

# Approaches for synthesising complex mental health interventions in meta-analysis

Deborah M Caldwell, Nicky J Welton

School of Social and Community Medicine, University of Bristol, Bristol, UK

**Correspondence to** Dr Deborah M Caldwell, d.m.caldwell@bristol.ac.uk



Editor's choice  
Scan to access  
more  
free content



CrossMark

## ABSTRACT

Clinical and statistical heterogeneity are commonplace in meta-analysis of mental health interventions. One possible source of this heterogeneity is the complexity of the intervention being evaluated. Complexity may relate to the intervention, or to the way in which it is implemented; however, the most common interpretation of a complex intervention is one which has multiple, potentially interacting components. In this article we outline different analytical strategies suggested for incorporating intervention complexity in a meta-analysis.

## INTRODUCTION

According to the Centre for Evidence-Based Medicine's hierarchies of evidence, systematic reviews of homogenous well-conducted randomised controlled trials (RCTs) provide the best evidence for evaluating intervention effectiveness.<sup>1</sup> Homogeneity implies that each study included in a meta-analysis is estimating a single, true underlying relative intervention effect, so that any differences in estimates between studies are due to sampling error alone.<sup>2</sup> If homogeneity does not hold, a single, fixed treatment effect meta-analysis model should not be assumed and a random effects model may be more appropriate.<sup>2,3</sup> In systematic reviews of interventions for mental health, it could be argued that the homogeneity assumption is unlikely to hold in general. This is because the *clinical* variation observed across patient populations, therapist fidelity, intervention and comparator conditions, and outcomes (both patient and clinician reported) can give rise to *statistical* heterogeneity in meta-analysis. Indeed, heterogeneity might be considered inevitable.<sup>4</sup>

In systematic reviews of mental health interventions, the presence of statistical heterogeneity may be attributable to the *complexity* of the intervention being evaluated leading to potentially important differences across studies. 'Complexity' itself is a contested term,<sup>5</sup> however the Medical Research Council (MRC) have described the characteristics of complex interventions as having:

- ▶ 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.<sup>6</sup>

These characteristics may, or may not, be present in every complex intervention. Petticrew *et al*<sup>7</sup> classify characteristics of complexity as (1) those which relate to the intervention itself (such as multiple interacting components and flexibility of implementation) and (2) those which relate to the interventions' causal pathway (such as interaction with context, multiple mediators and moderators of effect).

Strategies to handle complex interventions in meta-analysis range from 'lumping' all interventions together<sup>8</sup> to sophisticated statistical modelling techniques.<sup>9</sup> The aim of this paper is to give an overview of the different analytical strategies suggested for incorporating intervention complexity in a systematic review, illustrated using a subset of studies from a Cochrane review examining psychological therapies for reducing depressive symptoms postcoronary heart disease.<sup>10</sup> We consider complexity only as it relates to the intervention and conceptualise a complex intervention as one, which has multiple, potentially interacting components. This is the most common interpretation.<sup>8</sup> Interested

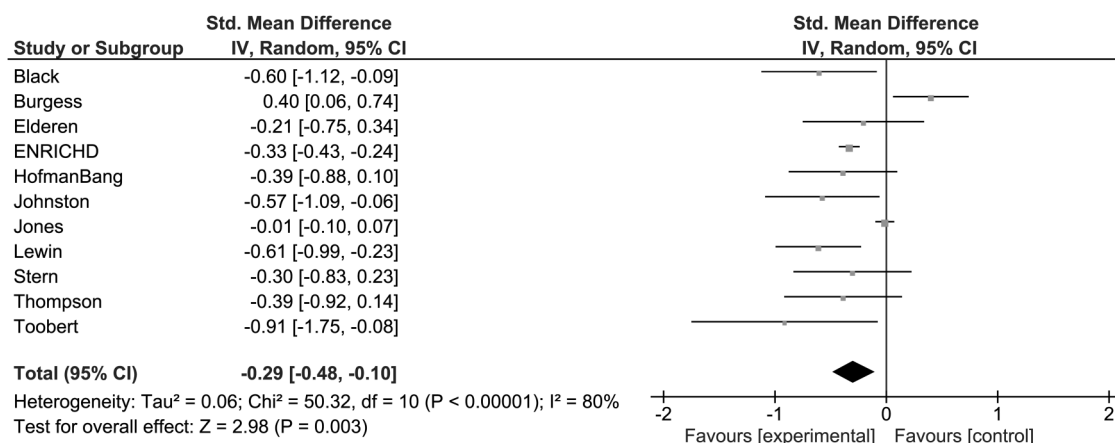
readers are referred to a special edition of the *Journal of Clinical Epidemiology* for consideration of strategies for handling to handle other aspects of complexity in systematic reviews.<sup>11</sup>

## FORMULATING THE RESEARCH QUESTION—LUMPING OR SPLITTING?

The analytical strategy for synthesising complex interventions should be pre-specified and begins with the formulation of a sensible research question, which in turn depends on the purpose of the review.<sup>12,13</sup> The specification of a review's objectives shapes whether the analytical strategy will 'lump' or 'split' interventions. For example, 'in principle' research questions such as 'Do psychological therapies (as a whole), reduce depression after coronary heart disease?' might take a lumping approach to analysis, since this question seeks to understand effectiveness in general. However, when complex interventions are 'lumped' together to form a single comparator, any between interventions variation is masked and is likely to manifest as increased, but unexplained, heterogeneity. Of course, the decision to lump interventions may also be taken for practical reasons, such as when there are few eligible studies for inclusion in the review.

Consider figure 1, which is adapted from a Cochrane review of 36 psychological interventions for coronary heart disease.<sup>10</sup> The outcome of interest here is reduction in depressive symptoms, for which 11 studies were included. The comparison is any psychological intervention versus control, where control is defined as standard care/treatment as usual (TAU). A fixed-effect meta-analysis was conducted by the authors and a standardised mean reduction of  $-0.18$  (95% CI  $-0.24$  to  $-0.12$ ) suggests that psychological interventions may affect a modest reduction in depression postcoronary heart disease. However, the  $p$  value for the  $\chi^2$  statistic provides extremely strong evidence against the null hypothesis of homogeneity (ie, that interventions are estimating a single underlying treatment effect).

The  $I^2$  statistic suggests that 75% of the variation between studies is attributable to heterogeneity and not chance. To account appropriately for the observed between-study variation a random-effect meta-analysis may have been more appropriate.<sup>2</sup> However, this would still only answer an 'in principle' question of general effectiveness and results would not enable a clinician to select a specific psychological intervention for their patient. The meaningful analysis of complex interventions can therefore pose problems if a 'lumped' approach is followed. Further exploration can be achieved by conducting an a priori specified subgroup analysis.<sup>3</sup> Figure 2 shows subgroup analysis by mode of therapy delivery, however this does not appear to explain the observed heterogeneity (individual therapy  $I^2=77\%$ ; group therapy  $I^2=86\%$ ), and the test for subgroup differences is non-significant ( $p=0.31$ ). In principle, the interventions could be further subgrouped such as 'individual+weekly meetings' or 'group+weekly meetings+telephone support'. However, caution should be exercised



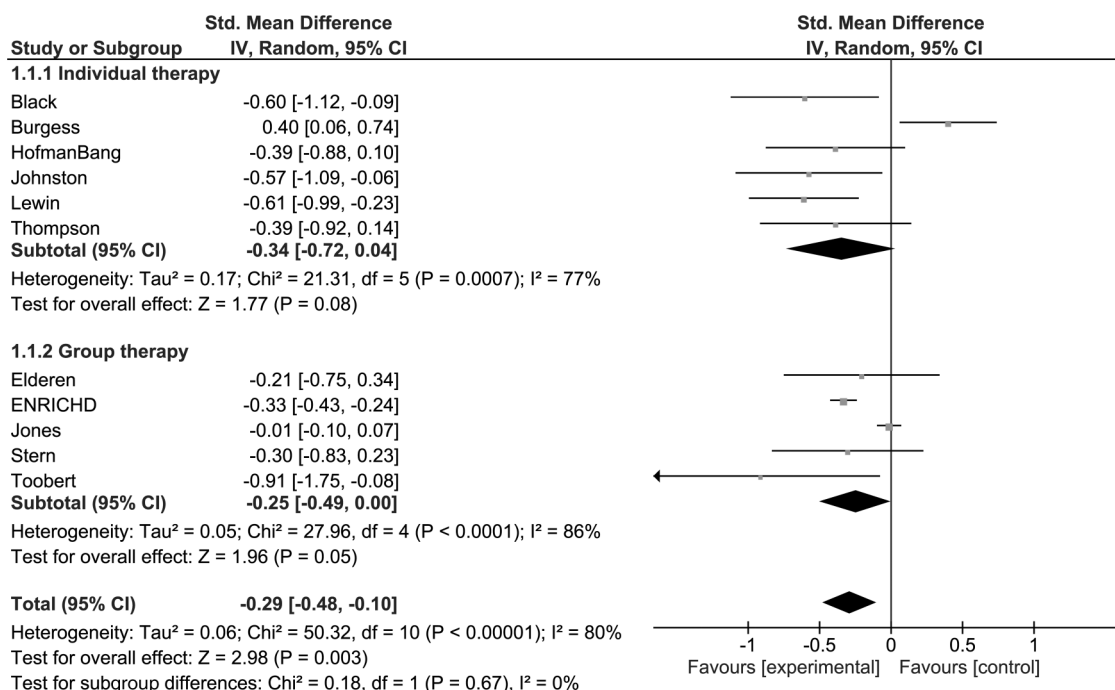
**Figure 1** Forest plot from a fixed effect meta-analysis comparing “any psychological intervention” vs control (treatment as usual). Outcome is reduction in depressive symptoms. Analysis conducted in Review Manager.<sup>43</sup>

since such analyses may suffer from low power due to the small number of included studies in each grouping. If the purpose of a review is to investigate which type of psychological intervention is effective, or which intervention characteristics are effective, then a review which categorises the intervention characteristics and ‘splits’ the analysis by intervention type may be the more appropriate and robust strategy. This can either be achieved as a series of separate reviews<sup>14–17</sup> or as separate analyses within the same review.<sup>18</sup>

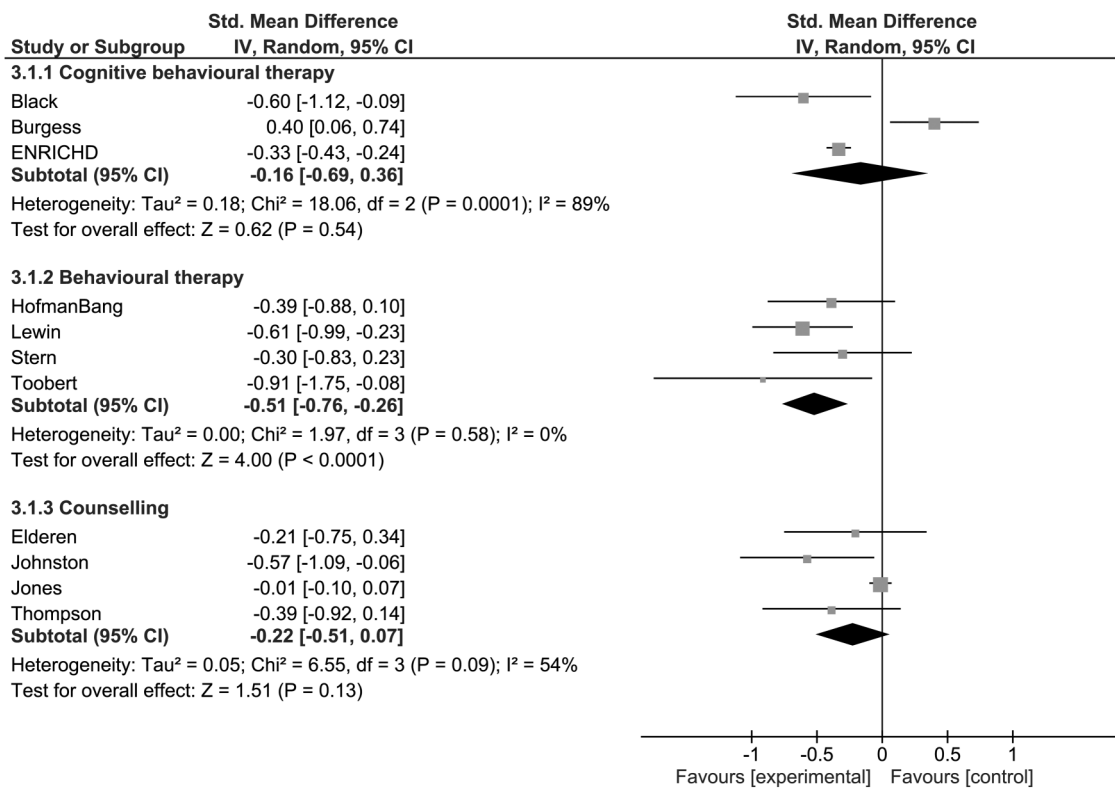
### CATEGORISATION OF INTERVENTION CHARACTERISTICS

There are a number of ways in which a splitting approach can be applied for meta-analyses of complex interventions. One possibility is to use the theoretical underpinning of the interventions to construct ‘clinically meaningful units’, which should be specified a priori.<sup>19</sup> In clinical psychology this might include classification by intervention modality such as cognitive-behavioural therapy (CBT), humanistic therapy or behavioural therapy (BT). In reviews of mental ill-health prevention,

interventions could be grouped by psychological or behavioural theory, such as the theory of planned behaviour, health belief model, social cognitive theory and so on.<sup>20 21</sup> In figure 3 the psychological interventions for coronary heart disease have been categorised according to intervention modality. The information was obtained from the Characteristics of Studies tables included in the original Cochrane review. The three modalities were CBT, BT and counselling-based interventions. Of these, we note only BT was associated with a reduction in depression and the I<sup>2</sup> is 0%. However, the between-study heterogeneity is still very high for CBT and counselling, and further investigation is warranted (note that estimates of heterogeneity become problematic when few studies are involved). One could disaggregate the intervention modalities further, for example, under CBT one might be interested in problem-solving therapies or rational-emotive behavioural therapies.<sup>22</sup> Note however that a balance needs to be found between a sufficiently detailed categorisation that can explain heterogeneity and sufficient numbers of studies for statistical power and to avoid spurious findings.



**Figure 2** Forest plot from a fixed effect meta-analysis comparing “any psychological intervention” vs control, interventions categorised as “individual therapy” and “group therapy”. Outcome is reduction in depressive symptoms. Analysis conducted in Review Manager.<sup>43</sup>



**Figure 3** Random effects meta-analysis comparing psychological intervention vs control (usual or standard care), interventions analysed as “clinically meaningful units”. Outcome is reduction in depressive symptoms. Analysis conducted in Review Manager.<sup>43</sup> Tau is the between-study variance. Its square root is the estimated standard deviation of underlying effects across studies.

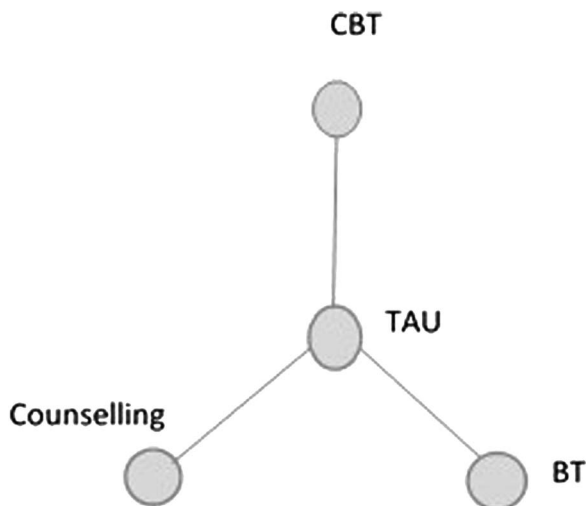
**COMPONENTS-BASED NETWORK META-ANALYSIS**

The obvious difficulty for standard, pairwise meta-analyses which seek to disaggregate complex interventions is that there are typically too few studies to allow clinically useful ‘splitting’ and inevitably, some degree of aggregation is needed for a meta-analysis to be conducted. In network meta-analyses (NMAs), however, the potential for

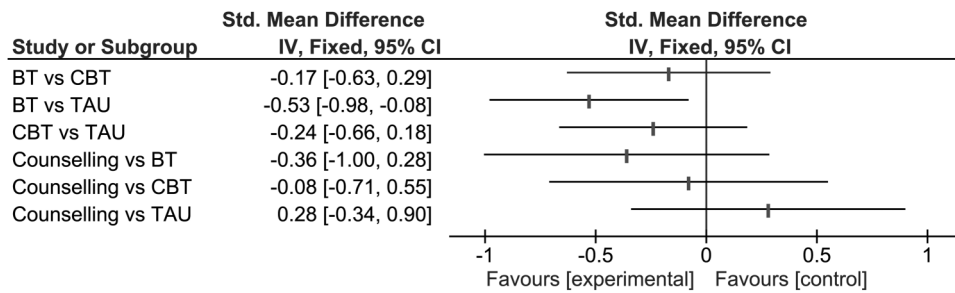
disaggregating the intervention is more promising.<sup>19</sup> A NMA is an extension of traditional pair-wise meta-analysis to include multiple interventions, as long as these interventions form a connected network of evidence (see figure 4, which shows a ‘star network’ where all interventions have been compared with the same common comparator—TAU). A key advantage of NMA is that it produces summary estimates of relative effectiveness regardless of whether interventions have been compared directly and ranks them according to the outcome measured (eg, effectiveness or safety). An additional advantage of NMA is that it allows more studies to be combined, as long as they connect to the network (eg, CBT vs counselling studies could be added to figure 4), bringing increased precision in the estimated intervention effects and the potential to explore statistical heterogeneity. For further details on the statistical methodology readers should see refs. 23–25 and for a discussion of the implications for systematic review methodology see ref. 26.

The ‘clinically meaningful unit’ classification approach explored above has been applied to a network of psychotherapies for treating depression,<sup>27</sup> treating acute depression in primary care<sup>28</sup> and psychotherapies for panic disorder.<sup>29</sup> Indeed, the ‘clinically meaningful unit’ analysis presented above in figure 3 could be re-analysed as an NMA with three interventions, and sharing a common heterogeneity parameter. Figure 4 depicts the network structure for this analysis; note that all active psychological interventions are compared to the ‘usual care’ node forming a star-shaped network. Just as in the pairwise meta-analyses above, it is assumed that the standard/TAU comparators are similar enough to be combined with the additional assumption that this must now apply across all interventions.<sup>23</sup>

In figure 5 the findings from the NMA are reported not only for the comparisons on which there is direct evidence but also for those where it is absent for example, BT versus CBT. There is substantial uncertainty surrounding the pairwise estimates of intervention effect,



**Figure 4** Each circle (node) represents an intervention component defined as a ‘clinically meaningful unit’ as extracted from the psychological interventions for coronary heart disease review<sup>8</sup> for the outcome of depression. The solid lines indicate where there was direct information available between comparisons. CBT, cognitive-behavioural therapy; TAU, treatment as usual; BT, behavioural therapy.



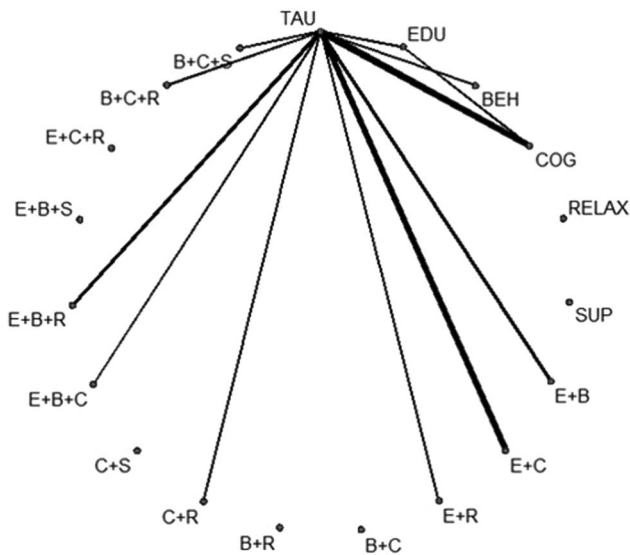
**Figure 5** Standardised mean difference estimates from NMA of each psychological intervention versus every other. Interventions analysed as “clinically meaningful units”. Outcome is reduction in depressive symptoms. Analysis conducted using OpenBUGS<sup>44</sup> and results plotted using RevMan.<sup>43</sup>

and the only comparison that reaches conventional statistical significance is BT vs TAU (note the NMA was performed in a Bayesian framework, which accounts for the wider CIs when compared to figure 3). On the basis of this NMA, CBT is ranked 3rd (95% CIs 1st to 4th) best in terms of reducing depressive symptoms, BT is ranked 1st (95% CIs 1st to 3rd) and counselling is ranked 2nd (95% CIs 1st to 4th). TAU is the ‘worst’ intervention. Note the CIs around the rankings reflect the considerable uncertainty observed in the effect estimates. In NMA a single between-study heterogeneity parameter is typically assumed.<sup>23</sup> Here the estimate of  $\tau^2$  is 0.11 which might be considered to represent a moderate level of heterogeneity. Note that the NMA assumes that the heterogeneity is the same regardless of which comparison is being made. This may not be appropriate here, since we found more heterogeneity in the CBT vs TAU comparison than for the other comparisons

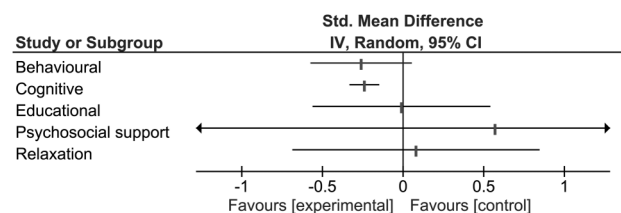
(figure 3). This suggests that the CBT classification may be too broad to capture the complex nature of CBT interventions.

### MULTICOMPONENTS-BASED NMA

Within an NMA framework, the analyst has greater flexibility to evaluate complex multicomponent interventions and to investigate whether interventions with a particular component(s) are more likely to be effective. Components are defined as the ‘active ingredients’, ‘intervention techniques’ or ‘elements of an intervention that have the potential to causally influence outcomes’.<sup>8</sup> As such they may be classified on practical elements, for example, activities, mode of delivery, setting and/or on theoretical underpinnings of the intervention. If there are common components across all interventions in the network, the components effectively become the intervention ‘nodes’ in the network and an NMA can be conducted. Figure 6 represents a multicomponents-based network plot for the coronary heart disease example. Welton *et al*<sup>30</sup> conducted a components-based NMA for the coronary heart disease network. Interventions were classified according to five key components; educational, behavioural, cognitive, relaxation and psychosocial support. Describing their model as a meta-regression-based extension to NMA, three models were evaluated in a Bayesian framework: (1) an additive main effects model which assumes that the effect of each component adds (ie, no synergistic or antagonistic effects), (2) a two-way interaction model (allowing pairs of components to have either a bigger or smaller effect than would be expected from the sum of their effects alone) and (3) a full-interaction model for interventions described as having >2 components (eg, cognitive+behavioural+support). To illustrate, their results for the depression outcome are shown in figure 7 for the main effects additive model. This analysis answers the question ‘Which intervention component has the greatest probability of being most effective?’ Compared to the broader categorisation used in figure 5, having broken down interventions into their component parts heterogeneity is now reduced;  $\tau^2=0.03$ . There is some evidence that an intervention with a cognitive and/or behavioural component(s) was associated with a reduction in standardised mean



**Figure 6** Network plot for a multi-component-based network meta-analysis: components identified from a full interaction model for psychological interventions for coronary heart disease. Each node represents a component, or combination of components, identified from the interventions included in the NMA for psychological interventions for coronary heart disease review<sup>8</sup> for outcome of depression. The diagram depicts all possible combinations of components from the full interaction model. The solid lines indicate where there was direct information available between component comparisons. 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, for example, ‘E+B’ is educational and behavioural components.



**Figure 7** Components based NMA: additive main effect model for psychological interventions relative to usual care. Analysis conducted using OpenBUGS<sup>44</sup> and results plotted using RevMan.<sup>43</sup>

depression score; for the cognitive component the pooled Standardised mean difference (SMD) was  $-0.26$  (95% credible interval:  $-0.55$  to  $0.02$ ) and for behavioural it was SMD  $-0.24$  (95% credible interval:  $-0.42$  to  $-0.06$ ).

## DISCUSSION

Component-based NMA is an option for the synthesis of complex interventions in the presence of heterogeneity. Of course, intervention categorisation is only one dimension contributing to heterogeneity in meta-analyses of complex interventions. In the above example, heterogeneity was explained by intervention definition but this may not be the case for all examples, where additional factors may cause residual heterogeneity (eg, an imbalance of effect modifiers across studies). A possible source of confounding here is the control intervention. In RCTs in clinical psychology and psychiatry control interventions may take several forms—waiting list controls are common as are no intervention controls.<sup>31</sup> A psychological placebo, where the intervention is regarded as inactive by the researchers but is judged as active by the participants, may be used. Similarly an attention placebo could be used where the control mimics the theoretically inactive elements of an intervention, but not the active elements.<sup>32–33</sup> Reviewers of complex interventions should also be mindful that TAU and standard care may differ across settings, contexts and countries, even though systematic reviews have traditionally lumped these into a single control.<sup>9–34</sup> Unfortunately, due to the small number of studies in the psychotherapy for coronary heart disease meta-analysis, further disaggregating by control intervention is of questionable value.

Component-based systematic reviews are becoming increasingly common as analysts realise the importance of identifying and investigating heterogeneity, regardless of its inevitability.<sup>35–36</sup> However, one difficulty in a components-based approach is the identification of distinct components from the published literature.<sup>37</sup> Complex interventions may not be described in sufficient detail to allow dismantling of key ingredients. Recent reporting guideline initiatives, such as CREdICI 2 (Criteria for Reporting the Development and Evaluation of Complex Interventions in healthcare: revised guideline)<sup>38</sup> and CONSORT-SPI (Social and Psychological Interventions)<sup>39</sup> seek to address this. The MRC's recent guidance on process evaluations may also help identification of components, and assessment of delivery and fidelity of complex interventions.<sup>40</sup> How to classify complex interventions and disaggregate the multiple interacting components within them is an area of ongoing interest. Several taxonomies have been developed; some designed for use in specific clinical areas and others are generic.<sup>41</sup> Further research is needed to assess the application of taxonomies across clinical areas. Logic models describing the mechanisms of action and casual pathways of interventions are increasingly used to structure systematic reviews of complex interventions<sup>42</sup> and could also be used to inform the classification of intervention components. What is clear, however, is that whichever approach the analyst chooses to categorise interventions it is desirable that components be specified a priori, and published in a protocol before data extraction to avoid data-driven decisions and reduce the likelihood of spurious findings.

**Funding** DMC is supported by a Medical Research Council Population Health Scientist fellowship award G0902118.

This work was undertaken with the support of The Centre for the Development and Evaluation of Complex Interventions for Public Health Improvement (DECIPHer), a UKCRC Public Health Research Centre of Excellence. Joint funding (MR/K023233/1) from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, the Welsh Government and the Wellcome Trust, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged.

**Competing interests** None declared.

**Provenance and peer review** Commissioned; externally peer reviewed.

doi:10.1136/eb-2015-102275

Received 15 November 2015; Revised 16 November 2015;

Accepted 19 November 2015

## REFERENCES

- Centre for Evidence Based Medicine *Levels of Evidence*. 2015 (cited 15 October 2015). <http://www.cebm.net>
- Nikolakopoulou A, Mavridis D, Salanti G. Demystifying fixed and random effects meta-analysis. *Evid Based Ment Health* 2014;**17**:53–7.
- Higgins J, Green S, eds. *Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]*. The Cochrane Collaboration, 2011. <http://www.cochrane-handbook.org>
- Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *J R Stat Soc Ser A Stat Soc* 2009;**172**:137–59.
- Wong G. Is complexity just too complex? *J Clin Epidemiol* 2013;**66**:1199–201.
- Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;**337**:a1655.
- Petticrew M, Anderson L, Elder R, et al. Complex interventions and their implications for systematic reviews: a pragmatic approach. *J Clin Epidemiol* 2013;**66**:1209–14.
- Kühne F, Ehmcke R, Härter M, et al. Conceptual decomposition of complex health care interventions for evidence synthesis: a literature review. *J Eval Clin Pract* 2015;**21**:817–23.
- Jakobsen JC, Hansen JL, Storebø OJ, et al. The effects of cognitive therapy versus 'no intervention' for major depressive disorder. *PLoS ONE* 2011;**6**:e28299.
- Rees K, Bennett P, West R, et al. Psychological interventions for coronary heart disease. *Cochrane Database Syst Rev* 2004;(2):CD002902.
- Tugwell P, Knottnerus JA, Idzerda L. Complex interventions—how should systematic reviews of their impact differ from reviews of simple or complicated interventions? *J Clin Epidemiol* 2013;**66**:1195–6.
- Squires JE, Valentine JC, Grimshaw JM. Systematic reviews of complex interventions: framing the review question. *J Clin Epidemiol* 2013;**66**:1215–22.
- Anderson LM, Petticrew M, Chandler J, et al. Introducing a series of methodological articles on considering complexity in systematic reviews of interventions. *J Clin Epidemiol* 2013;**66**:1205–8.
- Churchill R, Moore TH, Furukawa TA, et al. 'Third wave' cognitive and behavioural therapies versus treatment as usual for depression. *Cochrane Database Syst Rev* 2013;**10**:CD008705.
- Hunot V, Moore TH, Caldwell DM, et al. 'Third wave' cognitive and behavioural therapies versus other psychological therapies for depression. *Cochrane Database Syst Rev* 2013;**10**:CD008704.
- Shinohara K, Honyashiki M, Imai H, et al. Behavioural therapies versus other psychological therapies for depression. *Cochrane Database Syst Rev* 2013;**10**:CD008696.
- Caldwell D, Hunot V, Moore TH, et al. Behavioural therapies versus treatment as usual for depression. *Cochrane Database Syst Rev* 2010;(9):CD008697.
- Petticrew M, Rehfuss E, Noyes J, et al. Synthesizing evidence on complex interventions: how meta-analytical, qualitative, and mixed-method approaches can contribute. *J Clin Epidemiol* 2013;**66**:1230–43.
- Melendez-Torres GJ, Bonell C, Thomas J. Emergent approaches to the meta-analysis of multiple heterogeneous complex interventions. *BMC Med Res Methodol* 2015;**15**:47.
- Faggiano F, Vigna-Taglianti FD, Versino E, et al. School-based prevention for illicit drugs use: a systematic review. *Prev Med* 2008;**46**:385–96.
- Michie S, Prestwich A. Are interventions theory-based? Development of a theory coding scheme. *Health Psychol* 2010;**29**:1–8.
- Churchill R, Moore TH, Caldwell D, et al. Cognitive behavioural therapies versus other psychological therapies for depression. *Cochrane Database Syst Rev* 2010;(9):CD008698.
- Mavridis D, Giannatsi M, Cipriani A, et al. A primer on network meta-analysis with emphasis on mental health. *Evid Based Ment Health* 2015;**18**:40–6.
- Dias S, Sutton AJ, Ades AE, et al. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making* 2013;**33**:607–17.
- Dias S, Welton NJ, Sutton AJ, et al. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. *Med Decis Making* 2013;**33**:641–56.
- Caldwell DM. An overview of conducting systematic reviews with network meta-analysis. *Syst Rev* 2014;**3**:109.
- Barth J, Munder T, Gerger H, et al. Comparative efficacy of seven Psychotherapeutic interventions for patients with depression: a network meta-analysis. *PLoS Med* 2013;**10**:e1001454.
- Linde K, Sigterman K, Kriston L, et al. Effectiveness of psychological treatments for depressive disorders in primary care: systematic review and meta-analysis. *Ann Fam Med* 2015;**13**:56–68.
- Pompoli A, Furukawa TA, Imai H, et al. Psychological therapies for panic disorder with or without agoraphobia in adults. *Cochrane Database Syst Rev* In press.

30. **Welton NJ**, Caldwell DM, Adamopoulos E, *et al*. Mixed treatment comparison meta-analysis of complex interventions: psychological interventions in coronary heart disease. *Am J Epidemiol* 2009;**169**:1158–65.
31. **Button KS**, Munafò MR. Addressing risk of bias in trials of cognitive behavioral therapy. *Shanghai Arch Psychiatry* 2015;**27**:144–8.
32. **Furukawa TA**, Noma H, Caldwell DM, *et al*. Waiting list may be a placebo condition in psychotherapy trials: a contribution from network meta-analysis. *Acta Psychiatr Scand* 2014;**130**:181–92.
33. **Zhu Z**, Zhang L, Jiang J, *et al*. Comparison of psychological placebo and waiting list control conditions in the assessment of cognitive behavioral therapy for the treatment of generalized anxiety disorder: a meta-analysis. *Shanghai Arch Psychiatry* 2014;**26**:319–31.
34. **Cuijpers P**, Smit F, Bohlmeijer E, *et al*. Efficacy of cognitive-behavioural therapy and other psychological treatments for adult depression: meta-analytic study of publication bias. *Br J Psychiatry* 2010;**196**:173–8.
35. **Greaves CJ**, Sheppard KE, Abraham C, *et al*. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. *BMC Public Health* 2011;**11**:119.
36. **Squires JE**, Sullivan K, Eccles MP, *et al*. Are multifaceted interventions more effective than single-component interventions in changing health-care professionals' behaviours? An overview of systematic reviews. *Implement Sci* 2014;**9**:152.
37. **Guise J-M**, Chang C, Viswanathan M, *et al*. Agency for Healthcare Research and Quality Evidence-based Practice Center methods for systematically reviewing complex multicomponent health care interventions. *J Clin Epidemiol* 2014;**67**:1181–91.
38. **Möhler R**, Köpke S, Meyer G. Criteria for Reporting the Development and Evaluation of Complex Interventions in healthcare: revised guideline (CRReDECI 2). *Trials* 2015;**16**:204.
39. **Gardner F**, Mayo-Wilson E, Montgomery P, *et al*. Editorial Perspective: The need for new guidelines to improve the reporting of trials in child and adolescent mental health. *J Child Psychol Psychiatry* 2013;**54**:810–12.
40. **Moore GF**, Audrey S, Barker M, *et al*. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ* 2015;**350**:h1258.
41. **Michie S**, Richardson M, Johnston M, *et al*. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. *Ann Behav Med* 2013;**46**:81–95.
42. **Anderson LM**, Petticrew M, Rehfues E, *et al*. Using logic models to capture complexity in systematic reviews. *Res Synth Methods* 2011;**2**:33–42.
43. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.
44. Available from <http://www.openbugs.net/>