>
<td>$n_{cf}$</td>
<td>$n_c$</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>$n_s$</td>
<td>$n_f$</td>
<td>$n$</td>
</tr>
</tbody>
</table>

One effect size for binary data is the (natural) logarithm of the odds ratio (Hartung et al., 2008, p. 20)

$$d_{or} = \ln \left[ \frac{c(n_{es})/c(n_e - n_{es})}{c(n_{cs})/c(n_c - n_{cs})} \right], \quad \text{where e.g., } c(x) = x + 0.5$$

Addition of 0.5 to get around a problem if there is zero count in any of the cells (Hartung et al., 2008, p. 117)

An estimate of its variance as an estimator is

$$\hat{\text{Var}}[d_{or}] = \frac{1}{c(n_{es})} + \frac{1}{c(n_e - n_{es})} + \frac{1}{c(n_{cs})} + \frac{1}{c(n_c - n_{cs})}.$$
Meta-analysis techniques

Variance ratio

Example claim: Men have higher variation in intelligence than women.

We should test this

\[ \sigma_{\text{men}}^2 > \sigma_{\text{women}}^2 \]

The logarithm of the ratio between the two variations (lnvr) results in a good statistics (Shaffer, 1992)

\[ d_{vr} = \ln \left( \frac{s_c^2}{s_e^2} \right) \tag{5} \]

This is better than the variance ratio \( s_e^2 / s_c^2 \) (or standard deviation difference \( s_e - s_c \))
Meta-analysis techniques

Funnel plot with multiple studies

Scatter plot of the effect sizes and their standard errors (related to variance and sample size) for multiple studies.

May indicate publication bias, if researchers of small studies only publish if they see an effect then the funnel plot becomes asymmetric.
Meta-analysis techniques

Forrest plot with multiple studies

Strakowski SM
Altshuler LL
Getz GE
Blumberg HP
DeBello MP
Blumberg HP
Chang K (A)
Frazier JA (A)

Meta-analysis (fixed effect)
Meta-analysis (random effects)

Forest plot shows the effect size for 8 different studies (Amygdala volume in bipolar disorder).

The squares are the 8 effect sizes \( (d_i) \) and the lines indicate 0.05-confidence interval: \( d_i \pm 1.96 \sqrt{\text{Var}[d_i]} \)
Combing effect sizes across studies

Meta-analytic effect size: Inverse variance-weighting (in so-called fixed effect) for weighted averging of studies \((\text{Hartung et al., 2008, p. 36})\)

\[
    d_{\text{meta}} = \frac{\sum_i w_i d_i}{\sum_i w_i} \quad (6)
\]

where \(d_i\) is the effect size for the \(i\)th study and the weight for the \(i\)th study is determined as the inverse variance

\[
    w_i = \frac{1}{\text{Var}[d_i]} \quad (7)
\]

Variance:

\[
    \text{Var}[d_{\text{meta}}] \approx \frac{1}{\sum_i w_i} \quad (8)
\]

when the number of subject for study \(i\) increases \((n_i \rightarrow \infty)\), then the variance decreases \((\text{Var}[d_i] \rightarrow 0)\), the weight increases \((w_i \rightarrow \infty)\) and the meta-analytic variance decreases \((\text{Var}[d_{\text{meta}}] \rightarrow 0)\)
Random effects analysis

“Random effects” in meta-analysis adds an extra parameter that models the variation between studies.

One approach is the so-called DerSimonian-Laird (Hartung et al., 2008, p. 108)
Free tools for meta-analysis

R with `meta` package by Guido Schwarzer

RevMan and Archie of the Cochrane Library ([Elamin et al., 2009](#))

Open science meta-analysis

Brede Wiki & its meta-analysis service

Brede Wiki for personality genetics

Online data and meta-analyses
Meta-analysis techniques

Brede Wiki with data

Store numerical data on a spreadsheet-like page in a MediaWiki. Describe the data in structured format on the wiki.

Data in the Brede Wiki primarily from large meta-analyses by Matthew Kempton, Institute of Psychiatry, and his coworkers (Kempton et al., 2008; Kempton et al., 2010; Kempton et al., 2011)

Finn Årup Nielsen

September 19, 2012
Meta-analysis techniques

Meta-analysis with the Brede Wiki

Major depressive disorder

- Brain volume abnormalities in major depressive disorder: a meta-analysis of magnetic resonance imaging studies.
- Structural neuroimaging studies in major depressive disorder: meta-analysis and comparison with bipolar disorder.
- Hippocampal volume in primary unipolar major depression: a magnetic resonance imaging study.
- Reduced caudate and nucleus accumbens response to rewards in unmedicated individuals with major depressive disorder.
- Regional central blood flow abnormalities in depressed patients with cognitive impairment.

Treatment response and genetics

- Tryptophan hydroxylase, 5-HTP, 5-HT2A, 5-HT2C, 1FQ1, and 1FQ2 polymorphisms and major depression.
- Genetic association analysis of serotonin 2A receptor gene (HTR2A) with bipolar disorder and major depressive disorder in the Japanese population.

Meta-analyses

<table>
<thead>
<tr>
<th>Topic</th>
<th>Data</th>
<th>Raw data</th>
<th>Meta-analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amygdala — Major depressive disorder — MaND</td>
<td>Data</td>
<td>CSV</td>
<td>Meta-analysis</td>
</tr>
<tr>
<td>Amygdala — Bipolar disorder — BiND</td>
<td>Data</td>
<td>CSV</td>
<td>Meta-analysis</td>
</tr>
<tr>
<td>Amygdala — Obsessive-compulsive disorder — ObND</td>
<td>Data</td>
<td>CSV</td>
<td>Meta-analysis</td>
</tr>
</tbody>
</table>

Finn Årup Nielsen
Mass meta-analysis

With numerical data and information about it in the wiki it is possible to download and analyze all meta-analysis together.

Here a L’Abbé-like plot of many of the meta-analyses in the Brede Wiki with effect magnitude on the x-axis and its uncertainty on the y-axis.
Meta-analysis techniques

Brede Wiki for personality genetics

<table>
<thead>
<tr>
<th>Effect</th>
<th>Std</th>
<th>P</th>
<th>Studies</th>
<th>Subjects</th>
<th>Gene</th>
<th>Polymorphism</th>
<th>Trait</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.854</td>
<td>0.223</td>
<td>0.00013</td>
<td>2</td>
<td>107</td>
<td>ESR1</td>
<td>TA repeat</td>
</tr>
<tr>
<td>2</td>
<td>-1.102</td>
<td>0.289</td>
<td>0.00014</td>
<td>2</td>
<td>245</td>
<td>HTR3A</td>
<td>C178T</td>
</tr>
<tr>
<td>3</td>
<td>-0.779</td>
<td>0.220</td>
<td>0.00039</td>
<td>1</td>
<td>90</td>
<td>ESR1</td>
<td>TA repeat</td>
</tr>
<tr>
<td>4</td>
<td>-0.445</td>
<td>0.135</td>
<td>0.00098</td>
<td>1</td>
<td>247</td>
<td>TH</td>
<td>TCAT repeat</td>
</tr>
<tr>
<td>5</td>
<td>-0.401</td>
<td>0.123</td>
<td>0.00108</td>
<td>1</td>
<td>315</td>
<td>DRD4</td>
<td>Exon 3 VNTR</td>
</tr>
<tr>
<td>6</td>
<td>0.165</td>
<td>0.051</td>
<td>0.00118</td>
<td>13</td>
<td>1747</td>
<td>MAOA</td>
<td>uVNTR</td>
</tr>
<tr>
<td>7</td>
<td>-0.393</td>
<td>0.123</td>
<td>0.00135</td>
<td>1</td>
<td>315</td>
<td>DRD4</td>
<td>Exon 3 VNTR</td>
</tr>
<tr>
<td>8</td>
<td>-1.355</td>
<td>0.427</td>
<td>0.00152</td>
<td>1</td>
<td>125</td>
<td>HTR3A</td>
<td>C178T</td>
</tr>
<tr>
<td>9</td>
<td>-0.758</td>
<td>0.240</td>
<td>0.00161</td>
<td>1</td>
<td>122</td>
<td>SLC6A4</td>
<td>5-HTTLPR</td>
</tr>
<tr>
<td>10</td>
<td>-0.174</td>
<td>0.055</td>
<td>0.00163</td>
<td>16</td>
<td>1791</td>
<td>SLC6A4</td>
<td>5-HTTLPR</td>
</tr>
</tbody>
</table>

Meta-analysis across traits and polymorphisms

Large-scale data mining across all recorded personality traits and genetic polymorphisms and present the result on the wiki.

Order meta-analytic results, e.g., with respect to $P$-value
**Meta-analysis techniques**

**Neuroimaging meta-analysis**

Image-based meta-analysis if you got the summary images ([Salimi-Khorshidi et al., 2009](#)).

Coordinate-based meta-analysis if you got the stereotaxic coordinates ([Fox et al., 1997](#); [Nielsen and Hansen, 2002](#); [Turkeltaub et al., 2001](#)).

Convolve a smooth kernel on its stereotaxic coordinate

**Tools:** BrainMap’s GingerALE, Brede Toolbox
Meta-analysis techniques

BrainMap

One of the first and most comprehensive databases (Fox et al., 1994; Fox and Lancaster, 2002)

Presently 85007 locations from 2238 papers (2012 September)

Graphical Internet-based interface in Java, sleuth, with search facilities, e.g., on author, 3D coordinate, an others and GingerALE meta-analysis
Example on meta-analysis

Coordinate-based meta-analysis with the Brede Toolbox on pain studies

Volume threshold at statistical values determined by resampling statistics (Nielsen, 2005).

Red areas are the most significant areas: Anterior cingulate, anterior insula, thalamus. In agreement with “human” reviewer (Ingvar, 1999).
The Semantic MediaWiki allows you to construct a database on the Web without having to setup a standard database management system :-)

Semantic MediaWiki has a flexible schema: You can add fields after you have setup “table”, e.g., add a “is peer reviewed” field or “imaging modality” field. :-)

A full setup of a Semantic MediaWiki with forms and templates requires mastering of the somewhat obscure MediaWiki template language :-(

Finn Årup Nielsen
Semantic MediaWiki

Our Semantic MediaWiki instance http://wikilit.referata.com/ setup for a systematic review

Semantic MediaWiki allows you to define forms for input: text and categorical.

Semantic queries can be made so you can get the inputted data in comma-separated values suitable for further numerical processing

...and you can filter with the queries, e.g., peer-reviewed publications if that category is setup
Automated literature reviews with text mining

Download papers, extract words and represent them in a bag-of-words matrix, perform topic mining with an unsupervised multivariate analysis method, e.g., non-negative matrix factorization, to find themes (Nielsen et al., 2005).

For a recent example with NeuroSynth see (Poldrack et al., 2012).

Figure 2: Some of the topics found in a corpus on posterior cingulate neuroimaging.
Combining the cognitive ontology in Brede with the coordinate-based meta-analysis to constructed a functional atlas.
What can you do?

Report the summary statistics: mean, standard deviation and number of subjects.

Report summary statistics for all groups

Open Science
References


