Reviewing Psychological Treatments for Adult Depression

Steven D Hollon, PhD; Pim Cuijpers, PhD

1Professor of Psychology and Human Development and Professor of Psychiatry, Vanderbilt University, Nashville, Tennessee. Correspondence: Department of Psychology, Vanderbilt University, 306 Wilson Hall, Nashville, TN 37203; steven.d.hollon@vanderbilt.edu.
2Professor of Clinical Psychology, VU University, Amsterdam, the Netherlands.

The 2 articles that comprise this In Review1,2 use different methods to review their respective literatures. In the first, Dr Pim Cuijpers and colleagues1 conduct a meta-analysis of RCTs involving CBT in the treatment of adult depression and find CBT superior to various control conditions, comparable to alternative interventions, and better than medication alone when added in combination. In the second, Dr Jan Spijker and colleagues2 chose to do a qualitative review, given the limited number of relevant RCTs that focused on chronic major depressive disorder, and concluded that treatment was superior to its absence but that multiple successive interventions may be necessary. Both sets of conclusions seem reasonable given the available data.

We come from different traditions and have different levels of comfort with the approaches used in the reviews. Dr Hollon has spent his career doing RCTs and prefers reviews that emphasize the findings from well-conducted studies. Dr Cuijpers is an experienced quantitative researcher who deals with concerns regarding study quality, using meta-regressions to explore the study characteristics that influence the estimated effects. As we concur that both reviews were nicely implemented examples of their genres, we chose to use this editorial to comment on differences arising from our different perspectives.

Do quantitative reviews obscure differences between different treatments? Dr Cuijpers and colleagues1 concluded, based on their meta-analysis, that CBT was superior to control conditions and not different than alternative interventions. Hollon and Ponniah3 arrived at similar but not identical conclusions in a recent qualitative review that required a minimum number of well-conducted studies to draw an inference regarding treatment impact. Using the criteria adapted from US Food and Drug Administration guidelines to identify empirically supported treatments, a treatment was said to be efficacious if it was better than its absence in at least 2 studies conducted by different research groups (possibly if only one) and specific if it was superior conditions that controlled for the generic effects of simply going into treatment.4 Using these criteria, Hollon and Ponniah3 concluded that CBT was efficacious and specific, and came to the same conclusions regarding IPT, BA, PST, and ADs. Where they differed from Dr Cuijpers and colleagues1 was in concluding that both psychodynamic psychotherapy and EXP were only possibly efficacious (owing to the paucity of supportive trials in fully clinical populations) and drew no conclusions at all regarding supportive psychotherapy (to the extent that it differed from EXP). Quantitative reviews that estimate the average magnitude of effect across all published studies in a literature typically find no differences among bona fide therapies, whereas qualitative reviews that look for high-quality studies sometimes do. For example, ADs represent the current standard of treatment for depression and pill-placebos a particularly rigorous control. We can find studies in which the psychotherapy of interest was as efficacious as medications and superior pill-placebo for CBT,5,6 IPT,7 BA,8 and PST,9 but not for other kinds of psychotherapies.

Abbreviations

AD antidepressant
BA behavioural activation
CAU care as usual
CBT cognitive-behavioural therapy
ES effect size
EXP experiential psychotherapy
IPT interpersonal psychotherapy
PST problem-solving therapy
RCT randomized controlled trial

WWW.TheCJP.CA
If we were to choose among the different psychotherapies, do we choose one that has been shown to be at least the equal of the AD (when it is superior to the pill-placebo)? The typical quantitative review suggests that this does not matter. The 2 approaches especially diverge when translated into treatment guidelines. Although quantitative reviews suggest that all bona fide treatments are equally effective, guidelines tend to emphasize the best-supported treatments. If no comparisons with medication or other therapies are available, both approaches agree that such treatments should not be used in clinical practice.

Does quality of implementation matter? Closely related is the issue of quality of implementation. Dr Cuijpers and colleagues addressed study quality and found that the inclusion of lower-quality studies inflates the size of the psychotherapy effect. This was a nice methodological feature that highlights the risk of interpreting ESs in an uncritical fashion. However, the features they considered (adequate generation of allocation sequence, concealment of allocation to conditions, prevention of knowledge of the allocated intervention to assessors to outcomes, and dealing with incomplete data) all concern the confidence one can have in attributing the effects observed to the experimental manipulation (internal validity). Also important is whether the interventions were implemented in the manner intended. Quality of implementation can vary markedly across the interventions were implemented in the manner intended. For example, CBT was less efficacious than medication and no better than pill-placebo among patients, with more severe depressions in both the National Institute of Mental Health Treatment of Depression patients, with more severe depressions in both the National Institute of Mental Health Treatment of Depression compared with their milder depressions, whereas the latter 2 trials used experienced CBT provided with session-by-session supervision. It is not clear that it is reasonable to lump all 4 trials together. In a meta-analysis of the association between study quality and outcome in treatments of depression, no association was found between ES and the use of a manual, the training of therapists, and whether treatment integrity was checked in the trial. However, it is not clear that the factors included in those meta-regressions would have captured the differences described above.

Does severity moderate outcome? The 2 trials that assessed pretreatment severity as a moderating variable found evidence of specificity only among patients with more severe depressions. This relation also seems to hold for medication treatment. Simply put, it appears that treatment specificity may only matter when dealing with patients with more severe depressions. For most patients, anything is better than nothing (most treatments work better than their absence) and for most patients nothing is better than anything else (it does not matter what you do so long as you do something). That may account for why supportive psychotherapy fares as well as it does in quantitative reviews; if most patients do not require an intervention with specific effects to get better than nonspecific treatments are likely not to differ from treatments with specific effects when averaged across the literature. In that regard it is of interest to note that including mean pretreatment severity at the study level in meta-regressions failed to detect the moderating effect of severity found in patient-level data. Moreover, is it possible that we are missing other instances of moderation by relying on study-level indices rather than patient-level variation?

Is there variation in CAU? Dr Cuijpers and colleagues are to be congratulated for subdividing their control conditions into different categories; their meta-regressions indicated that ESs were larger for wait-lists (0.83) than for either CAU (0.59) or placebo controls (0.51). These categories likely reflect the very real distinction between determining whether a treatment works better than its absence (efficacy) and whether it works for reasons beyond the simple provision of generic treatment (specificity). Our only concern is with the greater heterogeneity in the placebo and especially the CAU comparisons. Our reading of the literature suggests that CAU varies greatly across settings. In some instances it involves little actual care, whereas in other settings it can be quite rich and approaches what is provided in alternative treatments. The same concern may be raised about the placebo–other category. It may be useful to specify the actual amount of care provided in these control conditions and test whether greater specification can reduce heterogeneity.

How important are enduring effects? In raising these points, we want to emphasize that we found much to like about both reviews and thought their methods were outstanding. We strongly concur with Dr Cuijpers and colleagues who emphasized the importance of enduring effects not found for medication. At its best, psychotherapy is no better than medications regarding the speed and magnitude of its effects and securing that effect depends on the quality of implementation (at least among patients with more severe depression). That said, the main advantage of CBT is that it appears to have an enduring effect not found for medications. In a chronically recurrent disorder such as depression, that is a very welcome effect that can reduce the cost of treatment.

In conclusion, regarding the 2 different approaches to writing reviews on treatments of depression, we can only conclude that both are needed, and both have a unique contribution to our understanding of treatments. The 2 In Review papers in this issue are excellent illustrations of that.

Acknowledgements
The Canadian Psychiatric Association proudly supports the In Review series by providing an honorarium to the authors.
References