Identifying Effective Components of Complex Interventions: Component Network Meta-Analysis (I)

Deborah M. Caldwell, Nicky J. Welton
Population Health Sciences, Bristol Medical School, University of Bristol

Cochrane Webinar, 22nd April 2020
Outline

• Two-part talk: (I) concepts and (II) methods

• Concepts:
  • What are complex interventions?
  • Complexity and evidence synthesis
  • Intervention level network meta-analysis
  • What are components and why focus on them?
Outline

• Two-part talk: (I) concepts and (II) methods

• Methods:
  • Component Network Meta-Analysis Models
    • Common effect (“lumped” MA)
    • Additive component effects
    • Two-way interaction models
    • Full interaction models (“split” NMA)
  • Illustrative examples
What are complex interventions?

- Cochrane handbook (Ch17) refers to “intervention complexity”, rather than “complex intervention”
  i. the number of components in the intervention;
  ii. interactions between intervention components and/or interactions between the intervention and its context; and
  iii. the wider system within which the intervention is introduced.
MRC definition of complexity (interventions)

- A number of interacting components within the experimental and control interventions,
- A number and difficulty of behaviours required by those delivering or receiving the intervention,
- A number of groups or organisational levels targeted by the intervention,
- A number and variability of outcomes,
- A degree of flexibility or tailoring of the intervention permitted.

Craig et al 2008, BMJ; 337 doi: https://doi.org/10.1136/bmj.a1655
MRC definition of complexity (interventions)

- A number of interacting components within the experimental and control interventions,
- A number and difficulty of behaviours required by those delivering or receiving the intervention,
- A number of groups or organisational levels targeted by the intervention,
- A number and variability of outcomes,
- A degree of flexibility or tailoring of the intervention permitted.
**Intervention complexity and evidence synthesis**

- Systematic review of well-conducted RCTs provides *highest quality* evidence for evaluating intervention effectiveness

- Three (main) options for synthesis are
  
  i. non-quantitative synthesis (tabulation, narrative, graphical approaches)

  ii. standard meta-analysis methods (pairwise, fixed, random effects with meta-regression)

  iii. complex synthesis methods (NMA, MPES, MVMA)

Case study: An illustrative dataset

- Subset of studies from a 2004 Cochrane review examining psychological therapies for reducing depressive symptoms post-coronary heart disease.
  - inclusion criteria parallel group RCT, at least 6-months follow-up, and report at least one of the following outcomes: all cause mortality, cardiac mortality, non-fatal MI, total cholesterol, systolic or diastolic blood pressure, depression or anxiety

- Psychological intervention vs control (TAU)
- Depression symptoms, 11 studies
**Pairwise, random-effects meta-analysis**

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Std. Mean Difference IV, Random, 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>-0.60 [-1.12, -0.09]</td>
</tr>
<tr>
<td>Burgess</td>
<td>0.40 [0.06, 0.74]</td>
</tr>
<tr>
<td>Elderen</td>
<td>-0.21 [-0.75, 0.34]</td>
</tr>
<tr>
<td>ENRICHID</td>
<td>-0.33 [-0.43, -0.24]</td>
</tr>
<tr>
<td>HofmanBang</td>
<td>-0.39 [-0.88, 0.10]</td>
</tr>
<tr>
<td>Johnston</td>
<td>-0.57 [-1.09, -0.06]</td>
</tr>
<tr>
<td>Jones</td>
<td>-0.01 [-0.10, 0.07]</td>
</tr>
<tr>
<td>Lewin</td>
<td>-0.61 [-0.99, -0.23]</td>
</tr>
<tr>
<td>Stern</td>
<td>-0.30 [-0.83, 0.23]</td>
</tr>
<tr>
<td>Thompson</td>
<td>-0.39 [-0.92, 0.14]</td>
</tr>
<tr>
<td>Toobert</td>
<td>-0.91 [-1.75, -0.08]</td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>-0.29 [-0.48, -0.10]</strong></td>
</tr>
</tbody>
</table>

Heterogeneity: $\tau^2 = 0.06$; $\chi^2 = 50.32$, df = 10 ($P < 0.00001$); $I^2 = 80\%$

Test for overall effect: $Z = 2.98$ ($P = 0.003$)
Pairwise, random-effects meta-analysis

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Std. Mean Difference</th>
<th>Std. Mean Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IV, Random, 95% CI</td>
<td>IV, Random, 95% CI</td>
</tr>
<tr>
<td>Black</td>
<td>-0.60 [-1.12, -0.09]</td>
<td></td>
</tr>
<tr>
<td>Burgess</td>
<td>0.40 [0.06, 0.74]</td>
<td></td>
</tr>
<tr>
<td>Elderen</td>
<td>-0.21 [-0.75, 0.34]</td>
<td></td>
</tr>
<tr>
<td>ENRICHD</td>
<td>-0.33 [-0.43, -0.24]</td>
<td></td>
</tr>
<tr>
<td>HofmanBang</td>
<td>-0.39 [-0.88, 0.10]</td>
<td></td>
</tr>
<tr>
<td>Johnston</td>
<td>-0.57 [-1.09, -0.06]</td>
<td></td>
</tr>
<tr>
<td>Jones</td>
<td>-0.01 [-0.10, 0.07]</td>
<td></td>
</tr>
<tr>
<td>Lewin</td>
<td>-0.61 [-0.99, -0.23]</td>
<td></td>
</tr>
<tr>
<td>Stern</td>
<td>-0.30 [-0.83, 0.23]</td>
<td></td>
</tr>
<tr>
<td>Thompson</td>
<td>-0.39 [-0.92, 0.14]</td>
<td></td>
</tr>
<tr>
<td>Toobert</td>
<td>-0.91 [-1.75, -0.08]</td>
<td></td>
</tr>
<tr>
<td><strong>Total (95% CI)</strong></td>
<td><strong>-0.29 [-0.48, -0.10]</strong></td>
<td><strong>-0.98 [-0.98, 0.40]</strong></td>
</tr>
</tbody>
</table>

Heterogeneity: $\tau^2 = 0.06$; $\chi^2 = 50.32$, df = 10 ($P < 0.00001$); $I^2 = 80\%$

Test for overall effect: $Z = 2.98$ ($P = 0.003$)
Complex interventions: lumping or splitting

• ‘Lumping’ of interventions can mask heterogeneity,
• ‘In principle’ research question such as “Do psychological therapies (as a whole), reduce depression after coronary heart disease?”
• What is the purpose of the review?
  – If is to investigate which type of psychological intervention is effective, or which intervention characteristics are effective, then ‘splitting’ may be the more appropriate approach
Subgroup analyses for exploring complexity

• Guise et al (2014) ways of grouping studies:
  – Key characteristics of interventions (e.g. group therapy, individual therapy, self-help)
  – Compare subclasses of intervention (mutually exclusive subgroups such as type of therapy – CBT, BT, counselling)

• Melendez-Torres (2015) “Clinically meaningful units”
  – by modality or similar theory of change
Subgroup analysis (splitting - characteristic)

### 1.2.1 Individual therapy

- **Black**
  - Std. Mean Difference: -0.60 [-1.12, -0.09]
- **Burgess**
  - Std. Mean Difference: 0.40 [0.06, 0.74]
- **HofmanBang**
  - Std. Mean Difference: -0.25 [-0.74, 0.23]
- **Johnston**
  - Std. Mean Difference: -0.56 [-1.07, -0.05]
- **Lewin**
  - Std. Mean Difference: -0.61 [-0.99, -0.23]
- **Thompson**
  - Std. Mean Difference: -0.39 [-0.92, 0.14]
- **Subtotal (95% CI)**
  - Std. Mean Difference: -0.32 [-0.69, 0.06]

Heterogeneity:
- Tau² = 0.17; Chi² = 20.87, df = 5 (P = 0.0009); I² = 76%
- Test for overall effect: Z = 1.66 (P = 0.10)

### 1.2.2 Group therapy

- **Stern**
  - Std. Mean Difference: -0.30 [-0.83, 0.23]
- **Toobert**
  - Std. Mean Difference: -0.91 [-1.75, -0.08]
- **Subtotal (95% CI)**
  - Std. Mean Difference: -0.52 [-1.09, 0.05]

Heterogeneity:
- Tau² = 0.06; Chi² = 1.46, df = 1 (P = 0.23); I² = 31%
- Test for overall effect: Z = 1.78 (P = 0.07)

### 1.2.3 Mixed therapy

- **Elderen**
  - Std. Mean Difference: -0.21 [-0.75, 0.34]
- **ENRICHOD**
  - Std. Mean Difference: -0.33 [-0.43, -0.24]
- **Jones**
  - Std. Mean Difference: -0.01 [-0.10, 0.07]
- **Subtotal (95% CI)**
  - Std. Mean Difference: -0.18 [-0.45, 0.10]

Heterogeneity:
- Tau² = 0.05; Chi² = 24.58, df = 2 (P < 0.00001); I² = 92%
- Test for overall effect: Z = 1.27 (P = 0.21)

Test for subgroup differences:
- Chi² = 1.22, df = 2 (P = 0.54), I² = 0%
Subgroup analysis (intervention type)

### 3.1.1 Cognitive behavioural therapy

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Std. Mean Difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>-0.60 [-1.12, -0.09]</td>
</tr>
<tr>
<td>Burgess</td>
<td>0.40 [0.06, 0.74]</td>
</tr>
<tr>
<td>ENRICHD</td>
<td>-0.33 [-0.43, -0.24]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>-0.16 [-0.69, 0.36]</td>
</tr>
</tbody>
</table>

Heterogeneity: \( \tau^2 = 0.18; \chi^2 = 18.06, \text{df} = 2 (P = 0.0001); I^2 = 89\%

Test for overall effect: \( Z = 0.62 \) (\( P = 0.54 \))

### 3.1.2 Behavioural therapy

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Std. Mean Difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HofmanBang</td>
<td>-0.39 [-0.88, 0.10]</td>
</tr>
<tr>
<td>Lewin</td>
<td>-0.61 [-0.99, -0.23]</td>
</tr>
<tr>
<td>Stern</td>
<td>-0.30 [-0.83, 0.23]</td>
</tr>
<tr>
<td>Toobert</td>
<td>-0.91 [-1.75, -0.08]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>-0.51 [-0.76, -0.26]</td>
</tr>
</tbody>
</table>

Heterogeneity: \( \tau^2 = 0.00; \chi^2 = 1.97, \text{df} = 3 (P = 0.58); I^2 = 0\%

Test for overall effect: \( Z = 4.00 \) (\( P < 0.0001 \))

### 3.1.3 Counselling

<table>
<thead>
<tr>
<th>Study or Subgroup</th>
<th>Std. Mean Difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elderen</td>
<td>-0.21 [-0.75, 0.34]</td>
</tr>
<tr>
<td>Johnston</td>
<td>-0.57 [-1.09, -0.06]</td>
</tr>
<tr>
<td>Jones</td>
<td>-0.01 [-0.10, 0.07]</td>
</tr>
<tr>
<td>Thompson</td>
<td>-0.39 [-0.92, 0.14]</td>
</tr>
<tr>
<td>Subtotal (95% CI)</td>
<td>-0.22 [-0.51, 0.07]</td>
</tr>
</tbody>
</table>

Heterogeneity: \( \tau^2 = 0.05; \chi^2 = 6.55, \text{df} = 3 (P = 0.09); I^2 = 54\%

Test for overall effect: \( Z = 1.51 \) (\( P = 0.13 \))
Intervention level network meta-analysis

Allows more studies to be combined, as long as they connect to the network – evidence base is strengthened.

Greater potential to explore heterogeneity

Coherent relative effect estimates based on more evidence, potentially more robust and precise
Intervention level network meta-analysis

\[ SMD_{BC}^{Ind} = SMD_{AC}^{Dir} - SMD_{AB}^{Dir} \]
**NMA of psychological interventions for CHD**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>SMD</th>
<th>95% CrIs</th>
</tr>
</thead>
<tbody>
<tr>
<td>BT vs TAU</td>
<td>-0.54</td>
<td>(-1.01 to -0.07)</td>
</tr>
<tr>
<td>CBT vs TAU</td>
<td>-0.17</td>
<td>(-0.66 to 0.32)</td>
</tr>
<tr>
<td>CSL vs TAU</td>
<td>-0.26</td>
<td>(-0.72 to 0.17)</td>
</tr>
<tr>
<td>CBT vs BT</td>
<td>0.37</td>
<td>(-0.33 to 1.06)</td>
</tr>
<tr>
<td>CSL vs BT</td>
<td>0.28</td>
<td>(-0.39 to 0.93)</td>
</tr>
<tr>
<td>CSL vs CBT</td>
<td>-0.09</td>
<td>(-0.78 to 0.56)</td>
</tr>
</tbody>
</table>

BT is ranked 1st (95% CrIs: 1st to 3rd)
CBT is ranked 2nd (95% CrIs: 1st to 4th)
Counselling is ranked 3rd (95% CrIs: 1st to 4th)
MRC definition of complexity (interventions)

- A number of interacting components within the experimental and control interventions,
- A number and difficulty of behaviours required by those delivering or receiving the intervention,
- A number of groups or organisational levels targeted by the intervention,
- A number and variability of outcomes,
- A degree of flexibility or tailoring of the intervention permitted.
What are intervention components?

- Complex interventions often considered greater than the sum of their parts.
- Components are defined as the active ingredients, processes, intervention techniques or “elements of an intervention that have the potential to causally influence outcomes”
- “Directly related to an intervention theory of change, which proposes the mechanisms by which an intervention works”

Guise et al, AHRQ; 2014. Report:14-EHC003-EF.
Blase and Fixsen, US Department of Health and Human Services 2011
Why focus on intervention components in SRs?

- Can explain a source of ‘clinical’ heterogeneity
- To understand how an intervention works
- To identify core drivers of intervention effect
  - Which components are essential for effectiveness
- To allow decision makers to adapt interventions without compromising effectiveness
- To optimize interventions for future studies.
Framework for evaluating components in NMA

- A component-based NMA approach likened to a factorial trial
- Akin to treating the network of evidence as a set of ‘dismantling trials’ comparing different combinations of components against each other (Melendez-Torres 2015)
- Nicky will discuss the approaches to modelling
Approaches to component identification

1. Inductive & iterative classification; coding of published papers
   • Intervention component classification (Sutcliffe et al, 2015)
   • Constant comparative method (Hetrick et al, 2015)

2. Review of entire subject literature to develop a taxonomy, typically with Delphi consensus
   • E.g. Taxonomy of behaviour change interventions (Michie 2013)

3. Automated approaches: AI and machine learning to extract information from intervention evaluation reports (Michie 2017)

4. Author contact: de Bruijn (2020) contacted authors with a list of active and control components.
   • 35% of experimental and 26% of comparator BCTs could be identified from published materials.
Case study: Component classification

- Interventions were classified into 5 groups: educational, behavioral, cognitive, relaxation, and psychosocial support.
  - Educational (EDU): educating patients about cardiovascular health risks and basic anatomy
  - Behavioral (BEH): change in domains relevant to coronary heart disease e.g., smoking cessation courses, physical exercise training, food preparation classes, and nutritional counseling sessions.
  - Cognitive (COG): restructuring patients' beliefs and perceptions re. health and coronary disease
  - Relaxation (REL) focused on training patients in different relaxation techniques, such as yoga and breathing courses.
  - Psychosocial support (SUP) interventions included attempts to bring patients together to encourage practical and/or emotional support.
Network plot: component combinations

TAU/T: treatment as usual
EDU/E: educational
BEH/B: behavioural
COG/C: cognitive
RELAX/R: relaxation
SUP/S, support.

+ indicates a combination of components, e.g. ‘E+B’ is educational and behavioural components.
Limitations of approach

• Networks may be sparse or not connected
  – Can only estimate effects between specific combinations that are connected in the network of evidence
  – Estimates of effect, may be imprecise

• Interventions are not only source of complexity
  – Interaction of intervention with setting should be considered.

• Methods of identifying and specifying components requires more research
  – Currently a balance between sufficiently specific for policy impact and sufficiently general for meaningful analysis
  – Reporting of complex interventions e.g. TiDier should improve field
References


