The responsiveness of the different versions of the Hamilton Depression Scale

Per Bech

Psychiatric Research Unit, Psychiatric Centre North Zealand, Copenhagen University Hospital, Hillerød, Denmark

In clinical pharmacology, the evidence proving that a drug has a therapeutic effect in a specific medical condition is based on two major elements: superiority of the drug over placebo in randomized clinical trials within the medical condition under examination, and a systematic relationship between the dose of the drug and the magnitude of the response it elicits.

In their overview of antidepressants versus placebo, Khan and Brown (1) conclude that "no clear dose-response relationship has been established to date for most of the new antidepressants", while the superiority of the antidepressants over placebo in terms of effect size statistics is approximately 0.30, a level they find "less than impressive".

In their review of double-blind, placebo-controlled trials of antidepressants conducted from 1981 to 2008, Khan et al (2) observed that the Hamilton Depression Scale (HAM-D) had been used as an outcome scale in most of the trials. However, the HAM-D was actually used in two different versions, the 21-item version (HAM-D-21) and the 17-item version (HAM-D-17). The HAM-D-21 was used in one third of the trials and the HAM-D-17 in two thirds. Unfortunately, authors who use the HAM-D-21 rarely provide information about the results on the HAM-D-17. Khan and Brown (1) highlight now that the antidepressant-placebo difference seems to be higher in HAM-D-21 trials compared to the trials in which the HAM-D-17 has been used as outcome measure. This is a tautological finding in so far as the standard deviation of this difference is not available, which is the case for most of the trials reviewed by Khan et al (2).

Among the trials collected by Khan et al (2) it is possible, however, to identify ten publications in which the six-item HAM-D (HAM-D-6) is compared to both HAM-D-17 and HAM-D-21, or to HAM-D-28. The HAM-D-6 covers the core symptoms of depression: depressed mood, work and interests, guilt feelings, psychomotor retardation, psychic anxiety, and general somatic (fatigability). These six items have clinical and psychometric validity (3). In two of these ten trials, a dose-response relationship was investigated. Fabre et al (4) showed that sertraline was significantly superior to placebo at all three doses (50, 100, 200 mg daily) when using HAM-D-6, but only at 50 mg daily when using HAM-D-17. Liebowitz et al (5) showed that desvenlafaxine was superior to placebo at both 50 and 100 mg daily when using HAM-D-6, but only at 50 mg daily when using HAM-D-17.

An analysis of all placebo-controlled trials of desvenlafaxine showed that at doses of 200 or 400 mg daily the effect size was negative during the first week of treatment (superiority of placebo) when using HAM-D-17 but not when using the HAM-D-6, implying that the HAM-D-17 includes symptoms which might be side effects of the drug (6). In placebo-controlled trials of fluoxetine, over a dose range from 20 to 60 mg daily, the effect size using HAM-D-17 was approximately 0.30, but when using HAM-D-6 it was approximately 0.40 (3). For escitalopram, a dose of 10 mg daily obtained an effect size of 0.38 mg using HAM-D-6 and a dose of 20 mg daily gave an effect size of 0.61 (3).

Over the past decade, the goal when evaluating the effect of an antidepressant has been the event of remission rather than response (7). Remission in major depression is defined as a minimal level of the core symptoms of depression (7). The syndrome reflected by the HAM-D-6 is a unidimensional measure for specific drug targets, and a cut-off score below 5 indicates that the individual symptoms of the scale are only present to a very doubtful degree (remission).

Khan and Brown (1) refer to the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) as an example of a poor response to citalopram treatment. Indeed, when using the conventional definition of remission (HAM-D-17 <8), only approximately 36% of the patients achieved remission. However, when using a HAM-D-6 score of <5 as the definition of remission, approximately 45% of the patients in that study achieved remission ($p<0.001$) (8).

From a statistical point of view, failed trials are merely a consequence of insufficient power, as the inability to reject the null hypothesis is inherently associated with low statistical power. This has recently been illustrated in a re-analysis of a failed study which had used the HAM-D-17 to evaluate the effect of erythropoietin as augmentation in patients with treatment-resistant depression (9). By focusing on the HAM-D-6, fewer patients are needed to reject the null hypothesis.

Psychometrically, Khan and Brown (1) correctly focus on the use of the HAM-D-17 as the major factor for the "less than impressive" effect size of 0.30 and the lack of a dose-response relationship. However, their solution to go for a larger HAM-D version (HAM-D-21) is not justified. The solution is to go for the brief, clinically and psychometrically valid subscale (HAM-D-6).

The use of the HAM-D-6 as outcome measure in placebo-controlled clinical trials of antidepressants increases the effect size to 0.40, which is indicative of clinical significance. Using the HAM-D-6 as outcome measure, a dose-response relationship has been established for newer antidepressants

such as escitalopram and desvenlafaxine. Moreover, fewer patients are then needed to identify antidepressant effect in controlled trials, which has important ethical implications (fewer patients need to receive placebo).

In my opinion, we need to aim at establishing "dose-remission" rather than dose-response relationship in future trials of antidepressants. The HAM-D-6 contains the core symptoms of depression by which to define the event of remission.

References

1. Khan A, Brown WA. Antidepressants versus placebo in major depression: an overview. World Psychiatry 2015;14:294-300.
2. Khan A, Bhat A, Kolts R et al. Why has the antidepressant-placebo difference in antidepressant clinical trials diminished over the past three decades? CNS Neurosci Ther 2010;16:217-26.
3. Bech P. Is the antidepressive effect of second-generation antidepressants a myth? Psychol Med 2010;40:181-6.
4. Fabre LF, Abuzzahab FS, Amin M et al. Sertraline safety and efficacy in major depression: a double-blind fixed-dose comparison with placebo. Biol Psychiatry 1995; 38:592-602.
5. Liebowitz MR, Manley AL, Padmanabhan SK et al. Efficacy, safety, and tolerability of desvenlafaxine 50 mg/day and 100 mg/day in outpatients with major depressive disorder. Curr Med Res Opin 2008;24:1877-90.
6. Bech P, Boyer P, Germain JM et al. HAM-D17 and HAM-D6 sensitivity to change in relation to desvenlafaxine dose and baseline depression severity in major depressive disorder. Pharmacopsychiatry 2010;43:271-6.
7. Rush AJ, Kraemer HC, Sackeim HA et al. Report by the ACNP Task Force on response and remission in major depressive disorder. Neuropsychopharmacology 2006;31:1841-53.
8. Ostergaard SD, Bech P, Trivedi MH et al. Brief, unidimensional melancholia rating scales are highly sensitive to the effect of citalopram and may have biological validity: implications for the Research Domain Criteria (RDoC). J Affect Disord 2014;163:18-24.
9. Ostergaard SD, Bech P, Miskowiak KW. Fewer study participants needed to demonstrate superior antidepressant efficacy when using the Hamilton melancholia subscale (HAM-D6) as outcome measure. J Affect Disord (in press).

DOI 10.1002/wps.20248

What if a placebo effect explained all the activity of depression treatments?

Pim Cuijpers, Ioana A. Cristea

[1]Department of Clinical Psychology, Faculty of Psychology and Education, University of Amsterdam, Amsterdam, The Netherlands, [2]Department of Clinical Psychology and Psychotherapy, Babes Bolyai University, Cluj, Romania

Many randomized trials have shown that when depressed patients receive no active treatment, e.g. they are administered pill placebo, a large part of them improve anyway. This improvement can be partly explained by natural remission or by the patients' expectations that a treatment will have an effect on their problems (even when they receive pill placebo). The corollary is that many patients remit even when undergoing exotic therapies, such as Argentian tango, swimming with dolphins or horticulture (1-3).

This phenomenon makes it difficult to examine the additional effects of specific treatments. This is not only true for pharmacotherapy, but also for psychotherapies for depression. In a recent meta-analysis, we found that 62% of patients meeting criteria for major depression at baseline did no longer meet these criteria after treatment (4). But among the patients receiving only care-as-usual, 48% also no longer met criteria for major depressive disorder. So, therapists may think that more than 60% of patients get better because of the psychotherapy, while in fact the additional benefit of psychotherapy over usual care is only 14%. Khan and Brown (5) indicate that comparable outcomes take place for pharmacotherapy, with symptom reduction of about 40% with antidepressants and 30% with placebo. That is in line with Kline's conclusion in 1964 that "in the treatment of depression, one always has an ally in the fact that most depressions terminate in spontaneous remissions" (6).

Given this large proportion of patients who remit spontaneously, patients as well as therapists can easily be led into the idea that their treatment is highly successful, while in fact the effects of this treatment may be only moderate. This may also explain why the exotic treatments mentioned earlier are believed by some to be effective, while most clinicians would consider the specific effects of such treatments as not very credible. "But we see that patients get better" is a phrase that supporters of such therapies often use.

Due to the discrepancy between the relatively high rate of spontaneous remission and the low additional value of specific (pharmacological and psychological) treatments, several important issues arise. One question is whether these treatments do in fact have any effects. Of course, randomized trials show that pharmacotherapy and psychotherapy are effective for treating depression, with small effect sizes of 0.30 for antidepressants (5) and 0.25 for psychotherapies (7). But we also know that these effects are much higher when risk of bias is not taken into account. In fact, only the highest quality studies show such small effects, and only after publication bias has been adjusted for.

But suppose there is still a bias lingering in these trials. For example, since patients getting a placebo know that they are not receiving active medication because they experience no side effects, this breaks the blinding and serves to lower their expectations.