University of BRISTOL

Lumping and splitting in meta-analysis of complex interventions

Deborah Caldwell

University of Bristol

University of BRISTOL

Background

- Systematic review of well-conducted RCTs provides high quality evidence for *effectiveness*
  - Homogeneity: each study in a meta-analysis is estimating a single, true underlying relative intervention effect.
  - Any differences in estimates between studies due to sampling error alone
- Unlikely to hold for many public health, policy or social interventions.
  - Variation in population, interventions & comparator, context (delivery/ implementation/ adherence/ length, setting) outcomes etc.
- Heterogeneity might be considered inevitable?

University of BRISTOL

"Reasons or excuses for avoiding meta-analysis…"*

Table 1 | Reasons for not showing summary estimates in forest plots from systematic reviews in Cochrane database 2005 issue 4

Reason	No (%) of systematic reviews (n=135)*
Statistical heterogeneity too high	32 (24)
Different interventions compared	41 (30)
Different metrics or outcomes evaluated	26 (19)
Different metric of same outcome	7
Different outcome	20
Different study designs	21 (16)
Non-randomised studies	3
Other design issues	18
Different study participants, settings	21 (16)
Data with many counts per participant	5 (4)
Data too limited	11 (8)
Clinical heterogeneity (not otherwise specified)	5 (4)
Synthesis considered inappropriate (not specified why)	3 (2)
Non-normality of data	1 (1)
No reason given	10 (7)
Artefact†	3 (2)
Quantitative synthesis given in text	7 (5)

*Ioannidis 2008; BMJ 336;1413

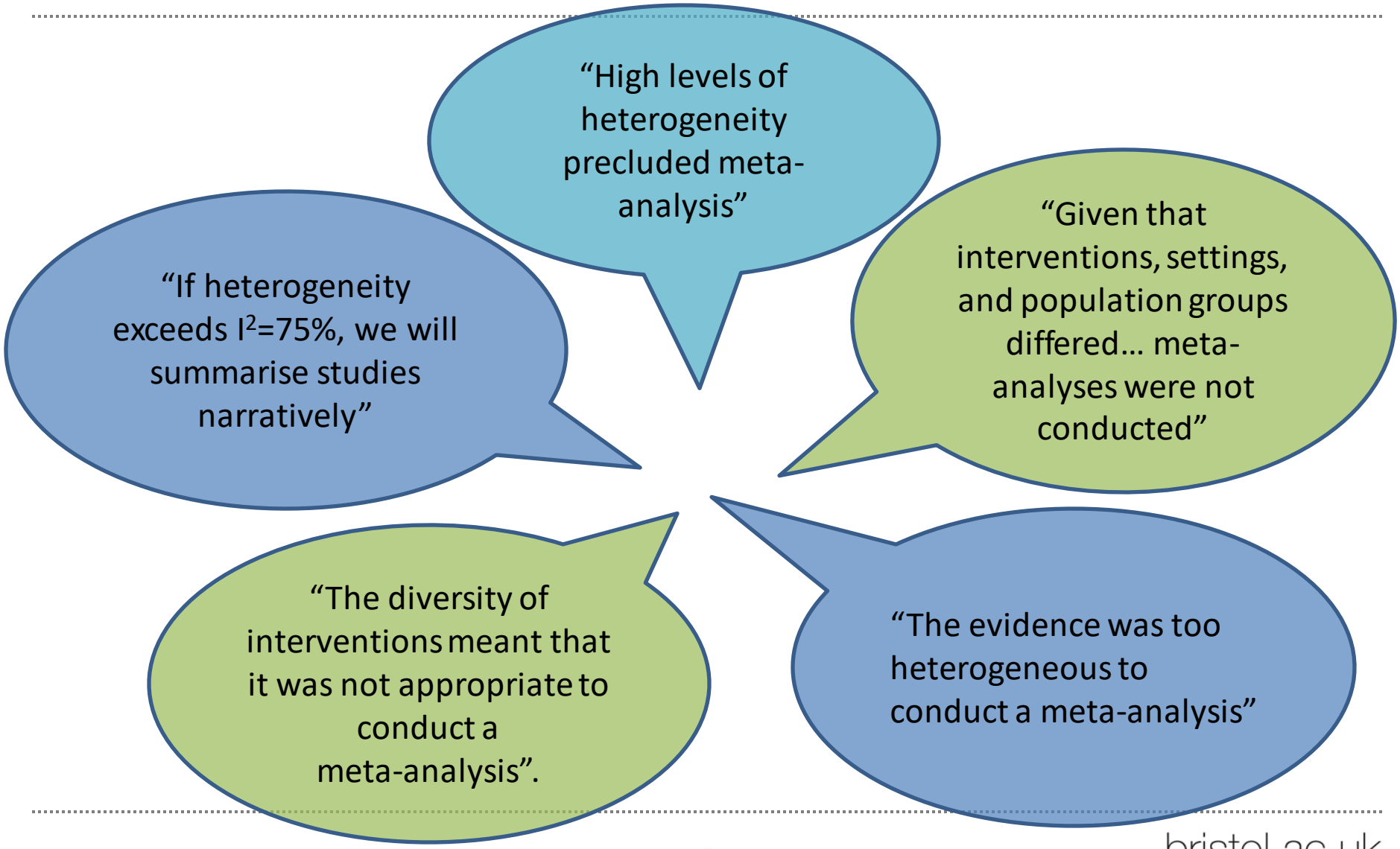
University of BRISTOL

Options for synthesis*

- non-quantitative synthesis (e.g. SWiM, tabulation, narrative, graphical approaches)
- standard meta-analysis methods (pairwise, fixed, random effects, meta-regression)
- advanced synthesis methods (NMA, MPES, MVMA)

*Higgins JPT et al *BMJ Global Health* 2019;**4:**e000858.

University of BRISTOL



MA, complexity and heterogeneity

"The challenge then is how to limit the boundaries of the review such that the engagement with heterogeneity can produce useful findings".

RESEARCH Open Ac

Meta-analysis, complexity, and heterogeneity: a qualitative interview study of researchers' methodological values and practices

Theo Lorenc[1*], Lambert Felix[2], Mark Petticrew[2], G J Melendez-Torres[3], James Thomas[4], Sian Thomas[2], Alison O'Mara-Eves[4] and Michelle Richardson[4]

"Traditionally, it was assumed that heterogeneity should be minimised to ensure the reliability of review findings. In the presence of complexity this may not be appropriate, since an adequate engagement with complex interventions and contexts demands the integration of heterogeneous types of data. In this context statistical heterogeneity is arguably to be expected, and may not be a useful indicator of problems with the data, but present opportunities for explanatory analysis."

University of BRISTOL

Example

- Subset of studies from a 2004 Cochrane review examining psychosocial interventions 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
- Consider the outcome of depression symptoms
- 11 studies, Published from 1983 to 2002

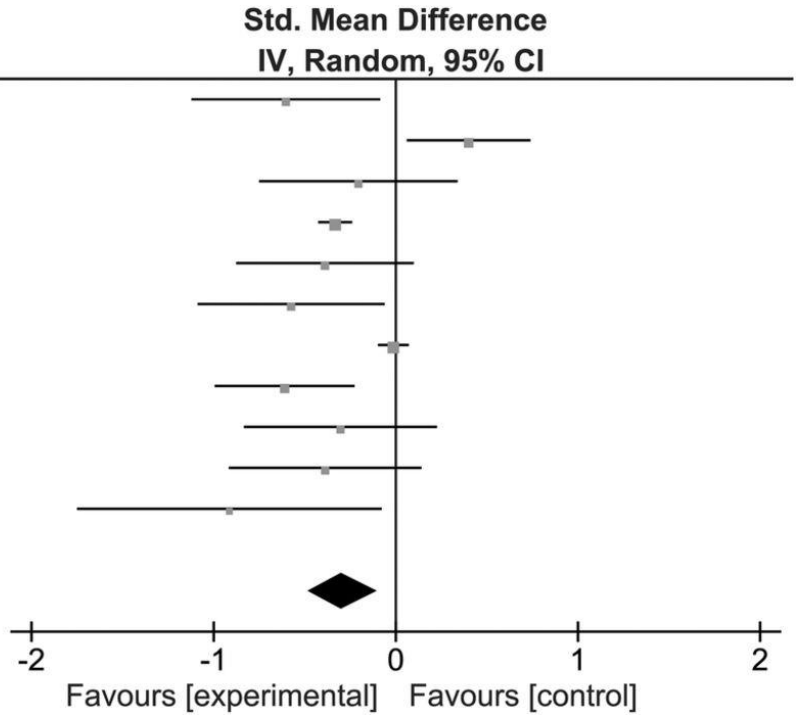
Welton et al AJE 2009:169; 1158-65
https://academic.oup.com/aje/article/169/9/1158/125216

Random-effects meta-analysis

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

Heterogeneity: Tau² = 0.06; Chi² = 50.32, df = 10 (P < 0.00001) I² = 80%

Test for overall effect: Z = 2.98 (P = 0.003)



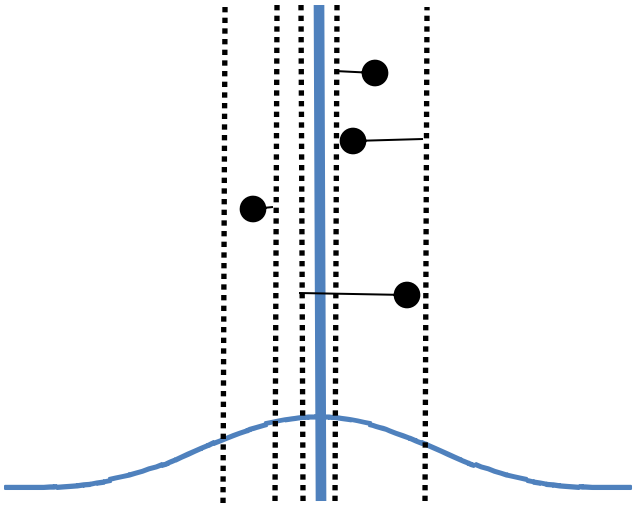
(Some) heterogeneity is inevitable

- Received wisdom is fixed effect MA inappropriate
- Random effect model can be difficult to interpret

Assumes:
true treatment effect in each study is randomly, normally distributed between studies, with variance τ^2

Estimates:
mean of the distribution of the trial-specific true treatment effects
i.e. summary estimate is the *average effect* across the included studies

May not reflect observed effect in any study

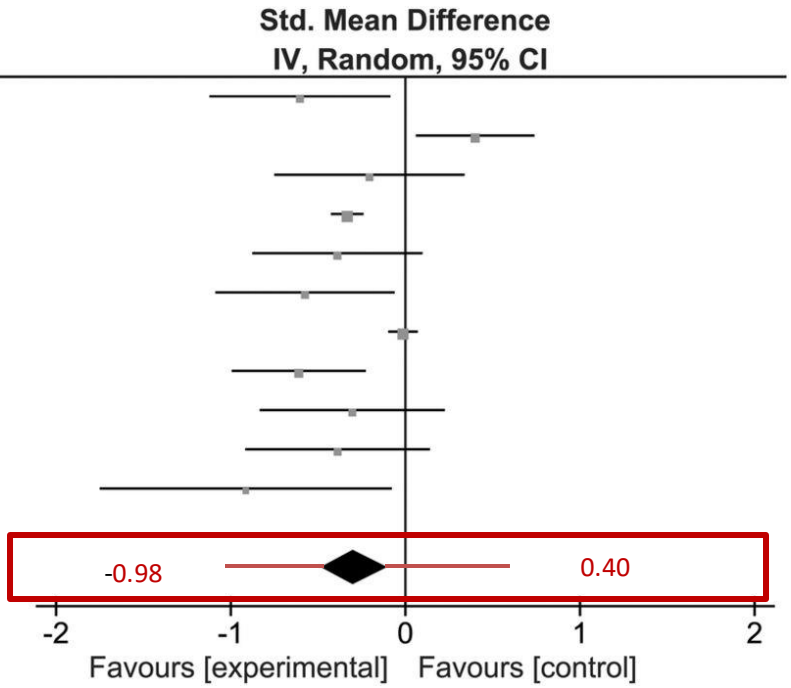


Predictive interval

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

Heterogeneity: Tau² = 0.06; Chi² = 50.32, df = 10 (P < 0.00001); I² = 80%

Test for overall effect: Z = 2.98 (P = 0.003)



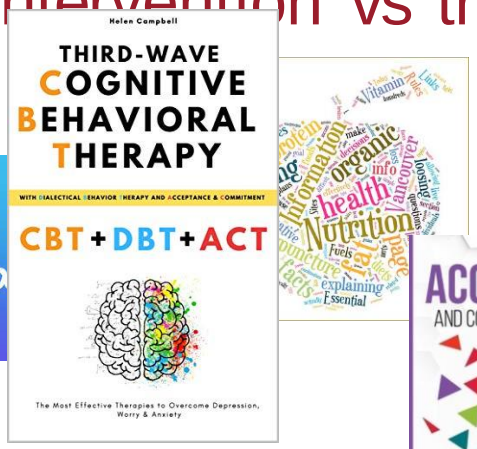
"How to limit the boundaries of the review*"

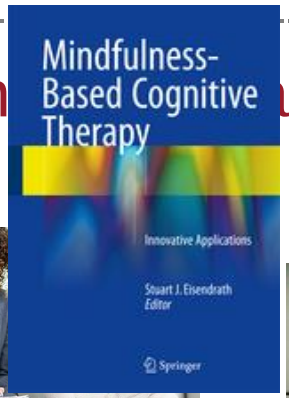
- What is the purpose of the review?
  - If it is literature summary, or an 'in principle' research question such as "Do psychosocial interventions reduce depression after coronary heart disease?", then 'lumping' interventions is reasonable.
  - If is to investigate which type of psychological intervention is effective, or which intervention characteristics are effective, then 'splitting' may be more appropriate.
- 'Lumping' of interventions can mask heterogeneity

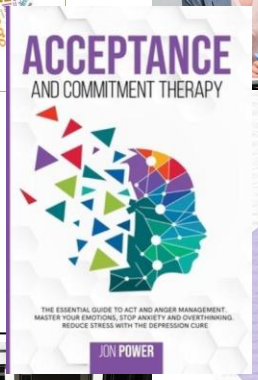
*Systematic Reviews 2016 5:192

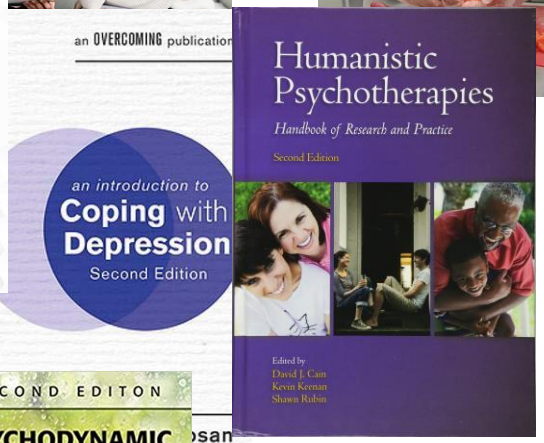
University of BRISTOL

Psychosocial intervention vs treatmen al

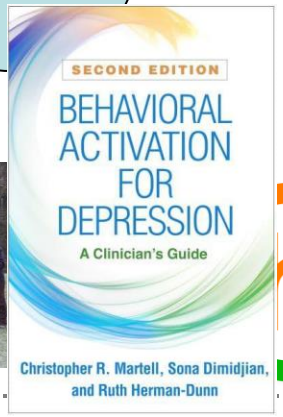


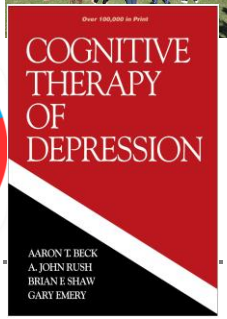


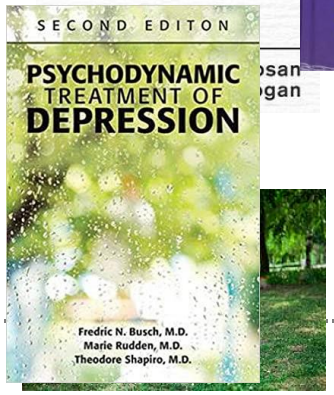




PSY







University of BRISTOL

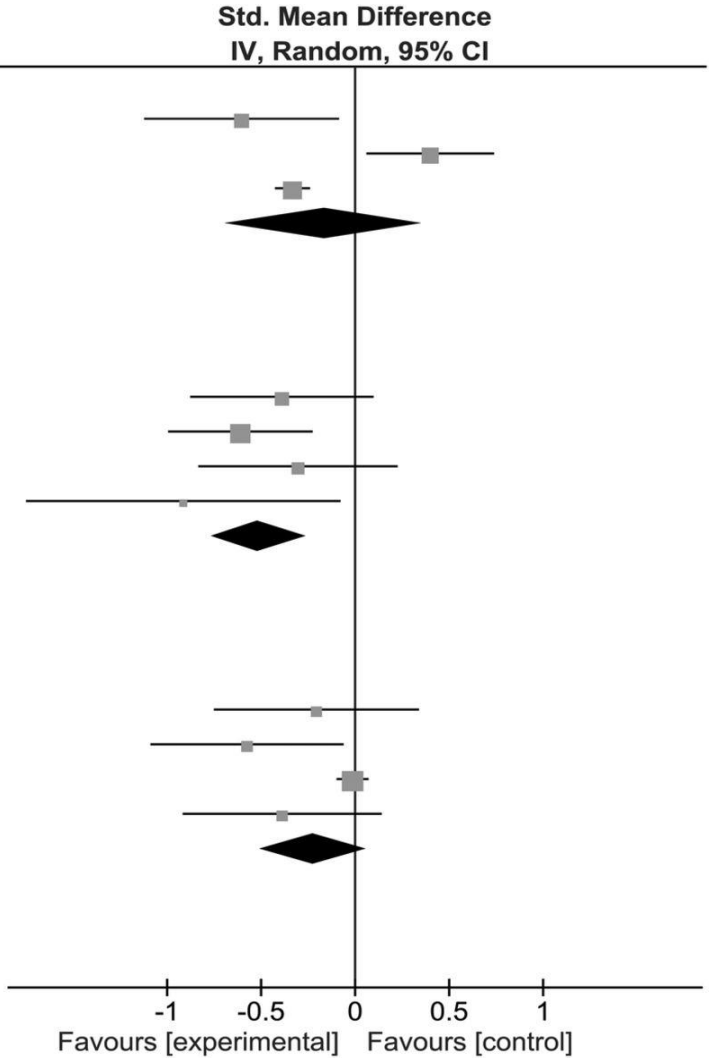
Subgroup analyses for complex interventions

- 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

Guise et al, AHRQ Report:14-EHC003-EF. 2014.
Melendez-Torres et al, *BMC Med Res Methodol* . 2015: 15

Subgroup analysis (intervention type)

Study or Subgroup	Std. Mean Difference IV, Random, 95% CI
3.1.1 Cognitive behavioural therapy	
Black	-0.60 [-1.12, -0.09]
Burgess	0.40 [0.06, 0.74]
ENRICHD	-0.33 [-0.43, -0.24]
Subtotal (95% CI)	-0.16 [-0.69, 0.36]
Heterogeneity: Tau² = 0.18; Chi² = 18.06, df = 2 (P = 0.0001); I² = 89%	
Test for overall effect: Z = 0.62 (P = 0.54)	
3.1.2 Behavioural therapy	
HofmanBang	-0.39 [-0.88, 0.10]
Lewin	-0.61 [-0.99, -0.23]
Stern	-0.30 [-0.83, 0.23]
Toobert	-0.91 [-1.75, -0.08]
Subtotal (95% CI)	-0.51 [-0.76, -0.26]
Heterogeneity: Tau² = 0.00; Chi² = 1.97, df = 3 (P = 0.58); I² = 0%	
Test for overall effect: Z = 4.00 (P < 0.0001)	
3.1.3 Counselling	
Elderen	-0.21 [-0.75, 0.34]
Johnston	-0.57 [-1.09, -0.06]
Jones	-0.01 [-0.10, 0.07]
Thompson	-0.39 [-0.92, 0.14]
Subtotal (95% CI)	-0.22 [-0.51, 0.07]
Heterogeneity: Tau² = 0.05; Chi² = 6.55, df = 3 (P = 0.09); I² = 54%	
Test for overall effect: Z = 1.51 (P = 0.13)	



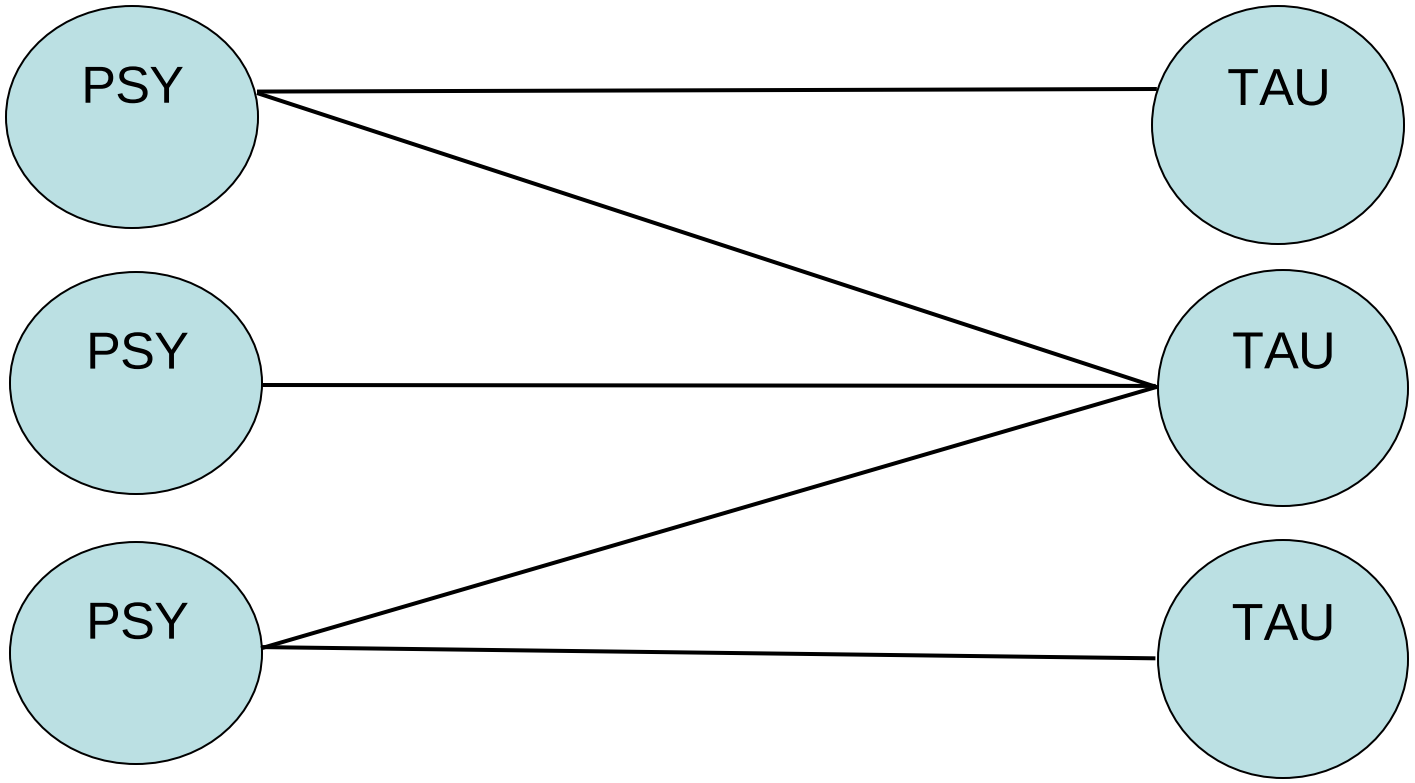

PSY
TAU

University of BRISTOL

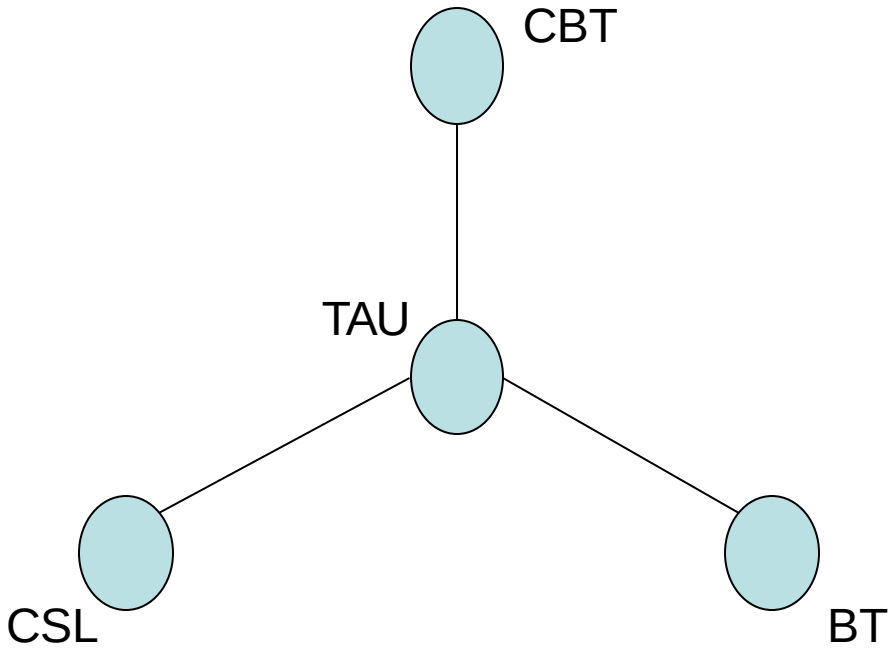
Challenges for subgroup analysis of complex interventions*

- Differences between subgroups may not be because of the characteristic used to define subgroups (in other words, there may be confounding because subgroup comparisons are observational by nature)
- Sufficient number of studies are required to provide convincing results (low power)
- Performing multiple subgroup analyses may involve multiple statistical tests, with inflated risk of spurious findings
- Characteristics may be correlated, hampering interpretation.

*Petticrew et al Journal of Clinical Epidemiology 2013:66; 1230-1243


PSY
PSY
PSY
TAU
TAU
TAU

Intervention level network meta-analysis

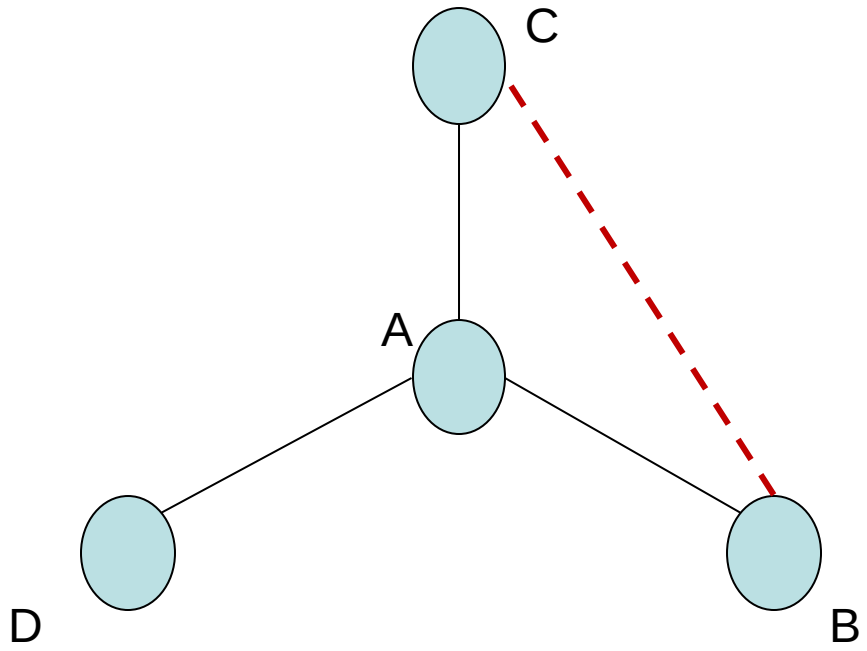


Coherent relative effect estimates based on more evidence, potentially more robust and precise.

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

Greater potential to explore heterogeneity?

Intervention level network meta-analysis



$$SMD_{BC}^{Ind} = SMD_{AC}^{Dir} - SMD_{AB}^{Dir}$$

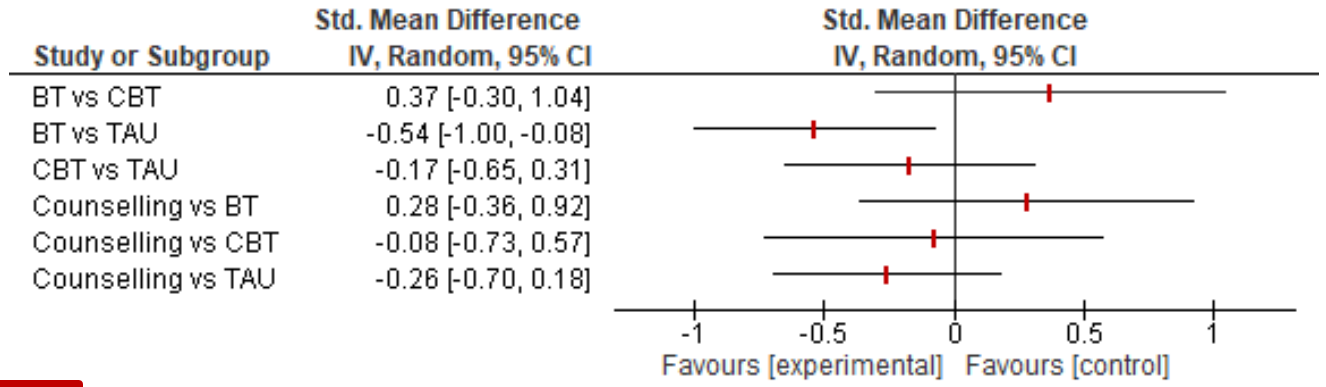
Validity of NMA:

AB, AC & AD studies similar across factors which may affect the outcome (modify treatment effect).

https://training.cochrane.org/handbook/current/chapter-11

University of BRISTOL

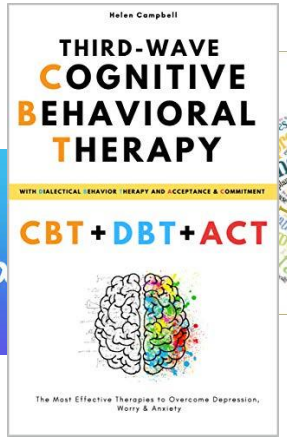
NMA of psychological interventions for CHD

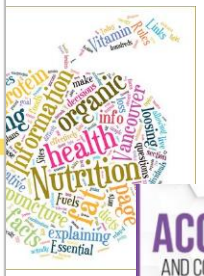


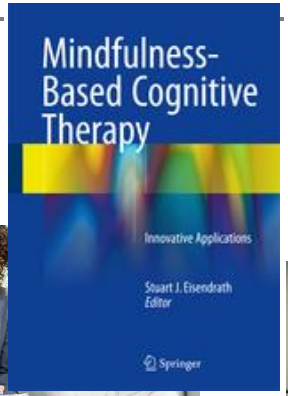
$\tau^2 = 0.11$

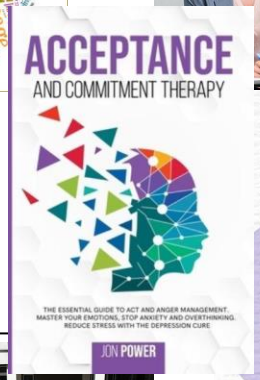
BT is ranked 1st (95% CrIs: 1st to 3rd)
CBT is ranked 2nd (95% CrIs: 1st to 4th)
Counselling is ranked 3rd (95% CrIs: 1st to 4th)

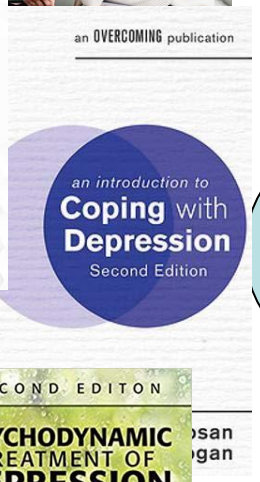
University of BRISTOL



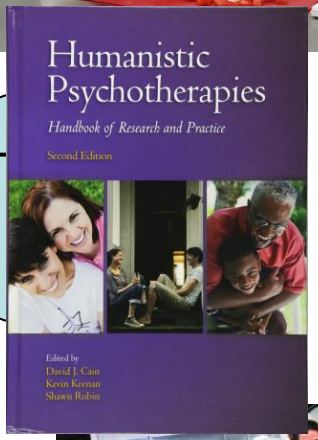


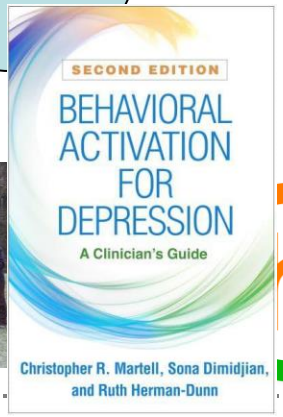


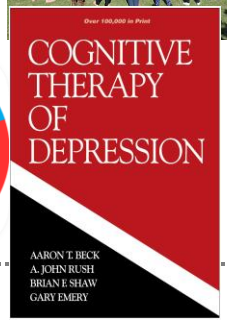


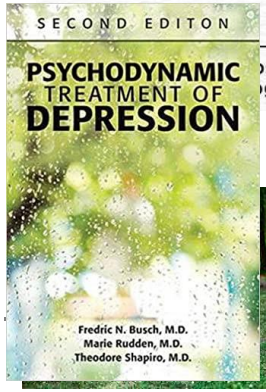


PSY





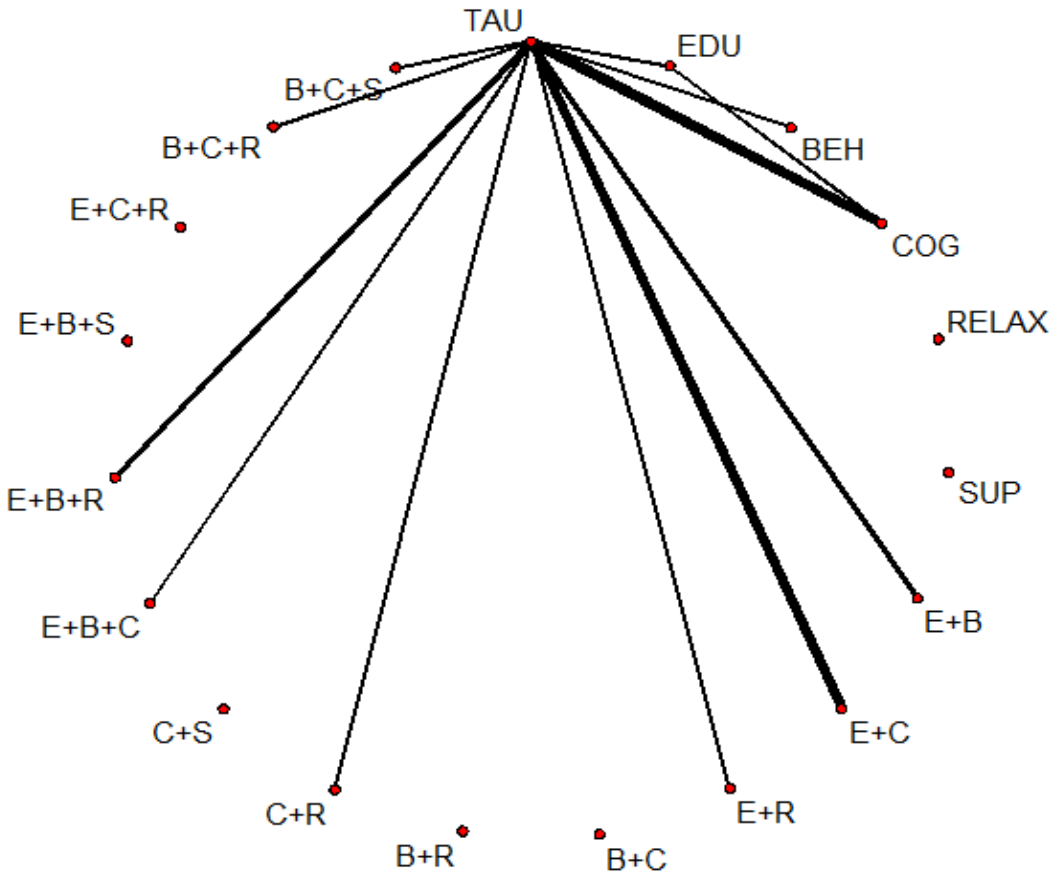




bristol.ac.uk

Trial	Modality	Delivered	Setting	Dose	Components
Black	Individual	Psychologist	Home	4 hours	BEH
Burgess	Individual	Nurse	Home	2 to 3 sessions	BEH, COG, SUP
Van-Elderen	Mixed	Not specified	Combination	2 sessions in hospital. 3 hrs(group)	COG, EDU
ENRICHD	Mixed	Not specified	Not clear	18 hrs	COG
HofmanBang	Individual	Not specified	Hospital	Not clear	BEH, EDU, REL
Johnston	Individual	Nurse	Not clear	3 to 4 hours	(1) COG, EDU (2) COG, EDU
Jones	Mixed	Not specified	Not clear	14 hrs	COG, REL
Lewin	Individual	Not specified	Home	6 weeks	BEH, EDU, REL
Stern	Group	Not specified	Not clear	13hrs	(1) BEH (2) COG
Thompson	Individual	Not specified	Not clear	2hrs	COG
Toobert	Group	'Therapist'	Home & retreat	7day retreat + follow-up	BEH, EDU, REL, SUP

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.

University of BRISTOL

Limitations

- 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
- A priori specification of characteristic, which may not be reported, might not be independent or never used on its own.
- Interventions are not only source of complexity
  - E.g. Interaction of intervention with setting or facilitator could also be considered.

References

- Ioannidis J, Patsopoulos N, Rothstein H. Reasons or excuses for avoiding meta-analysis in forest plots BMJ 2008; 336 :1413 doi:10.1136/bmj.a117
- Higgins J, López-López J, Becker B*, et al* Synthesising quantitative evidence in systematic reviews of complex health interventions *BMJ Global Health* 2019;**4:**e000858.
- Campbell M, McKenzie J, Sowden A, et al. Synthesis without meta-analysis (SWiM) in systematic reviews: reporting guideline BMJ 2020; 368 :l6890 doi:10.1136/bmj.l6890
- Lorenc T, Felix L, Petticrew, M. et al. Meta-analysis, complexity, and heterogeneity: a qualitative interview study of researchers' methodological values and practices. Syst Rev 5, 192 (2016).
- Welton N, Caldwell D, Adamopoulos E, Vedhara K. Mixed treatment comparison meta-analysis of complex interventions: psychological interventions in coronary heart disease. Am J Epidemiol. 2009;169(9):1158-65.
- Guise JM, Chang C, Viswanathan M, et al. Systematic Reviews of Complex Multicomponent Health Care Interventions [Internet]. Rockville (MD): *Agency for Healthcare Research and Quality* (US); 2014 Mar.
- Melendez-Torres, G.J., Bonell, C. & Thomas, J. Emergent approaches to the meta-analysis of multiple heterogeneous complex interventions. BMC Med Res Methodol 15, 47 (2015).
- Chaimani A, Caldwell D, Li T, Higgins J, Salanti G. Chapter 11: Undertaking network meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.2 (updated February 2021). Cochrane, 2021.
- Petticrew et al. Synthesizing evidence on complex interventions: how meta-analytical, qualitative, and mixed-method approaches can contribute. *JCE*. 2013: 66; 1230-1243

Copyright attributions for images on slides 12 and 21

- https://commons.wikimedia.org/wiki/File:Nordic_Walkers_near_Ilkley_020.JPG Nordic Walkers near Ilkley, northern England. 10 February 2010. This file is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license.
- https://commons.wikimedia.org/wiki/File:Spin_Cycle_Indoor_Cycling_Class_at_a_Gym.JPG Spin Cycle Indoor Cycling Class at a Gym. Photo by www.localfitness.com.au. 2009. This file is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license.
- https://www.flickr.com/photos/localmattersorg/10038526116/in/photolist-gi6vVq-gi6njV-gi6aZv-gi6bDj-gi5RXi-gi5CmF-gi6eXn-gi4D4r-gi4i3U-gi555u-gi4MrZ-gi5vUp-8xeCNS-nntN7s-tTkNu8-sWjCZC-8BenPj-8Bbg9R-Rt3KsA Attribution-NonCommercial 2.0 Generic (CC BY-NC 2.0). No changes made to the original file by #Cooking Matters at Faith Mission - Chicken Lesson
- https://www.flickr.com/photos/vvonstruen/6786472758/ Attribution-NonCommercial-ShareAlike 2.0 Generic (CC BY-NC-SA 2.0). No changes were made to the original image by Susan von Struensee.
- https://www.flickr.com/photos/synergybyjasmine/6808397021/ Attribution-NoDerivs 2.0 Generic (CC BY-ND 2.0). yoga-triangle-pose-synergybyjasmine. By Jasmine Kaloudis. No changes made to original image.
- https://www.flickr.com/photos/83905817@N08/7676623576/ Attribution-NonCommercial-ShareAlike 2.0 Generic (CC BY-NC-SA 2.0), No changes made to original file by RelaxingMusic.
- https://www.flickr.com/photos/queensucanada/26869936161/ Attribution-NonCommercial-NoDerivs 2.0 Generic (CC BY-NC-ND 2.0) FOE Revive & Thrive. No changes made to original image by Queen's University.
- https://www.flickr.com/photos/pennstatelive/33688898478/. Abington opioid counsellor. Attribution-NonCommercial-NoDerivs 2.0 Generic (CC BY-NC-ND 2.0). No changes made to original image by Penn State University.