Software and model selection challenges in meta-analysis

Software and model selection challenges in meta-analysis

MetaEasy, model assumptions and homogeneity

Evan Kontopantelis David Reeves

Health Sciences Primary Care Research Centre
University of Manchester

RSS Primary Care Study Group
Errol Street, 2 July 2012

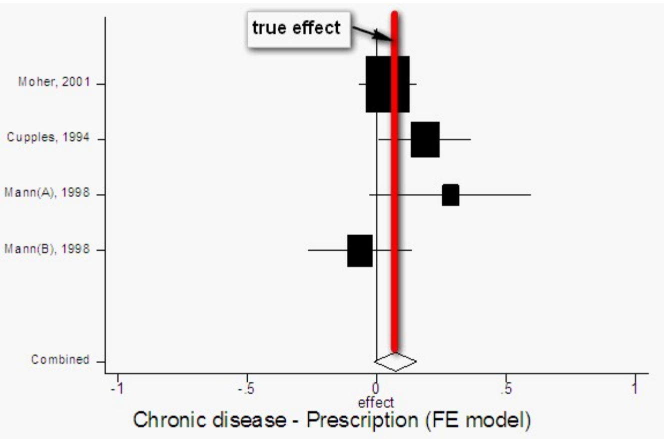
Outline

Heterogeneity

The big bad wolf

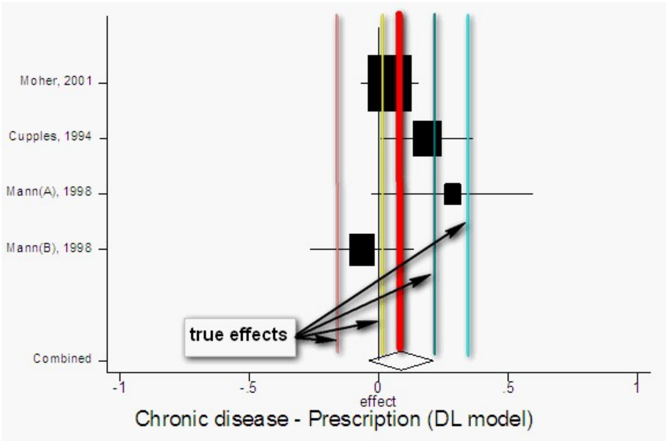
- When the effect of the intervention varies significantly from one study to another.
- It can be attributed to clinical and/or methodological diversity.
  - Clinical: variability that arises from different populations, interventions, outcomes and follow-up times.
  - Methodological: relates to differences in trial design and quality.
- Detecting quantifying and dealing with heterogeneity can be very hard.

Absence of heterogeneity



- Assumes that the true effects of the studies are all equal and deviations occur because of imprecision of results.
- Analysed with the fixed-effects method.

Presence of heterogeneity



- Assumes that there is variation in the size of the true effect among studies (in addition to the imprecision of results).
- Analysed with random-effects methods.

Challenges with meta-analysis

- Heterogeneity is common and the fixed-effect model is under fire.
- Methods are asymptotic: accuracy improves as studies increase. But what if we only have a handful, as is usually the case?
- Almost all random-effects models (except Profile Likelihood) do not take into account the uncertainty in $\hat{\tau}^2$. Is this, practically, a problem?
- DerSimonian-Laird is the most common method of analysis, since it is easy to implement and widely available, but is it the best?

Challenges with meta-analysis
...continued

- Can be difficult to organise since...
  - outcomes likely to have been disseminated using a variety of statistical parameters
  - appropriate transformations to a common format required
  - tedious task, requiring at least some statistical adeptness
- Parametric random-effects models assume that both the effects and errors are normally distributed. Are methods robust?
- Sometimes heterogeneity is estimated to be zero, especially when the number of studies is small. Good news?

Based on our original work...

Organising

- Data initially collected using data extraction forms.
- A spreadsheet is the next logical step to summarise the reported study outcomes and identify missing data.
- Since in most cases MS Excel will be used we developed an add-in that can help with most processes involved in meta-analysis.
- More useful when the need to combine differently reported outcomes arises.

Meta-analysis overview
A practical guide
Summary
the MetaEasy add-in
metaeff & metaan
Methods and performance
$\hat{\tau}^2 = 0$

What it can do

- Help with the data collection using pre-formatted worksheets.
- Its unique feature, which can be supplementary to other meta-analysis software, is implementation of methods for calculating effect sizes (& SEs) from different input types.
- For each outcome of each study...
  - it identifies which methods can be used
  - calculates an effect size and its standard error
  - selects the most precise method for each outcome

The University of Manchester MANCHESTER 1824
Kontopantelis, Reeves
Software and model selection challenges in meta-analysis

What it can do
...continued

- Creates a forest plot that summarises all the outcomes, organised by study.
- Uses a variety of standard and advanced meta-analysis methods to calculate an overall effect.
  - a variety of options is available for selecting which outcome(s) are to be meta-analysed from each study
- Plots the results in a second forest plot.
- Reports a variety of heterogeneity measures, including Cochran's Q, I^2, ${H_M}^2$ and $\hat{\tau}^2$ (and its estimated confidence interval under the Profile Likelihood method).

Advantages

- Free (provided Microsoft Excel is available).
- Easy to use and time saving.
- Extracted data from each study are easily accessible, can be quickly edited or corrected and analysis repeated.
- Choice of many meta-analysis models, including some advanced methods not currently available in other software packages (e.g. Permutations, Profile Likelihood, REML).
- Unique forest plot that allows multiple outcomes per study.
- Effect sizes and standard errors can be exported for use in other meta-analysis software packages.

Installing

- Latest version available from www.statanalysis.co.uk
- Compatible with Excel 2003, 2007 and 2010.
- Manual provided but also described in:
  - Kontopantelis E and Reeves D.
    MetaEasy: A Meta-Analysis Add-In for Microsoft Excel.
    Journal of Statistical Software, 30(7):1-25, 2009.

Play video clip

Meta-analysis overview
A practical guide
Summary
the MetaEasy add-in
metaeff & metaan
Methods and performance
$\hat{\tau}^2 = 0$

Stata implementation

- MetaEasy methods implemented in Stata under:
  - metaeff, which uses the different study input to provide effect sizes and SEs
  - metaan, which meta-analyses the study effects with a fixed-effect or one of five available random-effects models
- To install, type in Stata:
  - ssc install *<command name>*
    help *<command name>*
- Described in:
  - Kontopantelis E and Reeves D.
    metaan: Random-effects meta-analysis.
    The Stata Journal, 10(3):395-407, 2010.

The University of Manchester
MANCHESTER 1824
Kontopantelis, Reeves
Software and model selection challenges in meta-analysis

Many random-effects methods

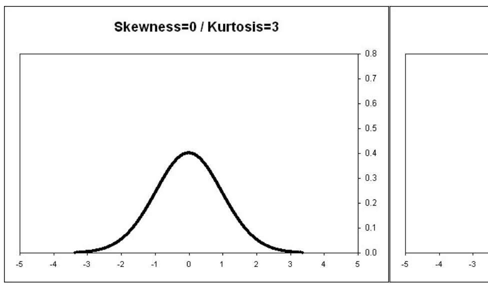
which to use?

- DerSimonian-Laird (DL): Moment-based estimator of both within and between-study variance.
- Maximum Likelihood (ML): Improves the variance estimate using iteration.
- Restricted Maximum Likelihood (REML): an ML variation that uses a likelihood function calculated from a transformed set of data.
- Profile Likelihood (PL): A more advanced version of ML that uses nested iterations for converging.
- Permutations method (PE): Simulates the distribution of the overall effect using the observed data.

Performance evaluation

our approach

- Simulated various distributions for the true effects:
  - Normal.
  - Skew-Normal.
  - Uniform.
  - Bimodal.
- Created datasets of 10,000 meta-analyses for various numbers of studies and different degrees of heterogeneity, for each distribution.



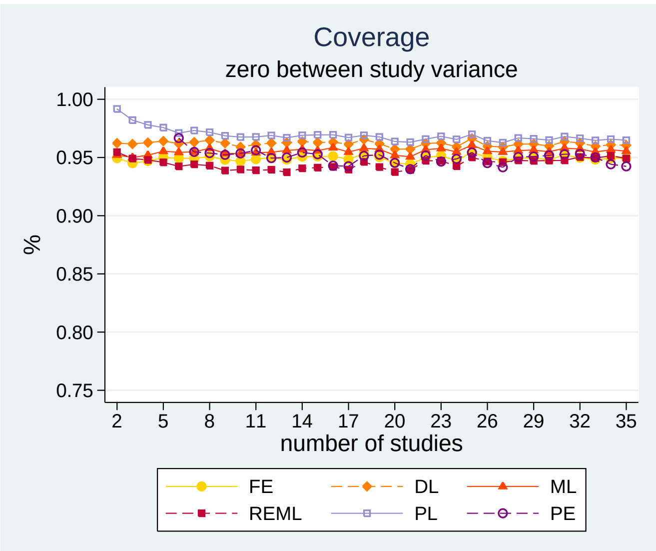
Performance evaluation

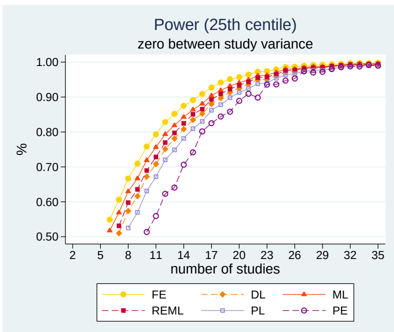
our approach

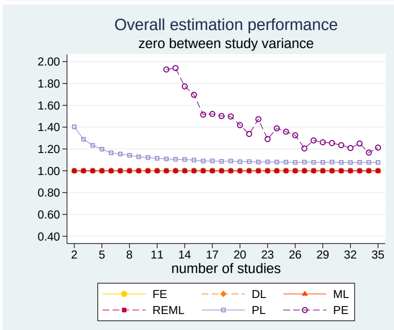
- Compared all methods in terms of:
  - Coverage, the rate of true negatives when the overall true effect is zero.
  - Power, the rate of true positives when the true overall effect is non-zero.
  - Confidence Interval performance, a measure of how wide the (estimated around the effect) CI is, compared to its true width.

Homogeneity

Zero between study variance

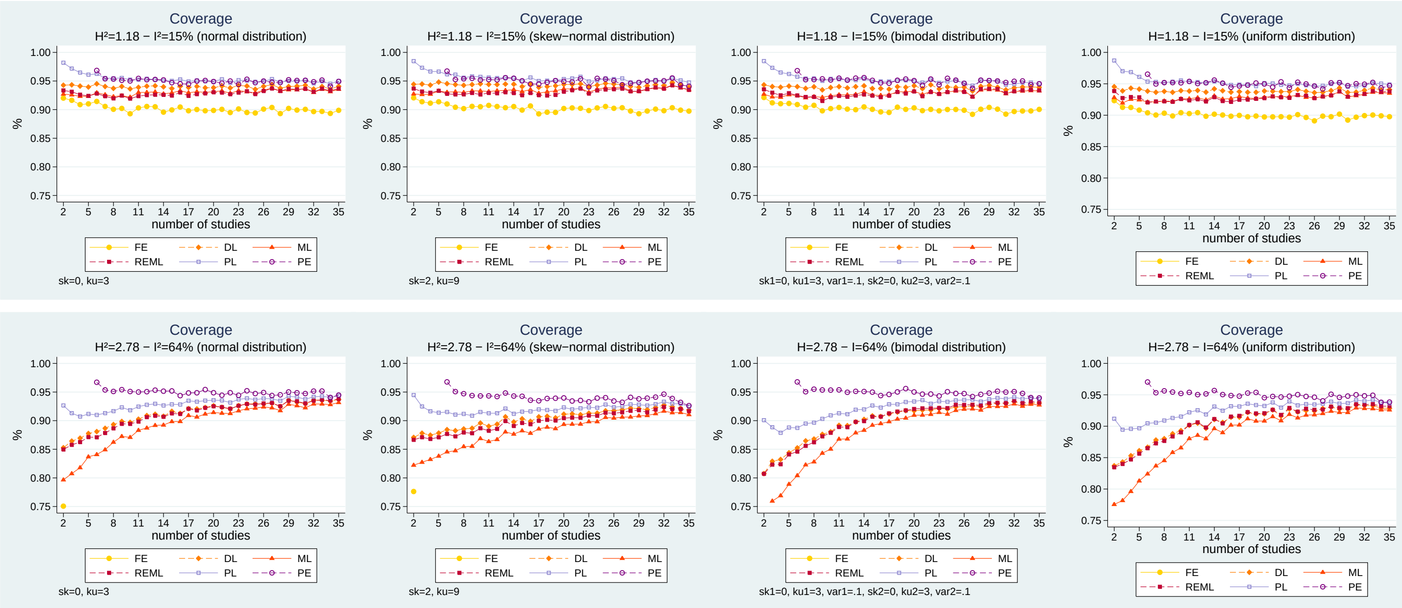






Coverage performance

Small and large heterogeneity under various distributional assumptions

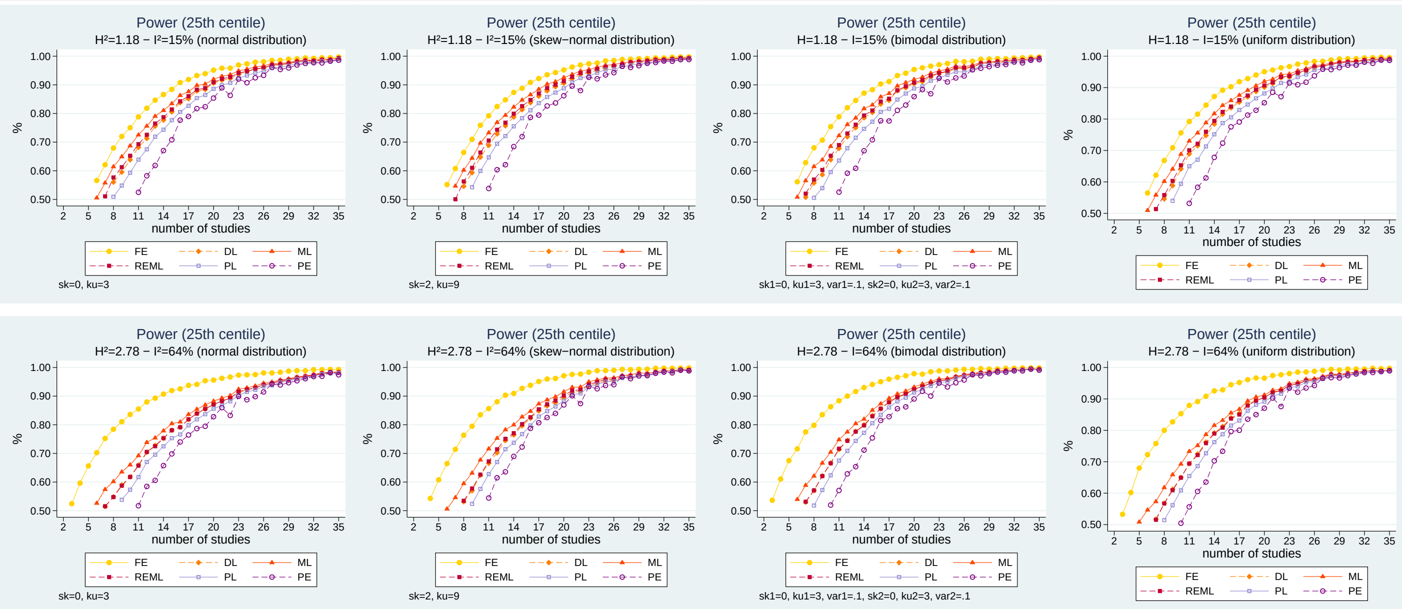


The University of Manchester
MANCHESTER 1824

Meta-analysis overview
A practical guide
Summary
the MetaEasy add-in
metaeff & metaan
Methods and performance
$\hat{\tau}^2 = 0$

Power performance

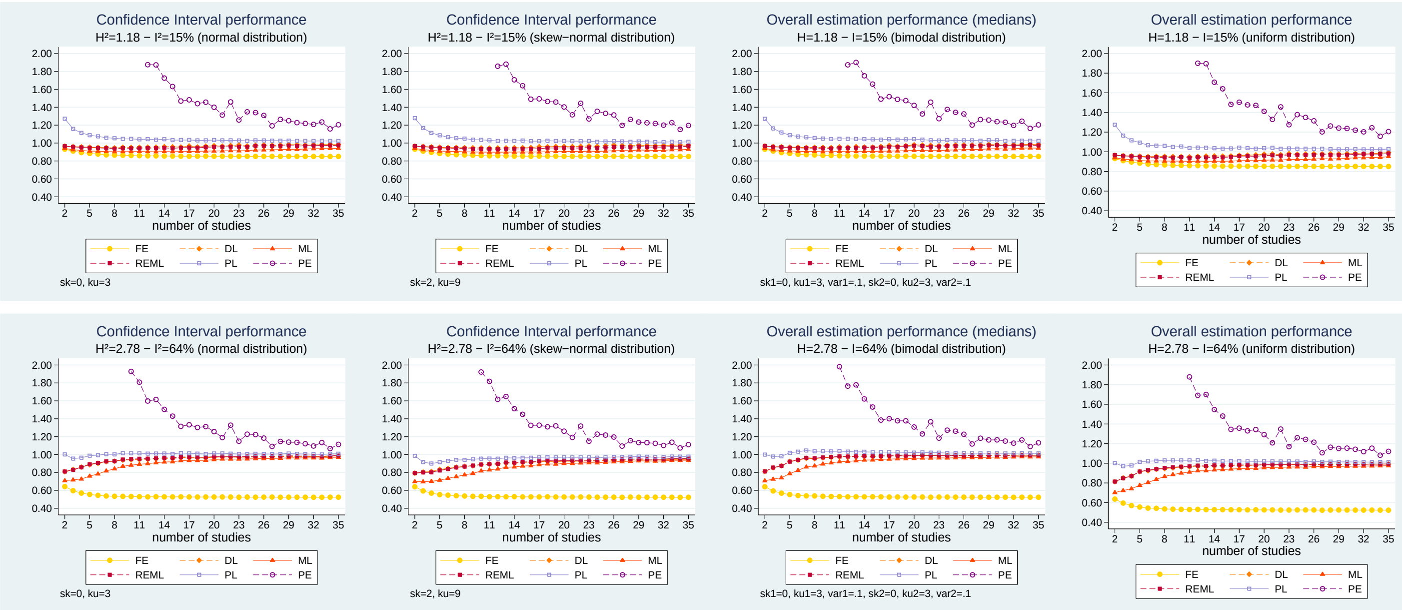
Small and large heterogeneity under various distributional assumptions



The University of Manchester
MANCHESTER 1824

CI performance

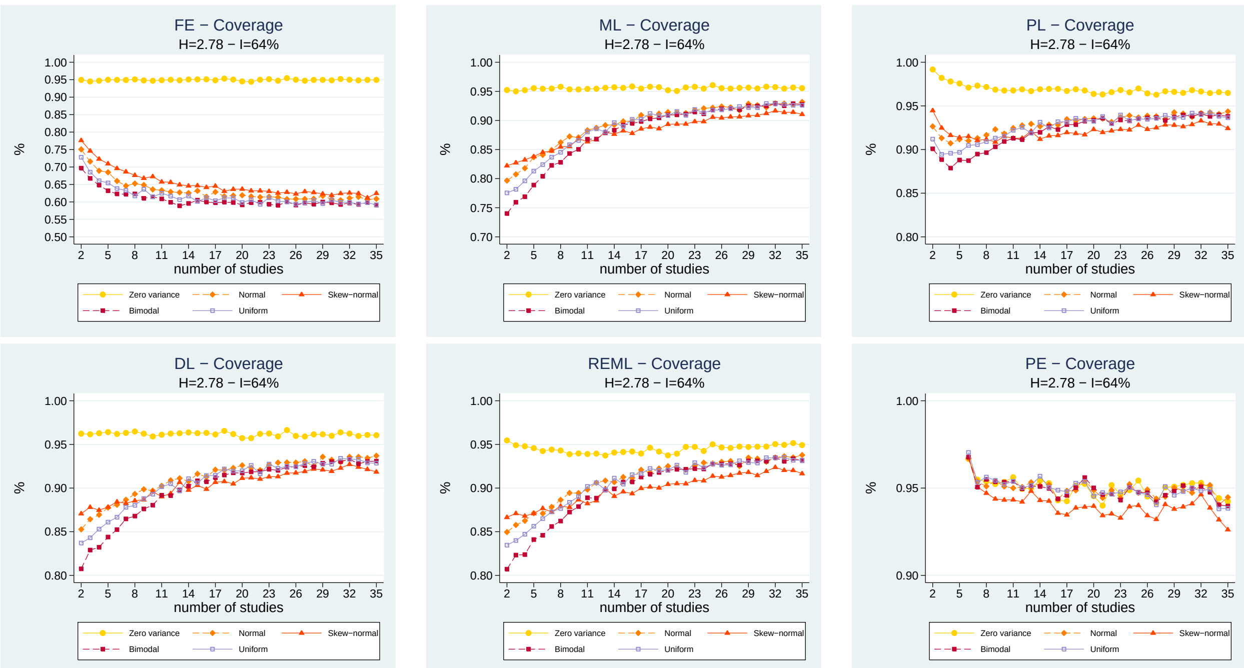
Small and large heterogeneity under various distributional assumptions



Meta-analysis overview
A practical guide
Summary
the MetaEasy add-in
metaeff & metaan
Methods and performance
$\hat{\tau}^2 = 0$

Coverage by method

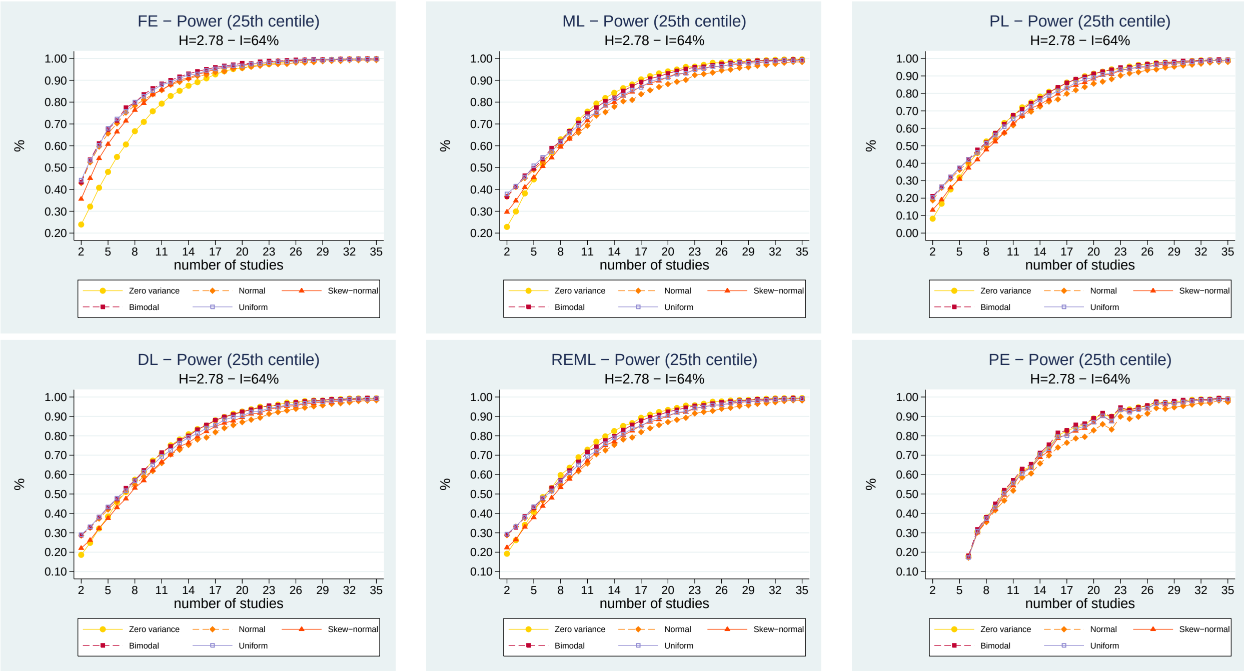
Large heterogeneity across various between-study variance distributions



Kontopantelis, Reeves
Software and model selection challenges in meta-analysis

Power by method

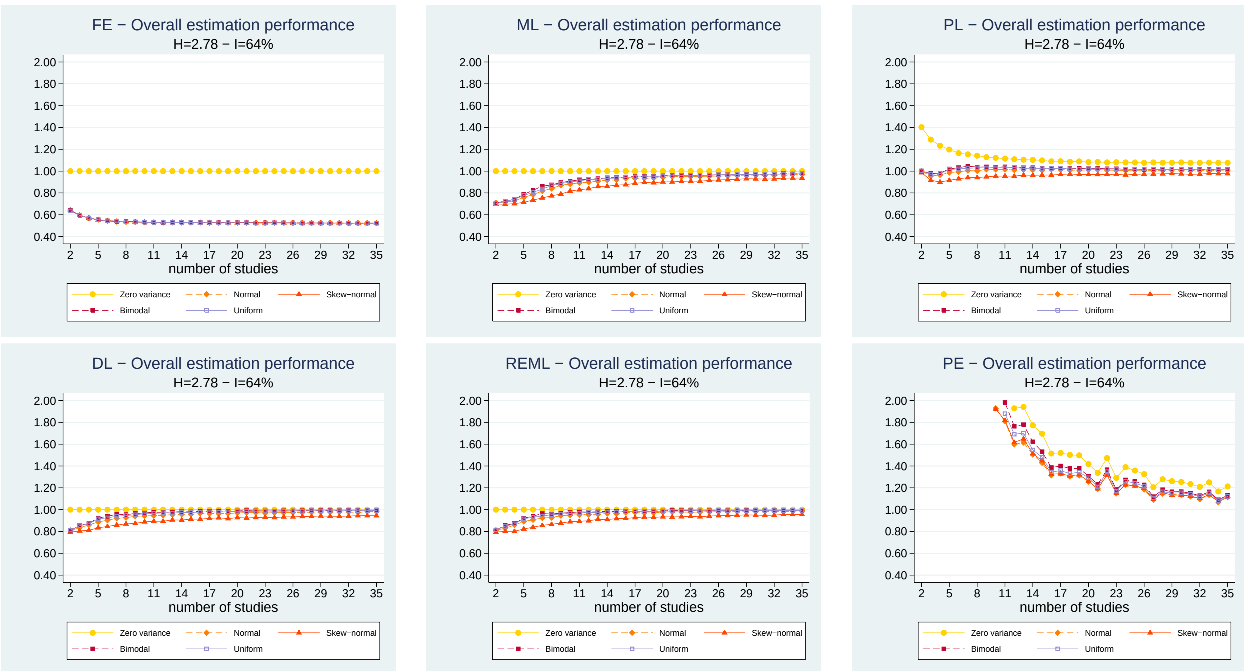
Large heterogeneity across various between-study variance distributions



Meta-analysis overview
A practical guide
Summary
the MetaEasy add-in
metaeff & metaan
Methods and performance
$\hat{\tau}^2 = 0$

CI performance by method

Large heterogeneity across various between-study variance distributions



Kontopantelis, Reeves
Software and model selection challenges in meta-analysis

Which method then?

- Within any given method, the results were consistent across all types of distribution shape.
- Therefore methods are highly robust against even severe violations of the assumption of normality.
- Choose PE if the priority is an accurate Type I error rate (false positive).
- But low power makes it a poor choice when control of the Type II error rate (false negative) is also important and it cannot be used with less than 6 studies.

Which method then?

- For very small study numbers ($\leq$5) only PL gives coverage >90% and an acccurate CI.
- PL has a ‘reasonable’ coverage in most situations, especially for moderate and large heterogeneiry, giving it an edge over other methods.
- REML and DL perform similarly and better than PL only when heterogeneity is low ($I^2 < 15\%$)
- The computational complexity of REML is not justified.

Bring on the champagne?

- Does not necessarily mean homogeneity.
- Most methods use biased estimators and not uncommon to get a negative $\hat{\tau}^2$ which is set to 0 by the model.
- We identified a large percentage of cases where the estimators failed to identify existing heterogeneity.
- In our simulations, for 5 studies and $I^2 \approx 29\%$:
  - 30% of the meta-analyses were erroneously estimated to be homogeneous under the DL method.
  - 32% for REML and 48% for ML-PL.

What does it mean?

- In these cases coverage was substandard and was over 10% lower than in cases where $\hat{\tau}^2 > 0$, on average.
- The problem becomes less profound as the number of studies and the level of heterogeneity increase.
- Better estimators are needed.
- There might be a large number of meta-analyses of ‘homogeneous’ studies which have reached a wrong conclusion.

What to take home

- MetaEasy can help you organise your meta-analysis and can be especially useful if you need to combine continuous and binary outcomes.
- Methods implemented in Stata under metaeff and metaan.
- A zero $\hat{\tau}^2$ is a reason to worry. Heterogeneity might be there but we cannot measure or account for in the model.
- If $\hat{\tau}^2 > 0$, even if very small, use a random-effects model.
- The DL method works reasonably well, under all distributions, especially for low levels of heterogeneity.
- Profile likelihood, which takes into account the uncertaintly in $\hat{\tau}^2$, works better when $I^2 \geq 15\%$.

- Comments, suggestions: e.kontopantelis@manchester.ac.uk

References

- Kontopantelis E, Reeves D.
  MetaEasy: A Meta-Analysis Add-In for Microsoft Excel.
  Journal of Statistical Software, 30(7):1-25, 2009.
- Kontopantelis E, Reeves D.
  metaan: Random-effects meta-analysis.
  The Stata Journal, 10(3):395-407, 2010.
- Kontopantelis E, Reeves D.
  Performance of statistical methods for meta-analysis when true study effects are non-normally distributed: A simulation study.
  Stat Methods Med Res, published online Dec 9 2010.