An Introduction to Network Meta-analysis

Gordon Guyatt
Department of Health Research
Methods, Evidence, and Impact
McMaster University
ATTENDANCE SLIDE

ETA: 2NKW

June 8, 2017
Conflicts of interest

- Co-Chair of GRADE working group
- Written about NMA including Users Guide
Network Meta-analysis

- What is a network meta-analysis?
- How does NMA generate effect estimates?
- Determining credibility of NMA
Many disease areas have many alternatives exist. Clinicians/patients need to know the relative merits. It is impractical to test each comparator directly. Simultaneous comparison multiple treatments is possible through "network meta-analysis", "mixed treatment comparisons", and "adjusted indirect comparisons".
Conventional meta-analysis
Pooled Estimate assumption

Direct comparison A versus B
Single best estimate of treatment effect

Assumes effect similar across

 Patients    Interventions    Outcomes    Methodology

“Homogeneity assumption”

In GRADE: Unexplained heterogeneity lowers evidence quality
Indirect Comparisons

Interested in A versus B available data A vs C, B vs C

- Alendronate (A)
  - Odds ratio 0.5

- Risedronate (B)
  - Odds ratio 1.0

- Placebo (C)
  - Odds ratio 0.5

⇒ Relative confidence in indirect vs direct? ⇒ Why?
Vulnerability of Indirect comparison

- Effect modifiers
  - Patients
  - Optimal interventions
  - Comparator
  - Cointerventions
  - Outcome measures
  - Risk of bias

- Important differences - intransitivity
Combine direct and indirect comparisons
- Additional assumption direct and indirect similar
- Coherence or consistency assumption
**Number of smokers failing to quit at 12 months, according to treatment group**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>No of trials</th>
<th>Odds ratio (95% CI)</th>
<th>I² (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bupropion v placebo</td>
<td>9</td>
<td>0.51 (0.36 to 0.73)</td>
<td>54</td>
</tr>
<tr>
<td>NRT patch v placebo</td>
<td>19</td>
<td>0.57 (0.48 to 0.67)</td>
<td>12</td>
</tr>
<tr>
<td>Bupropion v NRT patch:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct comparison</td>
<td>1</td>
<td>0.48 (0.28 to 0.82)</td>
<td>—</td>
</tr>
<tr>
<td>Adjusted indirect comparison</td>
<td>9+19</td>
<td>0.90 (0.61 to 1.34)</td>
<td>—</td>
</tr>
<tr>
<td>Combined (direct + indirect)</td>
<td>1+(9+19)</td>
<td>0.68 (0.37 to 1.25)</td>
<td>71</td>
</tr>
</tbody>
</table>
Network Meta-analysis Case Study: Which Approach to Nicotine Addiction Works Best

Comparison with Treatments

- **NRT**
  - Odds ratio: 1.54 (1.02-2.31)
  - Comparison with NRT + NRT
  - Comparison with NRT + antidepressant
  - Comparison with Varenicline
- **Antidepressants**
  - Odds ratio: 1.62 (1.00-2.63)
  - Comparison with NRT + NRT
  - Comparison with NRT + antidepressant
  - Comparison with Varenicline
- **Varenicline**
  - Odds ratio: 1.34 (0.71-2.56)
- **NRT + antidepressant**
  - Odds ratio: 1.70 (1.41-2.04)
  - Comparison with NRT + NRT
  - Comparison with NRT + antidepressant
  - Comparison with Varenicline
Combines effect estimates from direct and indirect comparisons

- Placebo
  - Nicotine replacement treatment (NRT)
  - Varenicline
  - Antidepressants + NRT
  - Antidepressants
Comparison with Treatments

NRT

Antidepressants

Varenicline

NRT + antidepressant

Comparison with Treatments

NRT

Antidepressants

Varenicline

NRT + antidepressant

Odds ratio
Direct Comparison

**Antidepressants**

**NRT**

- NRT+NRT: 1.63, $I^2=0\%$
  - 4 comparisons
- buspirone: 1.36, $I^2=0\%$
  - 2 comparisons
- rimonabant: 0.73, $I^2=0\%$
  - 2 comparisons
- clonidine: 2.68, $I^2=82\%$
  - 5 comparisons
- varenicline: 1.28, $I^2=0\%$
  - 3 comparisons
- 1.28
  - 1 comparison
- 4.85
  - 1 comparison
- 1.14, $I^2=63\%$
  - 6 comparisons
- 1.54, $I^2=46\%$
  - 5 comparisons
- 1.70, $I^2=0\%$
  - 3 comparisons
- 1.12
  - 1 comparison
- Direct evidence (3 trials)
  - 1.34 (0.71, 2.56)
    - I-squared=43.7%

**NRT superior**

**Antidepressants superior**

1.85, $I^2=13\%$
- 67 comparisons
1.88, $I^2=19\%$
- 29 comparisons
Indirect Comparison 1: First order loop

**Indirect evidence**

1.01 (0.81, 1.27)

**I-squared = 43.7%**

**Direct evidence (3 trials)**

1.34 (0.71, 2.56)

I-squared = 43.7%
Indirect Comparison 2: Second order loop

- **Indirect evidence**
  - 0.85 (0.38, 1.92)
  - I-squared = 43.7%

- **Direct evidence (3 trials)**
  - 1.34 (0.71, 2.56)
  - I-squared = 43.7%

- **Comparisons**
  - 1.63, $I^2=0%$
    - 6 comparisons
  - 1.63, $I^2=0%$
    - 4 comparisons
  - 1.88, $I^2=19%$
    - 29 comparisons
  - 2.68, $I^2=82%$
    - 5 comparisons
  - 1.85, $I^2=13%$
    - 67 comparisons
  - 1.70, $I^2=0%$
    - 3 comparisons
  - 1.14, $I^2=63%$
    - 6 comparisons
  - 1.36, $I^2=0%$
    - 2 comparisons
  - 0.73, $I^2=0%$
    - 2 comparisons
  - 1.54, $I^2=46%$
    - 5 comparisons
  - 1.28
    - 1 comparison
  - 1.14, $I^2=63%$
    - 6 comparisons
  - 1.12
    - 1 comparison
  - 1.88, $I^2=19%$
    - 29 comparisons
  - 1.54, $I^2=46%$
    - 5 comparisons
5 Paths to Indirectly Compare Antidepressants vs NRT

1. NRT + nonplacebo control
2. placebo and nonplacebo control
3. antidepressants + NRT
4. antidepressants
5. varenicline

Comparisons:
- NRT + NRT: 29 comparisons
- buspirone: 2 comparisons
- rimonabant: 4 comparisons
- clonidine: 6 comparisons
- antidepressants: 1 comparison
- varenicline: 3 comparisons
5 Paths to Indirectly Compare Antidepressants vs NRT

1. 1.01 (0.81, 1.27)
2. 0.85 (0.38, 1.92)
3. 0.89 (0.29, 2.77)
4. 1.56 (0.54, 4.49)
5. 1.31 (0.25, 6.76)
Comparative effectiveness of NRT vs. Antidepressants on prolonged abstinence

**Indirect evidence**

<table>
<thead>
<tr>
<th>Path</th>
<th>NRT superior</th>
<th>Antidepressants superior</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.70 (I²=0%)</td>
<td>1.01 (0.81, 1.27)</td>
</tr>
<tr>
<td>2</td>
<td>1.36 (I²=0%)</td>
<td>0.85 (0.38, 1.92)</td>
</tr>
<tr>
<td>3</td>
<td>0.73 (I²=0%)</td>
<td>0.89 (0.29, 2.77)</td>
</tr>
<tr>
<td>4</td>
<td>2.68 (I²=82%)</td>
<td>1.56 (0.54, 4.49)</td>
</tr>
<tr>
<td>5</td>
<td>1.28 (I²=0%)</td>
<td>1.31 (0.25, 6.76)</td>
</tr>
</tbody>
</table>

**Direct evidence**

- 3 trials pooled: 1.34 (0.71, 2.56) I-squared = 43.7%
- NRT superior: 1.14 (I²=63%) 6 comparisons
- Antidepressants superior: 1.54 (I²=46%) 5 comparisons

**I-squared** = 43.7%
Comparative effectiveness of NRT vs. Antidepressants on prolonged abstinence

Indirect evidence

<table>
<thead>
<tr>
<th>Path</th>
<th>NRT superior</th>
<th>Antidepressants superior</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.01 (0.81, 1.27)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.85 (0.38, 1.92)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.89 (0.29, 2.77)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.56 (0.54, 4.49)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.31 (0.25, 6.76)</td>
<td></td>
</tr>
</tbody>
</table>

Direct evidence

3 trials pooled

1.34 (0.71, 2.56)
I-squared = 43.7%

0.98
(95% 0.85-1.13)
Comparative effectiveness of NRT vs. Antidepressants on prolonged abstinence

Direct evidence

3 trials pooled

1.34 (0.71, 2.56)
I-squared = 43.7%

Indirect evidence

pooled estimate
1.01
(95% 0.88-1.15)

NRT superior
Antidepressants superior

<table>
<thead>
<tr>
<th>Path</th>
<th>NRT superior</th>
<th>Antidepressants superior</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.01 (0.81, 1.27)</td>
<td>0.98 (0.85-1.13)</td>
</tr>
<tr>
<td>2</td>
<td>0.85 (0.38, 1.92)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0.89 (0.29, 2.77)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1.56 (0.54, 4.49)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1.31 (0.25, 6.76)</td>
<td></td>
</tr>
</tbody>
</table>

I-squared = 43.7%
Presentation of Results

- Often presented with rankings
- Potentially very misleading
  - Small difference between ranks
  - Everything low or very low quality
  - First ranked lower quality than others
Comparative Effectiveness of Drug Treatments to Prevent Fragility Fractures: A Systematic Review and Network Meta-Analysis

Data Synthesis: This network meta-analysis included 116 trials (139,647 patients; median age, 64 yr; 86% females and 88% Caucasians; median follow-up, 24 months). Trials were at low to moderate risk of bias. Teriparatide had the highest risk reduction of fractures (odds ratios, 0.42, 0.30, and 0.50 for hip, vertebral, and nonvertebral fractures, respectively) and the highest probability of being ranked first for efficacy (probabilities of 42, 49, and 79% for hip, vertebral, and nonvertebral fractures, respectively). However, differences to denosumab, zoledronate, risedronate, ibandronate, and alendronate were not statistically significant. Raloxifene and bazedoxifene were likely less effective, although these data were limited. Calcium and vitamin D were ineffective given separately but reduced the risk of hip fractures if given in combination (odds ratio, 0.81; 95% confidence interval, 0.68–0.96).
<table>
<thead>
<tr>
<th>Comparison</th>
<th>Direct evidence OR (95% confidence interval)</th>
<th>Direct evidence confidence in estimates</th>
<th>Indirect evidence OR (95% credible interval)</th>
<th>Indirect evidence confidence in estimates</th>
<th>Network OR (95% credible interval)</th>
<th>Network confidence in estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teriparatide vs. Placebo</td>
<td>---</td>
<td>---</td>
<td>0.42 (0.10-1.82)</td>
<td>very low 3, 6</td>
<td>0.42 (0.10-1.82)</td>
<td>very low</td>
</tr>
<tr>
<td>Zoledronate vs. Placebo</td>
<td>---</td>
<td>---</td>
<td>0.50 (0.33-0.74)</td>
<td>high</td>
<td>0.50 (0.34-0.73)</td>
<td>high</td>
</tr>
<tr>
<td>Risedronate vs. Placebo</td>
<td>0.17 (0.05 to 0.59)</td>
<td>moderate 1</td>
<td>0.54 (0.36-0.75)</td>
<td>low 6</td>
<td>0.48 (0.31-0.66)</td>
<td>moderate</td>
</tr>
</tbody>
</table>
Users’ Guides for Credibility of the Systematic Review Process

Did the review explicitly address a sensible clinical question?
Was the search for relevant studies exhaustive?
Was the risk of bias of the primary studies assessed?
Did the review address possible explanations of between-study differences in results?
Did the review present results that are ready for clinical application?
Were selection and assessments of studies reproducible?
Did the review address confidence in effect estimates?
Analysis of the systematic reviews process in reports of network meta-analyses: methodological systematic review

Conclusions Essential methodological components of the systematic review process—conducting a literature search and assessing risk of bias of individual studies—are frequently lacking in reports of network meta-analyses, even when published in journals with high impact factors.

Very few addressed certainty in effects
Certainty differs across outcomes
# Certainty assessment criteria

<table>
<thead>
<tr>
<th>Study Design</th>
<th>Confidence in estimates</th>
<th>Lower if</th>
<th>Higher if</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomized trials</td>
<td>High</td>
<td>Risk of bias</td>
<td>Large Effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 Serious</td>
<td>+1 Large</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2 Very serious</td>
<td>+1 Very large</td>
</tr>
<tr>
<td></td>
<td>Moderate</td>
<td>Inconsistency</td>
<td>Dose response</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 Serious</td>
<td>+1 Evidence of a gradient</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2 Very serious</td>
<td>All plausible confounding</td>
</tr>
<tr>
<td>Observational studies</td>
<td>Low</td>
<td>Indirectness</td>
<td>+1 Would reduce a demonstrated effect or</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 Serious</td>
<td>+1 would suggest a spurious effect when results show no effect</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2 Very serious</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Very Low</td>
<td>Imprecision</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 Serious</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2 Very serious</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Publication bias</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 Likely</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-2 Very likely</td>
<td></td>
</tr>
</tbody>
</table>
Indirect Comparisons

Interested in A versus B
available data A vs C, B vs C

Intervention A

Moderate Quality

Placebo (C)

Intervention B

Low Quality
Combining direct and indirect

Interested in A versus B
Direct and Indirect available

- **Alendronate (A)**
  - Moderate Quality

- **Risedronate (B)**
  - High Quality

- **Placebo (C)**
  - Low Quality
Combine direct and indirect comparisons
- Additional assumption direct and indirect similar
- Coherence or consistency assumption
## Incoherence

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Direct evidence OR (95% confidence interval)</th>
<th>Direct evidence confidence in estimates</th>
<th>Indirect evidence OR (95% credible interval)</th>
<th>Indirect evidence confidence in estimates</th>
<th>Network OR (95% credible interval)</th>
<th>Network confidence in estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unreamed versus reamed</td>
<td>0.74 (0.45 – 1.24)</td>
<td>MODERATE</td>
<td>0.07 (0.01 – 0.46)</td>
<td>LOW</td>
<td>0.62 (0.37 – 1.03)</td>
<td>LOW</td>
</tr>
<tr>
<td>Unreamed versus external fixation</td>
<td>0.39 (0.23 – 0.65)</td>
<td>MODERATE</td>
<td>0.35 (0.08 – 1.56)</td>
<td>LOW</td>
<td>0.38 (0.23 – 0.62)</td>
<td>MODERATE</td>
</tr>
</tbody>
</table>

P-value for test of incoherence 0.02
Resuscitation in Sepsis: Colloids versus Crystalloids

Appendix Figure III. Forrest plot for mortality in direct comparisons of all crystalloids versus all colloids

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Colloids Events</th>
<th>Colloids Total</th>
<th>Crystalloids Events</th>
<th>Crystalloids Total</th>
<th>Weight</th>
<th>Odds Ratio M-H, Fixed, 95% CI</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haupt 1982</td>
<td>8</td>
<td>13</td>
<td>3</td>
<td>4</td>
<td>0.2%</td>
<td>0.53 [0.04, 6.65]</td>
<td>1982</td>
</tr>
<tr>
<td>Finfer 2004</td>
<td>185</td>
<td>603</td>
<td>217</td>
<td>615</td>
<td>20.4%</td>
<td>0.81 [0.64, 1.03]</td>
<td>2004</td>
</tr>
<tr>
<td>Brunhorst 2008</td>
<td>107</td>
<td>261</td>
<td>93</td>
<td>274</td>
<td>7.3%</td>
<td>1.35 [0.95, 1.92]</td>
<td>2008</td>
</tr>
<tr>
<td>Li 2008</td>
<td>14</td>
<td>30</td>
<td>20</td>
<td>30</td>
<td>1.5%</td>
<td>0.44 [0.15, 1.24]</td>
<td>2008</td>
</tr>
<tr>
<td>McIntyre 2008</td>
<td>9</td>
<td>21</td>
<td>7</td>
<td>19</td>
<td>0.6%</td>
<td>1.29 [0.36, 4.58]</td>
<td>2008</td>
</tr>
<tr>
<td>Dubin 2010</td>
<td>3</td>
<td>12</td>
<td>7</td>
<td>13</td>
<td>0.7%</td>
<td>0.29 [0.05, 1.57]</td>
<td>2010</td>
</tr>
<tr>
<td>Myburgh 2012</td>
<td>248</td>
<td>976</td>
<td>224</td>
<td>945</td>
<td>23.3%</td>
<td>1.10 [0.89, 1.35]</td>
<td>2012</td>
</tr>
<tr>
<td>Lu 2012</td>
<td>7</td>
<td>22</td>
<td>12</td>
<td>20</td>
<td>1.2%</td>
<td>0.31 [0.09, 1.10]</td>
<td>2012</td>
</tr>
<tr>
<td>Guidet 2012</td>
<td>40</td>
<td>99</td>
<td>32</td>
<td>95</td>
<td>2.7%</td>
<td>1.33 [0.74, 2.40]</td>
<td>2012</td>
</tr>
<tr>
<td>Pemer 2012</td>
<td>202</td>
<td>398</td>
<td>173</td>
<td>400</td>
<td>11.6%</td>
<td>1.35 [1.02, 1.79]</td>
<td>2012</td>
</tr>
<tr>
<td>Siegmund 2012</td>
<td>44</td>
<td>117</td>
<td>50</td>
<td>124</td>
<td>4.2%</td>
<td>0.89 [0.53, 1.50]</td>
<td>2012</td>
</tr>
<tr>
<td>CRYSTAL 2013</td>
<td>252</td>
<td>774</td>
<td>286</td>
<td>779</td>
<td>26.4%</td>
<td>0.83 [0.67, 1.03]</td>
<td>2013</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>3326</strong></td>
<td><strong>3318</strong></td>
<td><strong>100.0%</strong></td>
<td></td>
<td><strong>0.99 [0.89, 1.10]</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total events</strong></td>
<td><strong>1119</strong></td>
<td><strong>1124</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Heterogeneity:</strong> Chi² = 23.20, df = 11 (P = 0.02); I² = 53%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Test for overall effect:</strong> Z = 0.18 (P = 0.85)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fluid Resuscitation in Sepsis
A Systematic Review and Network Meta-analysis

Bram Rochwer, MD; Waleed Alhazzani, MD; Anees Sindi, MD; Diane Heels-Ansdell, MSc; Lehana Thabane, PhD; Alison Fox-Robichaud, MD; Lawrence Mbuagbaw, MSc; Wojciech Szczeklik, MD; Fayez Alshamsi, MD; Sultan Altayyar, MD; Wang-Chun Ip, MD; Guowei Li, MSc; Michael Wang, MD; Anna Wludarczyk, MD; Qi Zhou, PhD; Gordon H. Guyatt, MD; Deborah J. Cook, MD; Roman Jaeschke, MD; and Djillali Annane, MD, PhD, for the Fluids in Sepsis and Septic Shock Group

Ann Intern Med. doi:10.7326/M14-0178
www.annals.org

For author affiliations, see end of text.

This article was published online first at www.annals.org on 22 July 2014.
Fluid resuscitation in sepsis

- Sensible clinical question
  - RCTs, sepsis requiring fluids, mortality 90 days
- Comprehensive search
  - Multiple data bases, reviewed citations
- Duplicate eligibility, risk of bias, data abstraction
Table 4. NMA Results of 6-Node Analysis, Including Confidence Assessments

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Trials With Direct Comparisons, n</th>
<th>Direct Estimate (95% CI); Quality of Evidence</th>
<th>Indirect Estimate (95% CrI); Quality of Evidence</th>
<th>NMA Estimate (95% CrI)*; Quality of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-HES vs. saline</td>
<td>4</td>
<td>1.07 (0.89–1.29); Moderate†</td>
<td>0.59 (0.25–1.35); Very low ‡§</td>
<td>1.04 (0.87–1.25); Moderate†</td>
</tr>
<tr>
<td>H-HES vs. saline</td>
<td>3</td>
<td>0.64 (0.30–1.37); Moderate†</td>
<td>1.13 (0.71–1.80); Very low ‡‡</td>
<td>0.95 (0.64–1.41); Moderate‡</td>
</tr>
<tr>
<td>Albumin vs. saline</td>
<td>2</td>
<td>0.81 (0.64–1.03); Moderate†</td>
<td>0.96 (0.14–6.31); Very low ‡</td>
<td>0.82 (0.65–1.04); Moderate</td>
</tr>
<tr>
<td>Balanced crystalloid vs. saline</td>
<td>0</td>
<td>–</td>
<td>0.78 (0.58–1.05); Low ‡</td>
<td>0.78 (0.58–1.05); Low</td>
</tr>
<tr>
<td>Gelatin vs. saline</td>
<td>0</td>
<td>–</td>
<td>1.04 (0.46–2.32); Very low ‡‡</td>
<td>1.04 (0.46–2.32); Very low</td>
</tr>
<tr>
<td>H-HES vs. L-HES</td>
<td>0</td>
<td>–</td>
<td>0.91 (0.63–1.33); Low ‡</td>
<td>0.91 (0.63–1.33); Low</td>
</tr>
<tr>
<td>Albumin vs. L-HES</td>
<td>0</td>
<td>–</td>
<td>0.79 (0.59–1.06); Low ‡</td>
<td>0.79 (0.59–1.06); Low</td>
</tr>
<tr>
<td>Balanced crystalloid vs. L-HES</td>
<td>2</td>
<td>0.80 (0.61–1.04); Moderate§</td>
<td>0.44 (0.19–0.97); Moderate‡</td>
<td>0.75 (0.58–0.97); Moderate</td>
</tr>
<tr>
<td>Gelatin vs. L-HES</td>
<td>0</td>
<td>–</td>
<td>1.00 (0.44–2.21); Very low ‡‡</td>
<td>1.00 (0.44–2.21); Very low</td>
</tr>
<tr>
<td>Albumin vs. H-HES</td>
<td>2</td>
<td>1.40 (0.35–5.56); Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balanced crystalloid vs. H-HES</td>
<td>1</td>
<td>0.74 (0.52–1.05); Moderate†</td>
<td>1.35 (0.63–2.92); Very low ‡</td>
<td></td>
</tr>
<tr>
<td>Gelatin vs. H-HES</td>
<td>1</td>
<td>1.09 (0.55–2.19); Low</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balanced crystalloid vs. albumin</td>
<td>0</td>
<td>–</td>
<td>0.95 (0.65–1.38); Very low ‡‡</td>
<td>0.95 (0.65–1.38); Very low</td>
</tr>
<tr>
<td>Gelatin vs. albumin</td>
<td>0</td>
<td>–</td>
<td>1.26 (0.55–2.90); Very low ‡‡</td>
<td>1.26 (0.55–2.90); Very low</td>
</tr>
<tr>
<td>Gelatin vs. balanced crystalloid</td>
<td>0</td>
<td>–</td>
<td>1.34 (0.61–2.89); Very low ‡‡</td>
<td>1.34 (0.61–2.89); Very low</td>
</tr>
</tbody>
</table>

CrI = credibility interval; H-HES = high-molecular-weight hydroxyethyl starch; L-HES = low-molecular-weight hydroxyethyl starch; NMA = network meta-analysis.

* Higher of direct or indirect confidence.
† Rated down for imprecision.
‡ Rated down for indirectness.
§ Rated down for inconsistency ($I^2 = 80\%$; $P = 0.03$ for heterogeneity).
|| Rated down 2 levels for imprecision.
<table>
<thead>
<tr>
<th>Comparison</th>
<th>Trials With Direct Comparisons, n</th>
<th>Direct Estimate (95% CI); Quality of Evidence</th>
<th>Indirect Estimate (95% CrI); Quality of Evidence</th>
<th>NMA Estimate (95% CrI)*; Quality of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-HES vs. saline</td>
<td>4</td>
<td>1.07 (0.89–1.29); Moderate†</td>
<td>0.59 (0.25–1.35); Very low†††</td>
<td>1.04 (0.87–1.25); Moderate†††</td>
</tr>
<tr>
<td>H-HES vs. saline</td>
<td>3</td>
<td>0.64 (0.30–1.37); Moderate†</td>
<td>1.13 (0.71–1.80); Very low†††</td>
<td>0.95 (0.64–1.41); Moderate†††</td>
</tr>
<tr>
<td>Albumin vs. saline</td>
<td>2</td>
<td>0.81 (0.64–1.03); Moderate†</td>
<td>0.96 (0.14–6.31); Very low††</td>
<td>0.82 (0.65–1.04); Moderate††</td>
</tr>
<tr>
<td>Balanced crystalloid vs. saline</td>
<td>0</td>
<td>–</td>
<td>0.78 (0.58–1.05); Low</td>
<td>0.78 (0.58–1.05); Low</td>
</tr>
<tr>
<td>Gelatin vs. saline</td>
<td>0</td>
<td>–</td>
<td>1.04 (0.46–2.32); Very low†††</td>
<td>1.04 (0.46–2.32); Very low†††</td>
</tr>
<tr>
<td>H-HES vs. L-HES</td>
<td>0</td>
<td>–</td>
<td>0.91 (0.63–1.33); Low</td>
<td>0.91 (0.63–1.33); Low</td>
</tr>
<tr>
<td>Albumin vs. L-HES</td>
<td>0</td>
<td>–</td>
<td>0.79 (0.59–1.06); Low</td>
<td>0.79 (0.59–1.06); Low</td>
</tr>
<tr>
<td>Balanced crystalloid vs. L-HES</td>
<td>2</td>
<td>0.80 (0.61–1.04); Moderate§</td>
<td>0.44 (0.19–0.97); Moderate†</td>
<td>0.75 (0.58–0.97); Moderate†</td>
</tr>
<tr>
<td>Gelatin vs. L-HES</td>
<td>0</td>
<td>–</td>
<td>1.00 (0.44–2.21); Very low†††</td>
<td>1.00 (0.44–2.21); Very low†††</td>
</tr>
<tr>
<td>Albumin vs. H-HES</td>
<td>2</td>
<td>1.40 (0.35–5.56); Low‖</td>
<td>0.83 (0.52–1.33); Low</td>
<td>0.87 (0.55–1.36); Low</td>
</tr>
<tr>
<td>Balanced crystalloid vs. H-HES</td>
<td>1</td>
<td>0.74 (0.52–1.05); Moderate†</td>
<td>1.35 (0.63–2.92); Very low††</td>
<td>0.82 (0.60–1.13); Moderate†</td>
</tr>
<tr>
<td>Gelatin vs. H-HES</td>
<td>1</td>
<td>1.09 (0.55–2.19); Low‖</td>
<td>–</td>
<td>1.10 (0.54–2.21); Low</td>
</tr>
<tr>
<td>Balanced crystalloid vs. albumin</td>
<td>0</td>
<td>–</td>
<td>0.95 (0.65–1.38); Very low†††</td>
<td>0.95 (0.65–1.38); Very low†††</td>
</tr>
<tr>
<td>Gelatin vs. albumin</td>
<td>0</td>
<td>–</td>
<td>1.26 (0.55–2.90); Very low††</td>
<td>1.26 (0.55–2.90); Very low††</td>
</tr>
<tr>
<td>Gelatin vs. balanced crystalloid</td>
<td>0</td>
<td>–</td>
<td>1.34 (0.61–2.89); Very low††</td>
<td>1.34 (0.61–2.89); Very low††</td>
</tr>
</tbody>
</table>
The results of this NMA highlight potentially important differences in mortality between crystalloid solutions. Our findings suggest an advantage of balanced crystalloids versus saline (low confidence) and low- or high-molecular-weight starch (moderate confidence), with similar mortality results for balanced crystalloids and albumin (very low confidence) (Table 4). These differences were not detectable.

These results raise concerns about whether using mostly unbalanced crystalloids in the acute resuscitation of patients with sepsis is optimal. Our findings may partially explain why the effects of different colloid agents vary, with albumin seeming equivalent or superior to all alternatives. Starches, regardless of molecular weight, seem inferior to alternative resuscitation fluids (that is, albumin and balanced crystalloids). Data on gelatin are markedly less robust. Gelatin is associated with increased mortality relative to some other resuscitation fluids; however, only 1 small trial contributes directly to this analysis. The hetero-
Conclusions

- NMA is here to stay
  - That’s a good thing
  - Combine direct and indirect
    - Multiple inferences, stronger inferences
- Need to consider issues for conventional
- Intransitivity and incoherence
  - Beware of rankings – need to consider quality
- Still evolving – rapid progress