



## Anti-Inflammatory Treatments for Depression: Perspectives on How to Read a Meta-Analysis Critically

Chittaranjan Andrade, MD



Each month in his online column, Dr Andrade considers theoretical and practical ideas in clinical psychopharmacology with a view to update the knowledge and skills of medical practitioners who treat patients with psychiatric conditions.

Department of Psychopharmacology, National Institute of Mental Health and Neurosciences, Bangalore, India  
(candrade@psychiatrist.com).

### ABSTRACT

Inflammatory mechanisms have been implicated in many psychiatric disorders, including depression, and anti-inflammatory agents have been suggested as potential treatments. In this context, a systematic review and meta-analysis of nonsteroidal anti-inflammatory drugs, cytokine inhibitors, glucocorticoids, statins, minocycline, and pioglitazone, all of which are considered to have anti-inflammatory properties, examined the antidepressant benefits of these agents in 36 randomized controlled trials conducted in patients with major depressive disorder (MDD) and patients with medical diseases and associated depressive symptoms. The meta-analysis found that, overall, relative to placebo, these drugs had superior antidepressant effect in patients with MDD as well as in those with medical disease. With the exception of pioglitazone, every drug/category also outperformed placebo. However, the findings of this extensive meta-analysis do not guide theory; because the different anti-inflammatory agents studied have multiple pharmacodynamic actions, there is no assurance that their anti-inflammatory mechanism was responsible for the reported antidepressant benefits. The findings do not guide clinical practice, either, because of substantial clinical and statistical heterogeneity; no specific drug, dose, duration of treatment, and disorder were identified for application of the findings. Finally, the meta-analysis contained obvious and non-obvious errors, including the combination of endpoint and improvement scores in the same summary estimate (standardized mean deviation) and the use of percentage scores rather than absolute scores in the computations. These issues are explained so that readers can more easily consider or detect the limitations of meta-analyses that are published.

*J Clin Psychiatry* 2019;80(3):19f12907

**To cite:** Andrade C. Anti-inflammatory treatments for depression: perspectives on how to read a meta-analysis critically. *J Clin Psychiatry*. 2019;80(3):19f12907.

**To share:** <https://doi.org/10.4088/JCP.19f12907>  
© Copyright 2019 Physicians Postgraduate Press, Inc.

Inflammatory mechanisms may be involved in many psychiatric disorders, including dementia,<sup>1,2</sup> schizophrenia,<sup>3</sup> and depression.<sup>4-6</sup> The relationship between depression and inflammation is complex. Depression may predispose to inflammation, inflammation may predispose to depression, and each may predispose to the other in a bidirectional loop.<sup>7</sup> In this regard, proinflammatory cytokines have been suggested to modulate mood, behavior, and cognition by decreasing brain monoamine levels, altering neuroendocrine responses, promoting excitotoxicity through increased glutamate levels, and impairing neuroplasticity.<sup>7</sup>

If inflammatory mechanisms may cause or worsen depression, may anti-inflammatory drugs have antidepressant potential? This question was examined by Köhler-Forsberg et al<sup>8</sup> in a systematic review and meta-analysis of randomized controlled trials (RCTs) on the subject. The present article provides a critical examination of this meta-analysis so that readers can better understand certain important limitations of meta-analysis in general and this meta-analysis in particular.

### Anti-Inflammatory Treatments for Depression

Köhler-Forsberg et al<sup>8</sup> searched several electronic databases (including a clinical trial registry) and reference lists and identified 31 studies describing 36 RCTs of drugs with anti-inflammatory actions that had been examined for efficacy as monotherapy or as add-on therapy in patients with major depressive disorder (MDD) or somatic disease and depression. The authors included several different drugs and groups of drugs with anti-inflammatory action: nonsteroidal anti-inflammatory drugs (NSAIDs), cytokine inhibitors, glucocorticoids, statins, minocycline, and pioglitazone. Most of the studies were 6 to 12 weeks in duration; a few, conducted in patients with medical conditions such as multiple sclerosis, rheumatoid arthritis, and cardiovascular disease, lasted from 6 months to many years.

Several primary outcomes were listed in this meta-analysis: reduction in depression ratings at the treatment endpoint; response rate at the treatment endpoint; remission rate at the treatment endpoint; and the serious adverse event rate. Several secondary outcomes and subanalyses were also listed.

Important findings from the meta-analysis are presented in Table 1. In summary, anti-inflammatory drugs were superior to placebo for antidepressant outcomes overall, in patients with MDD, and in patients with somatic disease and depressive symptoms. With the exception of pioglitazone, all drugs and categories of anti-inflammatory drugs outperformed placebo. With the exception of the glucocorticoid analysis, analysis of all categories of drugs was associated with high to very

high heterogeneity. To the extent that the authors provided information, it appeared the benefits were limited to the use of anti-inflammatory drugs as antidepressant augmentation agents.

### Meta-Analysis: General Notes

If two studies are conducted with identical or near-identical study protocols but at different places or at different points in time, individual patient data can be pooled into a single analysis as though the two studies were a single study.<sup>9,10</sup> This can happen, for example, in industry-driven research where identical studies are conducted separately to meet regulatory requirements. More commonly, however, this happens in the context of multicenter studies in which the same study protocol is executed simultaneously across 2 or more study sites; even though there could be undocumented variations in the way diagnosis, treatment, and other study procedures are conducted across the sites, the data are pooled into a single analysis.

When studies are designed and conducted by different teams and in different centers, although the objectives of the research could be the same, variations in study design and execution would preclude pooling of individual patient data as in a pooled analysis. For example, if different teams of researchers study the antidepressant benefits of celecoxib, they may use different doses of the drug, different concurrent (conventional) antidepressants, and different doses of these antidepressants in RCTs of different durations, and they may rate the patients using different rating scales at different points in time. Such data, therefore, cannot be considered to have arisen from the same research protocol, and so the data cannot be pooled, even were individual patient data available to the research team that wanted to combine the findings from these studies.

This is where meta-analysis comes in. When studies are generally similar but not identical, meta-analysis offers a method for combining data and generating summary estimates of outcomes that could guide theory, practice, or both. In this regard, there are 3 ways in which the meta-analysis by Köhler-Forsberg<sup>8</sup> failed. These are discussed in the subsequent sections.

### Guiding Theory

Do the findings of the meta-analysis<sup>8</sup> guide theory in the field? For example, can one conclude, based on this meta-analysis, that anti-inflammatory treatments have antidepressant effects? No, and this is so for a rather curious but quite easily understood reason. Most of the categories of drugs studied in this meta-analysis have a wide range of pharmacodynamic effects. What is the assurance that, across the spectrum of drugs, it was the anti-inflammatory mechanism and only the anti-inflammatory mechanism that was responsible for the antidepressant benefit described in the summary estimates (Table 1)? In other words, it is quite possible that different categories of drugs may have exerted the putative antidepressant benefits in different ways, and not necessarily through anti-inflammatory mechanisms.

**Table 1. Important Findings<sup>a</sup> From the Meta-Analysis by Köhler-Forsberg et al<sup>8</sup>**

1. Relative to placebo, across all drugs and categories of drugs, anti-inflammatory agents significantly attenuated depression ratings (36 RCTs; N = 9,422; SMD, -0.49; 95% CI, -0.64 to -0.33). There was substantial heterogeneity in this analysis ( $I^2 = 89\%$ ).
2. Relative to placebo, across all drugs and categories of drugs, anti-inflammatory agents significantly attenuated depression ratings in patients with MDD (14 RCTs; N = 597; SMD, -0.64; 95% CI, -0.88 to -0.40). Heterogeneity was moderately high ( $I^2 = 51\%$ ).
3. Relative to placebo, across all drugs and categories of drugs, anti-inflammatory agents significantly attenuated the severity of depressive symptoms in patients with somatic disease (22 RCTs; N = 8,825; SMD, -0.41; 95% CI, -0.60 to -0.22). Heterogeneity was very high ( $I^2 = 93\%$ ).
4. In monotherapy trials, response and remission rates did not differ significantly between anti-inflammatory drugs and placebo. In add-on (antidepressant augmentation) trials, anti-inflammatory drugs were associated with higher response as well as remission rates.
5. Among individual drugs and categories of drugs, NSAIDs (13 RCTs), cytokine inhibitors (9 RCTs), glucocorticoids (2 RCTs), statins (7 RCTs), and minocycline (3 RCTs) but not pioglitazone (2 RCTs) were each significantly superior to placebo in attenuating depression ratings in patients with MDD or somatic disease with depression. With the exception of the glucocorticoid analysis, all these analyses were characterized by high to very high heterogeneity.
6. Across drugs and categories of drugs, anti-inflammatory agents were not associated with an increased risk of gastrointestinal symptoms (13 RCTs), pain/muscle aching (4 RCTs), infections (7 RCTs), and cardiovascular symptoms (3 RCTs).

<sup>a</sup>SMDs were calculated based on the difference in improvement scores between drug and placebo at treatment endpoint; improvement scores were calculated as the endpoint depression rating score minus the baseline depression rating score. Thus, more negative values indicate greater improvement and greater superiority of drug over placebo. Abbreviations: CI = confidence interval, MDD = major depressive disorder, NSAID = nonsteroidal anti-inflammatory drug, RCT = randomized controlled trial, SMD = standardized mean difference.

As a simple example, conventional antidepressant drugs have also been suggested to have anti-inflammatory effects.<sup>11-13</sup> So why weren't conventional antidepressants included in this meta-analysis; is it because we already have the monoamine and neuroplasticity hypotheses to explain their mechanisms of action? In like manner, might as yet unknown mechanisms, different from anti-inflammatory action, explain the antidepressant action, if any, of some or all of the categories of drugs found effective in this meta-analysis?

It must also be considered that antidepressant action may not have been a primary effect in many of the RCTs. For example, in the RCTs conducted in patients with medical disease and depression, if the depression was secondary to the inflammation and pain associated with the medical disease, then the anti-inflammatory treatment may have reduced the inflammation and pain, and only secondarily resulted in attenuation of the depression. This is not the same as a primary antidepressant action.

Finally, it may be noted that merely outperforming placebo in the attenuation of antidepressant ratings is not a sufficient criterion to assert antidepressant action; even a hypnotic drug might do this by improving sleep outcomes on a depression rating scale that is heavily weighted for insomnia-related items. An anti-inflammatory treatment

**It is illegal to post this copyrighted PDF on any website.**

may reduce pain and discomfort and thereby improve sleep and reduce somatic preoccupations, so these changes would likewise be associated with lower depression scores. What is necessary, therefore, is to demonstrate that the core symptoms of depression are attenuated (which was not done in the meta-analysis) or to demonstrate that response and remission rates are higher in drug as compared with placebo groups (which could only be done for a small number of RCTs).

### Guiding Practice

Do the findings of the meta-analysis<sup>8</sup> help guide the treatment of depression? The answer to this is also in the negative. The reason is again easy to understand. For the results of a meta-analysis to be relevant, the findings must be generalizable to clinical practice. In ordinary clinical practice, patients with MDD will be treated with conventional antidepressant drugs; only treatment-refractory patients are likely to be considered for augmentation such as with anti-inflammatory agents. However, the meta-analysis provided no information about the usefulness of anti-inflammatory drugs in refractory MDD; in fact, it yielded no usable information, either, about the choice of individual drug and dose for MDD. Likewise, there was no good information available on which medical condition would be best treated with which drug and in what dose or whether this should be done in monotherapy or as an antidepressant augmentation strategy. The meta-analysis therefore fails to guide clinical practice because the RCTs in the meta-analyses were widely different in terms of diagnoses, drugs, and other study-related characteristics. This is addressed further in the next section, which considers heterogeneity.

A note is made here that there were a number of subgroup analyses performed on, for example, celecoxib alone and statins alone. For these analyses to be clinically useful, they should each have been the subject of an entirely independent systematic review and meta-analysis because, in such a meta-analysis, primary and secondary outcome measures could have been specified, subgroup and sensitivity analyses planned, design issues related to individual RCTs examined, and conclusions relevant to clinical practice drawn. All these were not possible in the umbrella meta-analysis under consideration.<sup>8</sup>

### Heterogeneity

As already stated, meta-analysis is primarily intended to combine studies that are reasonably similar so that readers can easily extrapolate the findings of the meta-analysis to their own practice, based on the methods of the individual studies (eg, with regard to drug, dose, and duration of treatment) and the characteristics of the study samples (eg, age, sex distribution, clinical diagnosis, and presence of comorbidities).

Combining RCTs that vary considerably in design and other characteristics is like combining apples and oranges; what is the average of these fruit? In this context, NSAIDs, cytokine inhibitors, statins, glucocorticoids, and so on are the

different varieties of fruit that are being pooled. So, which of these treatments must one choose when applying the results of the meta-analysis to clinical practice? Considering the wide range of clinical diagnoses in the meta-analyzed RCTs, to what clinical sample may one generalize the summary statistics?

One way of considering these tricky questions is to assess statistical heterogeneity in each meta-analysis. If there is little to none, then it supports a view that variations in study and sample characteristics have not influenced the meta-analysis outcomes. If heterogeneity is substantial, then subgroup and sensitivity analyses and meta-regression analyses (in which study descriptors are regressed against measures of effect size) may identify the source of the heterogeneity, and hence the study characteristics that influence the meta-analysis results.

The authors of this meta-analysis<sup>8</sup> performed many subgroup analyses. However, with a few exceptions, most of the subgroup analyses included too few trials and too few patients for the analyses to suggest actionable outcomes. Furthermore, there was substantial statistical heterogeneity not just in the main analyses but in the subgroup analyses, as well. The authors did not perform sensitivity analyses and meta-regression analysis, and the source of the heterogeneity remained undetected. Studies that contributed disproportionately to the summary statistics remained unidentified.

### An Elephant in the Room

It is assumed that no mistakes occur when authors extract data and subject these to meta-analysis. However, an examination of the numbers in the Köhler-Forsberg et al meta-analysis<sup>8</sup> suggests that there is a rather obvious elephant in the room. The reader is invited to inspect the forest plots presented in the main analyses in Figures 2 and 3 in the published paper. Scrutiny of the mean depression scores in the Treatment and Placebo columns reveals 2 curiosities. One is that some of the values are rather small and close to zero in value, whereas others are very large, such as in the region of 40 to -75. The other curiosity is that some of the values are positive, and as high as over 20, whereas others are negative, and as high as over -75.

These curiosities are simply explained. The authors extracted depression scores at treatment endpoint from some of the RCTs, such as those by Haghghi et al,<sup>14</sup> Husain et al,<sup>15</sup> and Dean et al,<sup>16</sup> and improvement scores at treatment endpoint from other RCTs, such as those by Akhondzadeh et al<sup>17</sup> and Abbasi et al.<sup>18</sup> These can be seen in the very first forest plot, labeled as 1.1.1 in Figure 2 in the published paper.<sup>8</sup>

Endpoint depression scores cannot be a negative value, so these would have appeared as the positive values in the forest plots. Improvement scores are conventionally calculated as endpoint minus baseline scores. So these would have been the positive values that were close to zero, if patients worsened at endpoint, and the negative values if patients improved.

The authors combined endpoint and improvement score differences between Treatment and Placebo groups in their meta-analyses. This is not permitted because how depressed a patient is at the end of the study is conceptually completely different from how much the patient has improved at the end of the study. These variables will also have different variances and will have different impacts on the meta-analyses. However, the software that crunches the numbers will not know this and will treat all numbers as legitimate entries.

Here is a simple explanation for why treatment endpoint and change scores cannot be combined in the same analysis. Imagine a study in which depression scores improve from 25 at baseline to 20 at endpoint in the treatment group and another study in which the depression scores improve from 25 at baseline to 23 at endpoint in the same (treatment) group. It should be clear from these two studies that the treatment is not very effective. In the first study, the endpoint score is 20. In the second study, the change score is  $-2$ .

It is not logical to combine these scores as the average of 20 and  $-2$ . And if one *does* try to average 20 and  $-2$ , the unweighted result is 9, which seems to suggest that the drug was effective, after all. The error approximately cancels out when intervention and control groups are compared in terms of absolute change; that is, when the summary statistic computed is the mean difference. However, the combination of endpoint and change scores is problematic and not permitted when the summary statistic is the standardized mean difference (Higgins and Green<sup>19</sup>; section 9.4.5.2), as in the meta-analysis under examination.<sup>8</sup> Other published meta-analyses have also committed this mistake, as pointed out in an earlier article in this column.<sup>20</sup>

## A Second Elephant

The elephant that is the mixing of endpoint and change scores in the same analysis is easily spotted once the reader knows what to look for, as explained in the previous section. The meta-analysis<sup>8</sup> also contained another elephant, one that the reader would never identify without an inspection of the source data. As an example, for at least 1 RCT,<sup>21</sup> the meta-analysis extracted percentage scores (with standard error of the percentage scores) instead of absolute scores. This study was included in the forest plot labeled 1.1.2 in Figure 2 in the published paper.<sup>8</sup>

Why is the use of percentage data in place of actual numbers problematic? In one group, a patient's scores may

improve from 20 to 10; that's a 10-point improvement but a 50% change. In the other group, a patient may improve from 30 to 20; that's also a 10-point improvement, but only a 33% change. Had change (improvement) scores been analyzed, both patients would have been considered to have improved by a similar extent. When percentage change scores are analyzed, because different patients have different baselines, improvement will be a function not only of the actual improvement but also of what the baseline score was.

A further elephant is that, in this analysis, the authors appear to have used the standard error and not the standard deviation to calculate the standardized mean deviation in the forest plot.

## General Comments

Here is an additional point that readers need to keep in mind when reading meta-analyses such as this.<sup>8</sup> When a treatment is used as an add-on, it may be used in either non-refractory patients, to determine whether response can be increased or hastened, or in refractory patients, to determine whether nonresponders can be converted into responders. There is a risk of a ceiling effect in the former situation; that is, the primary drug may produce such a robust effect that the add-on drug may not have room to demonstrate its own benefits. There is a risk of a floor effect in the latter situation; that is, the sample may be so difficult to treat that even an add-on may not be of much help. In both situations, the add-on treatment may be disadvantaged. Therefore, combining monotherapy studies with add-on studies is again like combining apples and oranges; the average may represent neither.

This is not to say that such studies should not be combined. This is merely to alert the reader to the need to interpret the summary statistics appropriately and to look for the effect of this aspect of study design on the summary statistics in the subgroup analyses and meta-regression analyses, if conducted.

## Further Reading

Readers who are looking for a primer on meta-analysis could start with the excellent article by Streiner.<sup>22</sup> Readers are also referred to useful articles on how to read meta-analyses critically.<sup>23–25</sup> Finally, the Cochrane Handbook<sup>19</sup> is a detailed text for those who wish to learn about meta-analysis in depth.

**Published online:** May 21, 2019.

## REFERENCES

- McGeer PL, Rogers J, McGeer EG. Inflammation, antiinflammatory agents, and Alzheimer's disease: the last 22 years. *J Alzheimers Dis*. 2016;54(3):853–857.
- Ardura-Fabregat A, Boddeke EWGM, Boza-Serrano A, et al. Targeting neuroinflammation to treat Alzheimer's disease. *CNS Drugs*. 2017;31(12):1057–1082.
- Müller N. Inflammation in schizophrenia: pathogenetic aspects and therapeutic considerations. *Schizophr Bull*. 2018;44(5):973–982.
- Leonard BE. Inflammation and depression: a causal or coincidental link to the pathophysiology? *Acta Neuropsychiatr*. 2018;30(1):1–16.
- Dooley LN, Kuhlman KR, Robles TF, et al. The role of inflammation in core features of depression: insights from paradigms using exogenously-induced inflammation. *Neurosci Biobehav Rev*. 2018;94:219–237.
- Woelfer M, Kasties V, Kahlfuss S, et al. The role of depressive subtypes within the neuroinflammation hypothesis of major depressive disorder. *Neuroscience*. 2019;403:93–110.
- Bauer ME, Teixeira AL. Inflammation in psychiatric disorders: what comes first? *Ann N Y Acad Sci*. 2019;1437(1):57–67.
- Köhler-Forsberg O, N Lydholm C, Hjorthøj C, et al. Efficacy of anti-inflammatory treatment on major depressive disorder or depressive symptoms: meta-analysis of clinical trials. *Acta Psychiatr Scand*. 2019;139(5):404–419.
- Allgulander C, Hartford J, Russell J, et al. Pharmacotherapy of generalized anxiety disorder: results of duloxetine treatment from a pooled analysis of three clinical trials. *Curr Med Res Opin*. 2007;23(6):1245–1252.
- Tourian KA, Padmanabhan SK, Groark J, et al. Desvenlafaxine 50 and 100 mg/d in the

It is illegal to post this copyrighted PDF on any website.

- treatment of major depressive disorder: an 8-week, phase III, multicenter, randomized, double-blind, placebo-controlled, parallel-group trial and a post hoc pooled analysis of three studies. *Clin Ther.* 2009;31(pt 1):1405–1423.
11. Nazimek K, Strobel S, Bryniarski P, et al. The role of macrophages in anti-inflammatory activity of antidepressant drugs. *Immunobiology.* 2017;222(6):823–830.
  12. Gałecki P, Mossakowska-Wójcik J, Talarowska M. The anti-inflammatory mechanism of antidepressants: SSRIs, SNRIs. *Prog Neuropsychopharmacol Biol Psychiatry.* 2018;80(pt C):291–294.
  13. Dhami KS, Churchward MA, Baker GB, et al. Fluoxetine and its metabolite norfluoxetine induce microglial apoptosis. *J Neurochem.* 2019;148(6):761–778.
  14. Haghghi M, Khodakarami S, Jahangard L, et al. In a randomized, double-blind clinical trial, adjuvant atorvastatin improved symptoms of depression and blood lipid values in patients suffering from severe major depressive disorder. *J Psychiatr Res.* 2014;58:109–114.
  15. Husain MI, Chaudhry IB, Husain N, et al. Minocycline as an adjunct for treatment-resistant depressive symptoms: a pilot randomised placebo-controlled trial. *J Psychopharmacol.* 2017;31(9):1166–1175.
  16. Dean OM, Kanchanatawan B, Ashton M, et al. Adjunctive minocycline treatment for major depressive disorder: a proof of concept trial. *Aust N Z J Psychiatry.* 2017;51(8):829–840.
  17. Akhondzadeh S, Jafari S, Raisi F, et al. Clinical trial of adjunctive celecoxib treatment in patients with major depression: a double blind and placebo controlled trial. *Depress Anxiety.* 2009;26(7):607–611.
  18. Abbasi SH, Hosseini F, Modabbernia A, et al. Effect of celecoxib add-on treatment on symptoms and serum IL-6 concentrations in patients with major depressive disorder: randomized double-blind placebo-controlled study. *J Affect Disord.* 2012;141(2–3):308–314.
  19. Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions.* Chichester, West Sussex, England: Wiley-Blackwell; 2008.
  20. Andrade C. The use of statins for antipsychotic augmentation in schizophrenia: examination of meta-analyses with flawed methods and conclusions. *J Clin Psychiatry.* 2018;79(5):18f12562.
  21. Tyring S, Bagel J, Lynde C, et al. Patient-reported outcomes in moderate-to-severe plaque psoriasis with scalp involvement: results from a randomized, double-blind, placebo-controlled study of etanercept. *J Eur Acad Dermatol Venereol.* 2013;27(1):125–128.
  22. Streiner DL. Using meta-analysis in psychiatric research. *Can J Psychiatry.* 1991;36(5):357–362.
  23. Streiner DL. I have the answer, now what's the question? why metaanalyses do not provide definitive solutions. *Can J Psychiatry.* 2005;50(13):829–831.
  24. Lam RW, Kennedy SH. Using metaanalysis to evaluate evidence: practical tips and traps. *Can J Psychiatry.* 2005;50(3):167–174.
  25. Huf W, Kalcher K, Pail G, et al. Meta-analysis: fact or fiction? how to interpret meta-analyses. *World J Biol Psychiatry.* 2011;12(3):188–200.

You are prohibited from making this PDF publicly available.