Models for potentially biased evidence in meta-analysis using empirically based priors

Nicky Welton

Thanks to: Tony Ades, John Carlin, Doug Altman, Jonathan Sterne, Ross Harris

RSS Avon Local Group Meeting, 25th May 2010

Department of Community Based Medicine
Introduction

• Meta-Analysis / Evidence Synthesis
  • Pooling of information from a set of RCT’s comparing the same interventions
  • Pooled relative effect measure (e.g. Odds-Ratio)
  • To summarise a body of evidence (Cochrane)
  • Aid decision-making (Health Technology Assessment / NICE guidance)
Risk of Bias

- Pooled treatment effects and resulting decisions
  - rely on integrity of evidence on which they are based

- Randomised Controlled Trials (RCT’s)
  - considered the gold-standard evidence to inform relative treatment efficacy

- Even RCTs vary in quality
  - Is randomisation allocation adequately concealed?
  - Is there appropriate blinding?
  - Quality information routinely collected (Cochrane risk of bias tool)
Clozapine versus neuroleptic medication for schizophrenia

Concealment inadequate (H)
- Australia 1976: 11/33, 17/31
- Europe 1974: 6/110, 8/113
- Europe 1984: 8/39, 15/40
- Germany 1989: 2/16, 2/16
- Germany 1994: 1/18, 1/18
- Hong Kong 1974: 1/20, 1/20
- Japan 1977: 4/47, 0/41
- Switzerland 1975: 1/18, 2/19
- Taiwan 1997: 6/16, 11/15
- USA 1979a: 4/7, 5/8
- USA 1979b: 27/75, 36/76
- USA 1987: 2/21, 0/20
- USA 1994a: 6/25, 1/14
- USA 1994d: 46/136, 58/89
- USA 1996a: 88/205, 157/218
- USA 1997: 37/215, 49/235
- Subtotal: 215/807, 317/760

Concealment adequate (L)
- Canada 1977: 6/22, 9/28
- Romania 1976: 1/20, 0/20
- USA 1988: 15/126, 18/142
- USA 1994b: 12/37, 21/34
- USA 1996b: 3/10, 1/11
- Subtotal: 37/215, 49/235
- Overall: 252/1022366, 995/995
Should we include evidence at high risk of bias?

- Best available evidence approach
  - ignore evidence at high risk of bias
    … but evidence at low risk of bias may be relatively sparse

- All available evidence approach
  - Somehow combine high & low risk of bias evidence
  - In spirit of NICE, where focus is on decision analysis that reflects body of evidence available
How to estimate and adjust for bias?

- Internally within meta-analysis?
  - High risk evidence contributes mainly to bias estimation, and very little to treatment effect estimates

- Use external evidence as priors
  - Elicitation from experts (Turner at al. JRSSA 2009)
  - Evidence-based from previous meta-analyses (Welton et al. JRSSA 2009)
Meta-epidemiology

(Naylor, *BMJ* 1997; 315: 617-619)

- Identify a large number of meta-analyses
  - “Meta-meta-analysis”
- Record characteristics of individual studies (e.g. adequate allocation concealment or blinding)
- Compare treatment effects *within* each meta-analysis (e.g. not double blind vs. double blind)
- **Ratio of odds ratios** comparing trials at high risk of bias (H) with those at low risk of bias (L)
Allocation concealment: combined evidence

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Ratio of odds ratios (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schulz, 1995</td>
<td>0.66 (0.59, 0.73)</td>
</tr>
<tr>
<td>Moher, 1998</td>
<td>0.63 (0.45, 0.88)</td>
</tr>
<tr>
<td>Kjaergard, 2001</td>
<td>0.60 (0.37, 0.97)</td>
</tr>
<tr>
<td>Balk, 2002</td>
<td>0.95 (0.83, 1.09)</td>
</tr>
<tr>
<td>Egger, 2003</td>
<td>0.79 (0.70, 0.89)</td>
</tr>
<tr>
<td>Als-Neilsen, 2004</td>
<td>1.02 (0.93, 1.13)</td>
</tr>
<tr>
<td>Overall (I-squared = 88.6%, p = 0.000)</td>
<td>0.79 (0.66, 0.95)</td>
</tr>
</tbody>
</table>

NOTE: Weights are from random effects analysis
Estimating bias

- Previous studies have focussed on estimating mean bias
- But …
  - there may be more between trial (within-meta-analysis) heterogeneity in H studies
  - expect mean bias varies between meta-analyses
- Suggests hierarchical model for bias
  - Estimated from meta-epidemiological data
  - Used to inform evidence based priors
Likelihood and treatment effect model

- Binary outcomes, $r_{j,k,m^*}$
  - Meta-analysis $m^*$, study $j$, treatment $k$
  - Likelihood: $r_{j,k,m^*} \sim Bin(p_{j,k,m^*}, n_{j,k,m^*})$

- Logistic regression (bias indicator $C$):

  \[
  \text{logit}(p_{j,k,m^*}) = \begin{cases} 
  \mu_{j,m^*} & \text{Control Arm, } k=0 \\
  \mu_{j,m^*} + \delta_{j,m^*} + \beta_{j,m^*} C_{j,m^*} & \text{Treatment Arm, } k=1 
  \end{cases}
  \]

  Baseline  LogOR  Bias

- Random effects for treatment effects, log(OR)'s:

  $\delta_{j,m^*} \sim N(d_{m^*}, \tau_{m^*}^2)$
The Bias Model

- Study specific bias is exchangeable between studies, within meta-analysis:
  \[ \beta_{jm^*} \sim N(b_{m^*}, \kappa^2) \]

- Meta-analysis specific mean bias is assumed exchangeable between meta-analyses:
  \[ b_{m^*} \sim N(b_0, \phi^2) \]

- Uncertainty in overall mean bias:
  \[ b_0 \sim N(B_0, V_0) \]

- Use meta-epidemiological studies to provide inputs:
  \[ \kappa, \phi, B_0, V_0 \]
Forming Hierarchical Prior from Meta-Epidemiological Data

- Schulz et al (1995): 33 meta-analyses; 250 trials; 79 L and 171 H trials
- Same likelihood & treatment effect model:

\[ r_{j,k,m} \sim Bin(p_{j,k,m}, n_{j,k,m}) \]

\[
\text{logit}(p_{j,k,m}) = \begin{cases} 
\mu_{j,m} & \text{Control Arm, } k=0 \\
\mu_{j,m} + \delta_{j,m} + \beta_{j,m} C_{j,m} & \text{Treatment Arm, } k=1 
\end{cases}
\]

\[ \delta_{j,m} \sim N(d_m, \tau_m^2) \]
Forming Hierarchical Prior from Meta-Epidemiological Data

- Same bias model:
  \[ \beta_{j,m^*} \sim N(b_{m^*}, \kappa^2) \]
  \[ b_{m^*} \sim N(b_0, \varphi^2) \]

- Priors are given to \( b_0, \kappa \) and \( \varphi \)

- The resulting joint posterior for \( b_0, \kappa \) and \( \varphi \) then provides the inputs for the new meta-analysis, \( m^* \)
  - Could sample from the joint posterior to form prior
  - Or simply plug-in posterior summaries
Results from Schulz Analysis

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>95% Credible Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schulz Analysis (Fixed treatment effect; k2 fixed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b_0$</td>
<td>-0.46</td>
<td>0.108</td>
<td>-0.47</td>
<td>(-.66, -.25)</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>0.15</td>
<td>0.106</td>
<td>0.13</td>
<td>(0.01, 0.39)</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>0.11</td>
<td>0.085</td>
<td>0.10</td>
<td>(0.00, 0.30)</td>
</tr>
</tbody>
</table>

$\hat{\kappa} = 0.13 \quad \hat{\varphi} = 0.10$

- Ideally use joint posterior distribution for $\kappa$ and $\varphi$
- We found results robust to simply plugging in posterior medians for $\kappa$ and $\varphi$
## Results: Clozapine

<table>
<thead>
<tr>
<th>Data/Model</th>
<th>Mean (95% Credible Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adequate Concealment, Face Value</td>
<td>-0.065 (-1.68, 2.84)</td>
</tr>
<tr>
<td>Inadequate/Unclear Concealment, Face Value</td>
<td>-0.533 (-1.03, 0.13)</td>
</tr>
<tr>
<td>All studies, Face Value</td>
<td>-0.452 (-0.88, 0.08)</td>
</tr>
<tr>
<td>All studies, Bias Adjusted</td>
<td>-0.149 (-0.61, 0.43)</td>
</tr>
</tbody>
</table>
What have we assumed?

- Bias is exchangeable across trials within a meta-analysis (OK?)
- Mean bias is exchangeable across meta-analyses (BIG assumption)
  - More realistic if restrict to meta-analyses in similar clinical areas
  - … but then this reduces size of evidence base available to inform hierarchical prior
Comments

• The informational content of high risk trials is limited
  • Even large trials downweighted by $V_0 + \kappa + \varphi$
  • Even if infinitely many trials posterior variance depends on $V_0 + \varphi$

• Increasing no. of meta-analyses to form prior:
  • Potentially can reduce $V_0$, but not $\kappa$ or $\varphi$

• Reducing variety of meta-analyses may reduce $\varphi$
  • But only at the expense of increasing $V_0$ …
Posterior sd for treatment effect

No H trials, L trials only
Single large H trial
10 typical H trials
Infinite no. of H trials

\[ \sqrt{V_0 + \varphi^2} \]
Informational content of high risk studies

(a) Single large H trial
(b) 10 typical H trials
(c) Infinite no. of H trials

× Clozapine, Random Effects model
+ Clozapine, Fixed Effect model
Posterior sd of treatment effect

\[ \kappa = 0 \]

\[ \kappa = 0.62 \]

\[ \sqrt{V_0 + \varphi^2} \]
Consequences for Decision Modelling

- Decisions made by decision-makers such as NICE need to be accepted by patient groups, pharmaceutical industry …

- Down-weighting evidence may lead to appeal
  - if dependent on choice of model
  - if dependent on inclusion criterion for evidence-based prior

- Assessment of model fit & sensitivity analysis to model inputs crucial if decisions based on these models are to have credence in practice
BRANDO (Bias in Randomised and Observational Studies)

- Previous meta-epidemiological studies produced conflicting results
- Combine data from all existing empirical studies into a single database
  - Seven studies contributed data on both trial characteristics and intervention effects
- Final database contains data on 2572 trials
- Restricted to meta-analyses where it was clear in which direction the bias acts
Sensitivity to priors

• Extreme sensitivity to priors for variance parameters
  • Although mean bias estimates robust
• Simulation exercise
  • Using typical study results from BRANDO
• Most priors performed badly
• Inverse-Gamma priors for variance parameters (e.g. IG(.001,.001)) performed best
Summary

• Evidence that poor methodological quality introduces bias

• In pairwise meta-analysis
  • can adjust for and down-weight studies using external evidence

• Sensitivity analyses important

• Assumes exchangeability within and between meta-analyses