Epidemiology and Psychiatric Sciences (2018), **27**, 240–241. © Cambridge University Press 2017
doi:10.1017/S2045796017000725

COMMENTARY TO SPECIAL ARTICLE

The clinical significance of drug-placebo differences

First published online 23 November 2017

Key words: Adverse effects, antidepressants, depression, depression, psychotropic drugs.

Commentary on: Cristea IA, Naudet F (2017). Defending psychiatry or defending the trivial effects of therapeutic interventions? A citation content analysis of an influential paper. *Epidemiology and Psychiatric Sciences*. doi: 10.1017/S2045796017000750.

In 2012, Leucht *et al*. (2012) published a highly-cited meta-meta-analysis in which they claimed that effect sizes for psychiatric drugs in which they concluded that 'psychiatric drugs were not generally less efficacious than other drugs,' along with the seemingly contradictory caveat that 'Any comparison of different outcomes in different diseases can only serve the purpose of a qualitative perspective. The increment of improvement by drug over placebo must be viewed in the context of the disease's seriousness, suffering induced, natural course, duration, outcomes, adverse events and societal values' (p.97). In an interesting follow-up to this paper, Cristea and Naudet (this issue) examined citations of the Leucht *et al.* article and concluded that most of them ignored Leucht *et al.*'s caveats and had used the citation to claim that the effects of psychiatric medications were similar to those of treatments in general medicine.

Whereas in 2012, Leucht and colleagues reported effects sizes for general medicine to be similar to those found in psychotropic medications, in 2015, they reported effect sizes in general medicine that are substantially larger than those for psychiatric drugs (Leucht *et al*. 2015). For example, the drug-placebo SMDs reported by Leucht *et al.* (2012) for antidepressants in the treatment of major depressive disorder are about 0.32. This is larger than the effect sizes reported by Leucht *et al.* (2015) for statins (SMD = 0.15) and aspirin (SMD = 0.12) for the prevention of cardiovascular events, but considerably smaller than the effect sizes reported for proton pump inhibitors (SMD = 1.39), oxycodone plus paracetamol (SMD = 1.04) and levodopa (SMD = 0.93) in treating gastric reflux, post-operative pain and Parkinson's disease, respectively. However, these data must be interpreted in light of Leucht *et al.*'s caution about comparisons between outcomes across different conditions. The more important question to answer is whether drug-placebo differences within particular disorders are clinically important. Cristea and Naudet make the claim that the effects of psychiatric treatments are 'trivial.' In this commentary, I consider whether that conclusion is justified.

Cristea and Naudet's conclusion that the effect of psychiatric drugs are trivial is based on the conventional classification that standardised mean differences (SMD) of 0.20 as small, those of 0.50 as a medium, and those of 0.80 as large (Cohen, 1988). However, Cohen proposed his cut-offs for small, medium and large effect sizes with 'invitations not to employ them if possible. The values chosen had no more reliable a basis than my own intuition' (p. 532). More meaningful criteria for clinical significance of drug-placebo difference have to be set independently for each disorder. How can this be done?

Researchers in the field of chronic pain reduction have reached a consensus on the clinical significance of treatment-induced pain reduction can be assessed within their field, and their method of so doing can provide a model for other fields (Farrar *et al.* 2001; Dworkin *et al.* 2005). The most common primary outcome measure in clinical trials of chronic pain is a 0–11 numerical rating scale (NRS) of pain intensity. To establish criteria for clinical significance, Farrar and colleagues compared NRS ratings with Patient Global Impression of Change scale, a 7-point scale, with endpoints of 'very much worse and very much improved' and 'very much worse,' with 'no change' as the midpoint. They decided in advance, that a score of 'much improved' would be their criterion for clinical significance, and established that this corresponded to a NRS reduction of 2 points or 30%. These, then, became conventionally accepted criteria for clinical improvement in chronic pain.

What would happen if this strategy were employed to evaluate the clinical significance of changes in symptoms of depression? The most commonly used scale for assessing levels of depression is the Hamilton Rating Scale for Depression (HRSD). Leucht *et al.* (2013) have provided the data needed to establish empirically derived criteria for the clinical significance of HRSD differences, using the approach adopted by chronic pain researchers (Moncrieff & Kirsch, 2015). These data indicate that 'much improved' is equivalent to a 14-point improvement

Address for correspondence: I. Kirsch, Program in Placebo Studies, Harvard Medical School, Boston, USA.
(Email: irvkirsch@gmail.com)

on the HRSD, minimally improved is equivalent to a 7-point difference, and the mean rating for 'no change' is 3-points improvement on the HRSD. These correspond to effect sizes of SMDs of 0.21 for 'no change,' 0.74 for 'minimally improved,' and 1.37 for 'much improved.' A 3-point improvement on the HRSD or effect size of 0.50 was proposed by the National Institute of Health and Care Excellence (NICE, 2004), which sets official treatment guidelines for the National Health Service in the UK. Leucht's data reveal these criteria to be equivalent to no difference at all for HRSD scores and well below a minimal difference for effect sizes (Moncrieff & Kirsch, 2015).

Thus, using empirically derived criteria for clinical significance, the difference in outcome between antidepressants and placebo is indeed trivial, as claimed by Cristea and Naudet. It does not come close to what would be considered a minimal difference in global ratings of improvement, let alone the 'much improved' criterion consensually adopted for evaluating the clinical significance of pain medications. On the other hand, a recent meta-analysis indicates that antidepressant medication has been shown to increase the risk of serious adverse events (Jakobsen *et al*. 2017). As concluded by the authors of that analysis, 'the potential small beneficial effects seem to be outweighed by harmful effects' (p. 2).

I. Kirsch
Program in Placebo Studies, Harvard Medical School, Boston, USA

References

Cohen J (1988). *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates: Hillsdale, NJ.

Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Witter J (2005). Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* **113**, 9–19.

Farrar JT, Young JP Jr., LaMoreaux L, Werth JL, Poole RM (2001). Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* **94**, 149–158.

Jakobsen JC, Katakam KK, Schou A, Hellmuth SG, Stallknecht SE, Leth-Møller K, Gluud C (2017). Selective serotonin reuptake inhibitors versus placebo in patients with major depressive disorder. A systematic review with meta-analysis and Trial Sequential Analysis. *BMC Psychiatry* **17**, 58.

Leucht S, Hierl S, Kissling W, Dold M, Davis JM (2012). Putting the efficacy of psychiatric and general medicine medication into perspective: review of meta-analyses. *British Journal of Psychiatry* **200**, 97–106.

Leucht S, Fennema H, Engel R, Kaspers-Janssen M, Lepping P, Szegedi A (2013). What does the HAMD mean? *Journal of Affective Disorders* **148**, 243–248.

Leucht S, Helfer B, Gartlehner G, Davis JM (2015). How effective are common medications: a perspective based on meta-analyses of major drugs. *BMC Medicine* **13**, 253.

Moncrieff J, Kirsch I (2015). Empirically derived criteria cast doubt on the clinical significance of antidepressant-placebo differences. *Contemporary Clinical Trials* **43**, 60–62.

NICE (2004). Depression: Management of depression in primary and secondary care. Clinical practice guideline No 23. Retrieved from http://www.nice.org.uk/page.aspx?o=235213.

https://doi.org/10.1017/S2045796017000725 Published online by Cambridge University Press