Epidemiology and Psychiatric Sciences

cambridge.org/eps

The waiting list is an inadequate benchmark for estimating the effectiveness of psychotherapy for depression

Ioana A. Cristea[1,2]

[1]Department of Clinical Psychology and Psychotherapy, Babes-Bolyai University, Cluj-Napoca, Romania and [2]Department of Clinical, Neuro and Developmental Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Commentary to Special Article

Cite this article: Cristea IA (2019). The waiting list is an inadequate benchmark for estimating the effectiveness of psychotherapy for depression. *Epidemiology and Psychiatric Sciences* **28**, 278–279. https://doi.org/10.1017/S2045796018000665

Received: 23 October 2018
Accepted: 26 October 2018
First published online: 27 November 2018

Key words:
Psychotherapy; Depression; Randomised Controlled Trials; Systematic Reviews

Author for correspondence:
Ioana A. Cristea, E-mail: ioana.cristea@ubbcluj.ro

In a recent meta-analysis entitled 'Was Eysenck right after all?', Cuijpers *et al.* (2018) propose a sobering reassessment of the effects of psychotherapy for adult depression. Their approach consists of a series of sequential sensitivity analyses adjusting the pooled effect size of psychotherapies for biasing factors, such as the use of waiting list (WL) arms as controls, trial risk of bias and publication bias. They evidence an overall effectiveness of psychotherapies for depression, though with more modest estimates, reduced to half (Hedges' $g = 0.31$) after excluding studies affected by biases. Some aspects of the methodology, like the use of the trim and fill to provide an estimate deemed free from publication bias, are less than ideal (Peters *et al.*, 2007), but the overall findings have a familiar resonance. The notion that the effects of psychotherapy for depression are overestimated due to biases is supported by previous similar work from the same research group (Cuijpers *et al.*, 2016), as well as by independent research (Dragioti *et al.*, 2017).

In reply, Munder *et al.* (2019) challenge this assessment by proposing a thought-provoking re-analysis. While Cuijpers *et al.* (2018) appeal to Eysenck's famous dictum about psychotherapy not being effective seemed mostly rhetoric, Munder and colleagues opt for an almost verbatim interpretation, focusing their entire rebuttal on establishing whether this historical claim is upheld. They object to several methodological choices in the original meta-analysis, but the crux of their case focuses on trials with a WL arm. Claiming that WL can be beneficial or at least not harmful, Munder *et al.* (2019) go on to conclude that it represents an appropriate control group by which to weigh the effectiveness of psychotherapy. Consequently, they recalculate effect estimates contrasting psychotherapy with WL and report a reassuring $g = 0.71$, double the effect in the original meta-analysis.

The plea for WL-controlled trials is in stark contrast to the negative attention this type of control has garnered over the recent years. In an oft-cited network meta-analysis, Furukawa *et al.* (2014) branded the WL a 'nocebo', after showing that the odds of response for patients in the psychological placebo, and, more strikingly, even in the no treatment (NT) control groups were significantly higher than for those in the WL groups. Munder *et al.* (2019) dispute these results as not providing 'persuasive evidence' (p. 3) that WL is an inappropriate control, but their main refuting argument is a conjectural *post hoc ergo propter hoc*. They pinpoint factors, such as the inclusion of studies on participants who were mildly or moderately depressed or not seeking treatment, which might have confounded the comparison between NT and WL and may consequently account for the observed differences. Evidently, this is a possibility, but until demonstrated empirically, on actual data, it remains solely a supposition.

Moreover, converging evidence that WL is inferior to other types of control arms comes from a network meta-analysis involving main author Thomas Munder (Barth *et al.*, 2013), as well as from another meta-analysis (Khan *et al.*, 2012), both showing diminished response on depressive symptoms for WL participants, as well as significant inferiority to other control conditions, such as treatment as usual (TAU) or placebo. Undoubtedly, all meta-analytic methods have limitations, and underscoring them, as Munder and colleagues do, is important. Yet these general limitations do not obliterate the consistent finding that not only are the effects of active psychotherapies disproportionately higher when contrasted to WL than to other control groups, but also its corollary that WL is conducive of worse outcomes than other control conditions.

Furthermore, other lines of research questioned the adequacy of WL control arms. One meta-analysis (Palpacuer *et al.*, 2017) with a meta-regression analysis showed that the effects of psychotherapies for depression compared to WL were rendered non-significant after controlling for non-specific factors, such as recruitment method, provenience, number of treatment sessions, length of follow-up and researcher allegiance. Another study (Cooper and Conklin, 2015) showed that in trials of psychotherapy for depression, inactive control conditions (a category inclusive of WL and placebo, but not of TAU) were associated with higher overall drop-out rates than active ones. As a minimum, this implies that a relevant proportion

CAMBRIDGE UNIVERSITY PRESS

of patients is not interested or willing to remain in an WL-controlled trial. Relatedly, disappointment about being randomised to the control group has been repeatedly reported in intervention research, and related to participant drop-out (Lindström *et al.*, 2010; Skingley *et al.*, 2014).

These findings add to the fact that psychotherapy trials with WL arms are in conspicuously large numbers (Barth *et al.*, 2013; Cuijpers *et al.*, 2016; Dragioti *et al.*, 2017; Palpacuer *et al.*, 2017). Investigators might consider only opting for these designs on a case-by-case basis, after a careful and judicious cost-benefit analysis, weighing both patient benefit and potential harms, along with trial validity. WL-controlled trials should probably not be routine, at the very least in a field as saturated with research as psychotherapy for depression. If publication bias could be taken out of the equation and we could reasonably assume all conducted trials are published, pitting new interventions against WLs could have the benefit of screening out ineffective interventions (i.e., that don't outperform WL) before they are tested in large-scale trials, preventing the squandering of resources on treatments that don't work (Ioannidis, 2016). With publication bias pervasive in clinical research, the informational value of WL-controlled trials is limited. Published trials comparing psychological interventions to WL systematically produce very large effects (Cuijpers and Cristea, 2016; Fodor *et al.*, 2018). Patients extract little benefit from being on a WL, and other, more beneficial, control arms are available (e.g., TAU). In this regard, several recent trials employed ingenious recruitment strategies aimed at maximising the time patients spend on a WL for active treatment by randomising them to low-intensity interventions during the waiting period (Lovell *et al.*, 2017). Equipoise-stratified designs (Lavori *et al.*, 2001; Shalev *et al.*, 2012), allowing participants to refuse undesired treatment options and still be randomly assigned to the other arms, could represent other options to counteract participant loss and the ensuing selection bias at recruitment, in those cases where a WL control is deemed necessary by investigators, or where other types of control arms are difficult to implement or pose other risks. Researchers should routinely report the uptake of delayed treatment for participants on the WL, as well as changes in the primary and secondary outcomes for this group. If patients who receive treatment after the WL period experience less improvement than those receiving treatment immediately after recruitment, this would provide further clues regarding the possible iatrogenic effects of this type of control condition.

Ultimately, the debate over whether Eysenck was right or not in a contention dating more than 50 years back has a grating academic resonance. This matter is entirely inconsequential for patients, caregivers or clinicians looking into systematic reviews and meta-analyses for guidance in selecting a course of treatment. The emphasis should be on up-to-date evidence as to what types of control arms are the most advantageous options in terms of maximising patient welfare and the internal and external validity of a trial.

Author ORCIDs. Ioana A. Cristea 0000-0002-9854-7076.

Acknowledgement. None.

Financial support. IAC is supported by a grant from the Romanian Ministery of Research and Innovation, CNCS-UEFISCDI (project number PN-III-P1-1.1-TE-2016-1054). The funder had no role in the preparation of the manuscript or decision to submit.

Conflict of interest. The author has completed the Unified Competing Interest form at http://www.icmje.org/coi_disclosure.pdf (available upon request from the corresponding author) and has nothing to declare.

References

Barth J, Munder T, Gerger H, Nuesch E, Trelle S, Znoj H, Juni P and Cuijpers P (2013) Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. *PLoS Medicine* **10**, e1001454.

Cooper AA and Conklin LR (2015) Dropout from individual psychotherapy for major depression: a meta-analysis of randomized clinical trials. *Clinical Psychology Review* **40**, 57–65.

Cuijpers P and Cristea IA (2016) How to prove that your therapy is effective, even when it is not: a guideline. *Epidemiology and Psychiatric Sciences* **25**, 428–435.

Cuijpers P, Cristea IA, Karyotaki E, Reijnders M and Huibers MJ (2016) How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. *World Psychiatry* **15**, 245–258.

Cuijpers P, Karyotaki E, Reijnders M and Ebert DD (2018) Was Eysenck right after all? A reassessment of the effects of psychotherapy for adult depression. *Epidemiology and Psychiatric Sciences* 1–10. doi: 10.1017/S2045796018000057.

Dragioti E, Karathanos V, Gerdle B and Evangelou E (2017) Does psychotherapy work? An umbrella review of meta-analyses of randomized controlled trials. *Acta Psychiatrica Scandinavica* **136**, 236–246.

Fodor LA, Cotet CD, Cuijpers P, Szamoskozi S, David D and Cristea IA (2018) The effectiveness of virtual reality based interventions for symptoms of anxiety and depression: a meta-analysis. *Scientific Reports* **8**, 10323.

Furukawa TA, Noma H, Caldwell DM, Honyashiki M, Shinohara K, Imai H, Chen P, Hunot V and Churchill R (2014) Waiting list may be a nocebo condition in psychotherapy trials: a contribution from network meta-analysis. *Acta Psychiatrica Scandinavica* **130**, 181–192.

Ioannidis JP (2016) Most psychotherapies do not really work, but those that might work should be assessed in biased studies. *Epidemiology and Psychiatric Sciences* **25**, 436–438.

Khan A, Faucett J, Lichtenberg P, Kirsch I and Brown WA (2012) A systematic review of comparative efficacy of treatments and controls for depression. *PLoS ONE* **7**, e41778.

Lavori PW, Rush AJ, Wisniewski SR, Alpert J, Fava M, Kupfer DJ, Nierenberg A, Quitkin FM, Sackeim HA, Thase ME and Trivedi M (2001) Strengthening clinical effectiveness trials: equipoise-stratified randomization. *Biological Psychiatry* **50**, 792–801.

Lindström D, Sundberg-Petersson I, Adami J and Tönnesen H (2010) Disappointment and drop-out rate after being allocated to control group in a smoking cessation trial. *Contemporary Clinical Trials* **31**, 22–26.

Lovell K, Bower P, Gellatly J, Byford S, Bee P, McMillan D, Arundel C, Gilbody S, Gega L, Hardy G, Reynolds S, Barkham M, Mottram P, Lidbetter N, Pedley R, Molle J, Peckham E, Knopp-Hoffer J, Price O, Connell J, Heslin M, Foley C, Plummer F and Roberts C (2017) Low-intensity cognitive-behaviour therapy interventions for obsessive-compulsive disorder compared to waiting list for therapist-led cognitive-behaviour therapy: 3-arm randomised controlled trial of clinical effectiveness. *PLoS Medicine* **14**, e1002337.

Munder T, Flückiger C, Leichsenring F, Abbass AA, Hilsenroth MJ, Luyten P, Rabung S, Steinert C, Wampold BE (2019) Is psychotherapy effective? A re-analysis of treatments for depression. *Epidemiology and Psychiatric Sciences* 1–7. doi: 10.1017/S2045796018000355.

Palpacuer C, Gallet L, Drapier D, Reymann JM, Falissard B and Naudet F (2017) Specific and non-specific effects of psychotherapeutic interventions for depression: results from a meta-analysis of 84 studies. *Journal of Psychiatric Research* **87**, 95–104.

Peters JL, Sutton AJ, Jones DR, Abrams KR and Rushton L (2007) Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Statistics in Medicine* **26**, 4544–4562.

Shalev AY, Ankri Y, Israeli-Shalev Y, Peleg T, Adessky R and Freedman S (2012) Prevention of posttraumatic stress disorder by early treatment: results from the Jerusalem trauma outreach and prevention study. *Archives of General Psychiatry* **69**, 166–176.

Skingley A, Bungay H, Clift S and Warden J (2014) Experiences of being a control group: lessons from a UK-based randomized controlled trial of group singing as a health promotion initiative for older people. *Health Promotion International* **29**, 751–758.