Psychological Medicine

cambridge.org/psm

Invited Letter Rejoinder

Cite this article: Murray SB, Quintana DS, Loeb KL, Griffiths S, Crosby RD, Le Grange D (2019). Meta-analysis misunderstood: a cautionary tale in interpreting meta-analytic findings. *Psychological Medicine* **49**, 699–700. https://doi.org/10.1017/S0033291718003501

First published online: 21 December 2018

Key words:
Anorexia nervosa; eating disorders; meta-analysis; treatment outcome

Author for correspondence:
Stuart B. Murray, E-mail: drstuartmurray@gmail.com, stuart.murray@ucsf.edu

© Cambridge University Press 2018

CAMBRIDGE UNIVERSITY PRESS

Meta-analysis misunderstood: a cautionary tale in interpreting meta-analytic findings

Stuart B. Murray[1], Daniel S. Quintana[2], Katharine L. Loeb[3], Scott Griffiths[4], Ross D. Crosby[5] and Daniel Le Grange[1,6]

[1]Department of Psychiatry, University of California, San Francisco, CA, USA; [2]NORMENT, KG Jebsen Centre for Psychosis Research, Division of Mental Health and Addiction, University of Oslo, and Oslo University Hospital, Oslo, Norway; [3]School of Psychology, Fairleigh Dickinson University, Teaneck, NJ, USA; [4]Melbourne School of Psychological Sciences, University of Melbourne, Melbourne, VIC, Australia; [5]Sanford Research and University of North Dakota School of Medicine and Health Sciences, Fargo, ND, USA and [6]Department of Psychiatry and Behavioral Neuroscience, The University of Chicago, IL, USA (Emeritus)

Scientific inquiry is a continually evolving, shared enterprise that is dependent upon rigorous and well-reasoned discourse. Perhaps now more than ever, scrutiny around scientific methods and the validity of reported findings ought to be etched firmly into the fabric of the meta-science of clinical inquiry. It is with this in mind that we welcome discourse around our recent meta-analysis that focused on the delineation dimensions of anorexia nervosa (AN) symptoms in response to existing treatments (Murray *et al.*, in press).

Extending from our recent commentary framing the importance of deconstructing remission indices in AN (Murray *et al.*, 2018), this meta-analysis was undertaken to examine potential discrepancies between weight and cognitive symptoms of AN in response to existing treatments. The distinction between these two symptom domains is important because clinical and empirical observations suggest that there is often a disconnect – in timing if not in mechanisms – between pathways to recovery in each. Our research question was directed at treatment studies in general, and not on a particular intervention(s), and we relied on the original investigators' designation of specialized *v.* comparator modalities. Our analyses address these interacting sets of questions, that is, how do specialized *v.* comparator interventions fare in achieving positive change in weight *v.* psychological symptoms for AN? Findings reveal differential patterns of outcome between these two symptom domains, with implications for targeting specific mechanisms of AN pathology with appropriately-timed precision interventions. Study aims and results are consistent with recent research applying the transdiagnostic theory of core psychological mechanisms of eating disorder pathology to the treatment of AN (Fairburn *et al.*, 2013).

Meta-analysis is a powerful methodology and enjoys a commanding position in the hierarchy of evidence (Harbour and Miller, 2001). It aims to systematically and transparently draw together and evaluate research relating to a common scientific question. Important questions in medicine are typically studied more than once, in more than one setting, and with more than one proposed solution. In fact, there is no illness in the field of medicine for which the evidence base consists of only one treatment, at one prescribed dose, or for one standardized duration. This is certainly true of AN, and meta-analysis provides a potent methodology to meaningfully synthesize data across studies.

Notwithstanding, the commentators put forth a series of well-articulated criticisms of a hypothetical meta-analysis – one that broadly assesses the efficacy of treatments for AN, which would have undoubtedly limited such a hypothetical study's ability to draw reasonable conclusions had it applied our methodology to that goal. As a reminder, our research question was *not* to index the efficacy of treatments for AN. Instead, and as outlined in our introduction, our primary aim was to delineate potentially discrepant dimensions of response to existing treatments for AN according to weight- and psychological-based outcomes. As we have discussed elsewhere (Murray *et al.*, 2018), common methods for reporting outcomes in treatment trials for AN have either registered outcome as a function of weight alone, or as a function of a categorical grouping which combines some variation of weight and psychological outcome scores. Neither of these methods provides insight into the individual trajectories of these important dimensions of treatment response, which is essential to parse out as we move towards precision treatment efforts. To address *our* specific research goal in this meta-analysis, it is not necessary to restrict analyses to include only those studies investigating similar treatment approaches. Moreover, the moderator analyses conducted in our study were not designed to assess which treatment worked better for whom. Instead, these analyses were conducted to assess whether the delineated dimensions of treatment response, i.e. weight *v.* psychological outcomes differed according to study or patient characteristics.

Thus, it is important that the primary aim as well as the findings of this meta-analysis are not misunderstood or misinterpreted. Like any study, this meta-analysis is only able to address

https://doi.org/10.1017/S0033291718003501 Published online by Cambridge University Press

the research question for which it was designed, and criticisms that its methods are incompatible with alternative research questions are self-evident. Certainly, and more broadly, Lock and his colleagues astutely point out important limitations for meta-analyses that aim to assess the efficacy of treatments for AN, and their remarks appropriately serve as cautionary considerations for those researchers wishing to undertake such a study. Along with this line of discourse, for instance, we would concur with their sentiment that to 'include all studies one might find in a literature search is highly problematic', and we would concur that such a strategy can lead to spurious findings. Ultimately, it is the research question that determines the inclusion/exclusion criteria of such studies. Therefore, it is with this concern in mind that we exercised strict inclusion criteria for studies deemed suitable to contribute data to our specific research question. Indeed, our selection criteria systematically eliminated 99% of the initially generated literature, and two-thirds of the subsequent seemingly-eligible studies. We also concur with Lock and his colleagues' assertion that to detect a moderate between-group treatment effect size with 80% power†[1,2], studies with over 50 participants per group would be ideal. However, given the well-explicated challenges around patient recruitment in treatment trials for AN (Halmi *et al*., 2005; Halmi, 2008), most existing studies have not met this criterion. This relates to our finding of an elevated risk of bias in randomized controlled treatment trials of AN. In fact, in a recent meta-analysis of family-based treatment for AN conducted by the commentators, 83% of the included studies fell short of this benchmark (Couturier, Kimber and Szatmari, 2013).

Why does all this matter beyond considerations of clarifying our particular meta-analytic findings? It is important that meta-analytic methods and results are not misinterpreted or discredited on the basis of a question that was *not* asked. Such misinterpretation can hinder the field's ability to use good science to strategically shape advocacy efforts and public policy support for eating disorder treatment and funding (Roberto and Brownell, 2017). As the commentators rightfully state, 'Meta-analyses are necessary to determine whether consensus has been reached on a particular research question, thus either encouraging or discouraging further research on that question, and providing the evidence base for clinical decision-making.' In the case of our meta-analysis, an accurate appraisal of our specific research question, methods and results can make the difference between generating novel scientific hypotheses and strategies relating to treatment response and mechanisms in AN, *v.* eliciting hopelessness about AN outcomes among researchers, patients, caregivers, and clinicians alike.

Note

[1] Also related to power, the commentators correctly point out that it would be preferable to account systematically for the effect of time. We chose to use EOT as defined in the original articles rather than run analyses separately for different lengths of treatment, or evaluate length of treatment as a moderator, as we would have very limited power for detecting these effects.

[2] Similarly, the commentators also note that the confidence intervals for weight outcomes at EOT ($p = 0.006$) and follow-up ($p = 0.15$) are largely overlapping, and that it is very unlikely that these treatment effects differ from each other. We made no claims that these effects were different; in fact, we formally tested the difference between these effects, reporting no significant difference between weight outcomes at EOT and follow-up ($p = 0.35$). We agree that studies are most likely underpowered to detect treatment effects at follow-up.

References

Couturier J, Kimber M and Szatmari P (2013) Efficacy of family-based treatment for adolescents with eating disorders: a systematic review and meta-analysis. *International Journal of Eating Disorders* **46**, 3–11.

Fairburn CG, Cooper Z, Doll HA, O'Connor ME, Palmer RL and Dalle Grave R (2013) Enhanced cognitive behaviour therapy for adults with anorexia nervosa: a UK-Italy study. *Behaviour Research and Therapy* **51**, 2–8.

Halmi KA (2008) The perplexities of conducting randomized, double-blind, placebo-controlled treatment trials in anorexia nervosa patients. *American Journal of Psychiatry* **165**, 1227–1228.

Halmi KA, Agras WS, Crow S, Mitchell J, Wilson GT, Bryson SW and Kraemer HC (2005) Predictors of treatment acceptance and completion in anorexia nervosa: implications for future study designs. *Archives of General Psychiatry* **62**, 776–781.

Harbour R and Miller J (2001) A new system for grading recommendations in evidence based guidelines. *British Medical Journal* **323**, 334–336.

Murray SB, Loeb KL and Le Grange D (2018) Treatment outcome reporting in aNnorexia nervosa: time for a paradigm shift? *Journal of Eating Disorders* **6**, 10.

Murray SB, Quintana DS, Loeb KL, Griffiths S and Le Grange D (in press) Treatment outcomes for anorexia nervosa: a systematic review and meta-analysis of randomized controlled trials. *Psychological Medicine*. doi: 10.1017/S0033291718002088 [Epub ahead of print].

Roberto CA and Brownell KD (2017) Strategic science for eating disorders research and policy impact. *International Journal of Eating Disorders* **50**, 312–314.

†The notes appear after the main text.

https://doi.org/10.1017/S0033291718003501 Published online by Cambridge University Press