EDITORIAL

CrossMark

Possibly no baseline severity effect for antidepressants versus placebo but for antipsychotics. Why?

Stefan Leucht[1,2] · S. Z. Levine[3] · M. Samara[1] · A. Cipriani[4] · J. M. Davis[5,6] · T. A. Furukawa[7]

10 years ago, a meta-regression analysis [1] of 35 clinical trials submitted to the FDA found that the efficacy of antidepressants compared to placebo increases with baseline severity, and that a clinically significant effect defined as an effect size of at least 0.50 can only be expected in patients with a Hamilton Depression Rating Score of above 28. This was interpreted by Kirsch et al., to mean that antidepressants are efficacious for depression only for patients who are severely ill at baseline. This paper had a major impact on treatment guidelines and fueled a heated discussion about the usefulness of these agents [2].

One discussion was about the effect size threshold of 0.50. Many commonly used drugs in general medicine and psychotherapies have effect sizes below [3, 4]. But here we want to discuss that meta-regression of aggregate data (using averages of included trials) is limited by "ecological fallacy", where characteristics of groups may not necessarily apply to individuals. As the studies may differ in other aspects than just baseline severity (for example gender or age), the increasing drug-placebo difference with baseline severity may be confounded by these other factors. Epidemiologists generally follow-up ecological findings with studies of individual patients. Individual patient data (IPD) meta-analysis uses patient-level data and so can better control for such confounding factors. Due to the impact of Kirsch et al.'s findings, a number of such IPD analyses have subsequently been published with conflicting results. Fournier et al. [5] essentially confirmed Kirsch's finding in an IPD meta-analysis of six studies (718 participants overall) and found that patients with a HAM-D score below 23 had an effect size of less than 0.20 compared to placebo (a small effect size according to Cohen, $0.20 =$ small, $0.50 =$ medium, $\geq 0.80 =$ large). The major limitation may have been the inclusion of both studies with minor depression and major depression. By contrast, a large IPD meta-analysis by Gibbons et al. [6] of 37 trials (8477 patients) found no influence of baseline severity on treatment efficacy. They had access to all published and unpublished sponsor-conducted randomized-controlled studies of fluoxetine and venlafaxine. The inclusion of all trials from one company (irrespectively of the publication status) is a strength, because the major limitation of IPD meta-analysis is that usually not all relevant studies can be included. The only detail that may have been missing was an illustrative plot of antidepressants and placebo by baseline severity in addition to the statistical results. In another large IPD meta-analysis of 34 RCTs with 10,737 patients from the NEWMEDS registry Rabinowitz et al. [7] did also not detect a baseline severity effect, although some drugs were not "true" antidepressants (e.g., quetiapine) and only trials with a positive efficacy finding were included. Nor did we in a smaller sample of trials (6 RCTS with 2464 participants) in Japanese patients [8]. Thus, the majority and the largest IPD meta-analyses could not replicate the baseline severity effect described by Kirsch and Fournier. This would be an important clinical issue if there were guidelines which still follow Kirsch's initial report, although it should also be noted that effect sizes are generally not large (e.g., 0.3 in [6]), which is a general concern [3, 4].

✉ Stefan Leucht
Stefan.Leucht@tum.de

J. M. Davis
davisjm@uic.edu

[1] Department of Psychiatry and Psychotherapy, Technische Universität München, Klinikum rechts der Isar, Ismaningerstr. 22, 81675 Munich, Germany

[2] Department of Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

[3] University of Haifa, Haifa, Israel

[4] Department of Psychiatry, University of Oxford, Oxford, UK

[5] Psychiatric Institute, University of Illinois at Chicago, Chicago, IL, USA

[6] Maryland Psychiatric Research Center, Baltimore, MD, USA

[7] Department of Health Promotion and Human Behavior, Graduate School of Medicine/School of Public Health, Kyoto University, Kyoto, Japan

These findings regarding antidepressants contrast sharply with results on antipsychotic drugs where IPD analyses have more consistently confirmed baseline severity effects in placebo-controlled trials in people with schizophrenia and predominant positive symptoms [9, 10], in people with schizophrenia and predominant negative symptoms [10], in acute mania [11] and autism [12]. We note that most of these studies also addressed whether in the more mildly ill patients the difference between drug and placebo is still clinically meaningful—the other aspect of Kirsch et al.'s initial critique. In our analyses, we concluded that the numbers-needed-to-treat (NNT) may be low enough so that even more mildly ill patients with schizophrenia and mania may benefit sufficiently from antipsychotic drugs. But we also recommended that clinicians should wait longer before they initiate drug treatment to be sure about the diagnosis, that they should be more careful in dosing, and choose less side effect prone antipsychotic drugs for patients.

Where does the difference in the IPD analyses of major depressive disorder stem from? In our opinion, drug-placebo differences should naturally increase with baseline severity. For example, because an important placebo effect may only be present in mildly ill patients or because the severely ill patients may enable those with more "biological" forms of the disorder to benefit more from treatment. Interestingly, Furukawa et al. [13] did also not find a baseline severity effect in an IPD meta-analysis of cognitive behavior therapy of depression with an average effect size for the difference between CBT and pill placebo of 0.22 (95% 0.02–0.42). We can only speculate about the explanation of this discrepancy: antipsychotics are more efficacious than antidepressants in general, leaving more room for baseline severity effects in meta-regressions. For example, in our analysis of antipsychotics for people with schizophrenia with positive symptoms [10] and in acute mania [11], the mean effect sizes were approximately 0.6, while in a study on antidepressants [6] it was 0.30. There may also be more severely ill participants in RCTs on schizophrenia than in RCTs on depression. For example, the average PANSS at baseline of the three schizophrenia trials in Furukawa et al. [10] was 99 which roughly corresponds to a CGI of approximately 5.3 (in the markedly to severely ill range) [14], while in Furukawa et al. [8] on antidepressants, the mean HAM-D at baseline was 22.5 which corresponds to a CGI of 4.0–4.6 (in the moderately to markedly ill range) [15]. If the span between more severely ill patients and more mildly ill patients is wider, there may be more leeway for a significant correlation between baseline severity and drug-placebo differences. Depression may also be an even more heterogeneous disorder than schizophrenia where the positive symptoms hallucinations, delusions and thought disorder more clearly distinguish affected people from the general population than symptoms of depression. This greater heterogeneity may also imply unmeasured and therefore unknown factors that mask any baseline severity effect in depression trials. Finally, the side effects of antipsychotics are more severe and this may lead to more unblinding in antipsychotic trials and also to some extent play a role.

These are the most likely explanations but there is no definitive answer and the influence of baseline severity effect of antidepressants and CBT for major depressive disorder remains an enigma. We end with emphasizing that baseline severity is of course not the only point of discussion about the efficacy of antidepressants [16]. Publication bias, possible unblinding due to side effects and the subjectivity of the outcomes are other examples that still need scientific attention.

Compliance with ethical standards

References

1. Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ et al (2008) Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. PLoS Med 5:e45

2. Cipriani A, Geddes JR (2014) Placebo for depression: we need to improve the quality of scientific information but also reject too simplistic approaches or ideological nihilism. BMC Med 25:105

3. Leucht S, Hierl S, Kissling W, Davis JM (2010) Are psychiatric drugs more or less efficacious than general medicine medication? A review of meta-analyses. Br J Psychiatry. 200(2):97–106

4. Huhn M, Tardy M, Spineli LM, Kissling W, Forstl H et al (2014) Efficacy of pharmacotherapy and psychotherapy for adult psychiatric disorders: a systematic overview of meta-analyses. JAMA Psychiatry 71:706–715

5. Fournier JC, DeRubeis RJ, Hollon SD, Dimidjian S, Amsterdam JD et al (2010) Antidepressant drug effects and depression severity: a patient-level meta-analysis. JAMA 303:47–53

6. Gibbons RD, Hur K, Brown CH, Davis JM, Mann JJ (2012) Benefits from antidepressants: synthesis of 6-week patient-level outcomes from double-blind placebo-controlled randomized trials of fluoxetine and venlafaxine. Arch Gen Psychiatry 69:572–579

7. Rabinowitz J, Werbeloff N, Mandel FS, Menard F, Marangell L et al (2016) Initial depression severity and response to antidepressants v. placebo: patient-level data analysis from 34 randomised controlled trials. Br J Psychiatry 209:427–428

8. Furukawa TA, Maruo K, Noma H, Tanaka S, Imai H et al (2018) Initial severity of major depression and efficacy of new generation antidepressants: individual participant data meta-analysis. Acta Psychiatr Scand 137:450–458

9. Rabinowitz J, Werbeloff N, Caers I, Mandel FS, Stauffer V et al (2014) Determinants of antipsychotic response in schizophrenia: implications for practice and future clinical trials. J Clin Psychiatry 75:e308–e316

10. Furukawa TA, Levine SZ, Tanaka S, Goldberg Y, Samara M et al (2015) Initial severity of schizophrenia and efficacy of antipsychotics: participant-level meta-analysis of 6 placebo-controlled studies. JAMA Psychiatry 72:14–21

11. Samara MT, Goldberg Y, Levine SZ, Furukawa TA, Geddes JR et al (2017) Initial symptom severity of bipolar I disorder and the efficacy of olanzapine: a meta-analysis of individual participant data from five placebo-controlled studies. Lancet Psychiatry 4:859–867

12. Levine SZ, Kodesh A, Goldberg Y, Reichenberg A, Furukawa TA et al (2016) Initial severity and efficacy of risperidone in autism: results from the RUPP trial. Eur Psychiatry 32:16–20

13. Furukawa TA, Weitz ES, Tanaka S, Hollon SD, Hofmann SG et al (2017) Initial severity of depression and efficacy of cognitive-behavioural therapy: individual-participant data meta-analysis of pill-placebo-controlled trials. Br J Psychiatry 210:190–196

14. Leucht S, Kane JM, Kissling W, Hamann J, Etschel E et al (2005) What does the PANSS mean? SchizophrRes 79:231–238

15. Leucht S, Fennema H, Engel R, Kaspers-Janssen M, Lepping P et al (2013) What does the HAMD mean? J Affect Disord 148:243–248

16. Cipriani A, Furukawa TA, Salanti G, Chaimani A, Atkinson LZ et al (2018) Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. Lancet 391:1357–1366