Understanding the Antidepressant Debate in the Treatment of Major Depressive Disorder

Florian Naudet, Rémy Boussageon, Clément Palpacuer, Laurent Gallet, Jean-Michel Reymann, Bruno Falissard

THERAPEUTIQUE

Thérapie 2015
DOI: 10.2515/therapie/2014228

Understanding the Antidepressant Debate in the Treatment of Major Depressive Disorder

Florian Naudet[1,2,3], *Rémy Boussageon*[4], *Clément Palpacuer*[2], *Laurent Gallet*[3], *Jean-Michel Reymann*[2,5] *and Bruno Falissard*[1,6,7,8]

1 INSERM U669, Paris, France
2 Centre d'investigation clinique CIC-P INSERM 1414, Hôpital de Pontchaillou, Centre hospitalier universitaire de Rennes & Université de Rennes 1, Rennes, France
3 Centre hospitalier Guillaume Régnier, Service hospitalo-universitaire de psychiatrie, Rennes, France
4 Faculté de médecine de Poitiers, Département de médecine générale, Poitiers, France
5 Laboratoire de pharmacologie expérimentale et clinique, Faculté de médecine de Rennes, Rennes, France
6 Université Paris-Sud et Université Paris Descartes, UMR-S0669, Paris, France
7 AP-HP, Hôpital Paul Brousse, Département de santé publique, Villejuif, France

Text received November 11th, 2014; accepted November 20th, 2014

Keywords:
antidepressants;
randomised controlled trials;
placebo;
publication bias;
outcomes

Abstract – There is a long-standing polemic concerning the usefulness of antidepressants in the treatment of major depressive disorder. In this paper we propose to highlight some aspects of this controversy by exploring the mutual influence of psychopharmacology and trial methodologies. Indeed, antidepressant efficacy, if not proved, was accepted before antidepressant randomised controlled trials (RCTs) were run. While RCTs became a gold standard to meet the requirements of the regulatory bodies, methodological tools were required to measure outcomes and to test whether antidepressants provide statistically significant benefits as compared with a placebo. All these methodological options have nonetheless introduced fuzziness in our interpretation of study results, in terms of clinical meaningfulness and in terms of transposability to a real life settings. Additionally, selective publication raises concerns about the published literature, and results in many paradoxes. Instead of providing easy answers, the application of the RCT paradigm in MDD raises numerous questions. This is probably in the nature of all scientific studies, but it can be in contradiction with clinicians' expectations, who want to be sure that the treatment will (or will not) work for their individual patients.

Mots clés :
antidépresseurs ;
essais randomisés contrôlés ;
placebo ;
biais de publication ;
critères de jugements

Résumé – Comprendre le débat sur les antidépresseurs dans le traitement de l'épisode dépressif majeur. Il existe un vieux débat à propos de l'utilité des antidépresseurs dans le traitement de l'épisode dépressif majeur. Dans cet article, nous présentons certains aspects de la controverse en explorant l'influence mutuelle de la psychopharmacologie et de la méthodologie des essais. En effet, l'efficacité des antidépresseurs était, sinon prouvée, admise avant que les premières études contrôlées randomisées (ECR) ne soient conduites. Alors que les ECR devenaient, du point de vue des autorités sanitaires, le "gold standard" pour l'évaluation des médicaments, il devenait nécessaire d'adopter des outils méthodologiques permettant de mesurer des critères de jugement et de tester si les antidépresseurs permettaient l'obtention d'une différence statistiquement significative par rapport au placebo. Ces options méthodologiques ont néanmoins introduit du flou quand à l'interprétation des résultats des ECR, notamment en terme de significativité clinique et de transposabilité « à la vraie vie ». Au-delà, la publication sélective des ECR impacte la validité de la littérature publiée et résulte en de nombreux paradoxes. Ainsi, au lieu de fournir des réponses simples, l'application du paradigme de l'ECR à l'épisode dépressif majeur soulève de nombreuses questions. Il en va probablement de même pour toutes les études scientifiques, mais dans ce cas précis, cela rentre en contradiction avec les attentes des cliniciens qui veulent être sûr que leur traitement sera efficace (ou pas) pour leurs patients.

Abbreviations : see end of article.

1. Introduction

There is a long-standing but still active polemic concerning the usefulness of antidepressants in the treatment of major depressive disorder. Recently, some opinion leaders stated that antidepressants have no place in evidence-based medicine,[1] while others consider that this is an "irrational polemic" and have disputed psychological interventions for depression.[2] This debate could lead to a major public health problem, since treatments that are offered to patients (pharmacological or psychological) are being discredited by partisans of either side, and this risks depriving some patients with depression of useful treatments. The subject is too important to reduce to a mere opposition between "pro" and "anti" antidepressants;[3] and it deserves careful examination from different points of view. In this paper we propose to highlight some aspects of this controversy.

2. Birth of the concepts of antidepressant and major depressive disorder

In the case of depression, stimulants were used as the treatment during the 1940s. In the 1950s, new substances such as iproniazid and imipramine were viewed as specific to treating depression, whereas earlier stimulants were regarded as non-specific.[4] In 1958, Khun[5] presented imipramine as an antidepressant although its biological foundations were not established. He noted that "best responses were obtained in cases of endogenous depression showing the typical symptoms of mental and motor retardation, fatigue, feeling of heaviness, hopelessness, guilt, and despair" and that this "condition is furthermore characterized by the aggravation of symptoms in the morning with a tendency to improvement during the day". Promptly, the monoamine theory of depression emerged[6] with the work by Sigg[7] who demonstrated that imipramine can potentiate the effects of noradrenaline, by Burn and Rand[8] who described the uptake of noradrenaline by adrenergic nerves, by Marshall *et al.*[9] who reported that imipramine blocked the uptake of serotonin by platelets, by Axelrod *et al.*[10] who described the uptake of labelled noradrenaline by adrenergic nerves which could be blocked by imipramine, and by Dengler *et al.*[11] who reported similar data regarding noradrenaline uptake by brain tissue. Arvid Carlsson developed zimelidine, a new treatment blocking the uptake of serotonin without blocking the uptake of cathecholamines.[12] While zilmelidine had a very favourable safety profile, within a year and a half of its introduction, some case reports of Guillain-Barré syndrome emerged, apparently caused by the drug, prompting its manufacturer to withdraw it from the market. After its withdrawal, it was succeeded by fluoxetine and the other serotonin reuptake inhibitors (SRIs) which were considered as selective drugs with fewer adverse events.[13]

The idea of an antidepressant, and the discoveries about their putative biological properties, reshaped the concept of depression. A debate emerged concerning whether there was any value in distinguishing "endogenous depression" and milder conditions in relation with stressful events known as neurotic depression (the Khun perspective) and treating them differently, or whether there was no basis for separate categories of depression since they all lie on a continuum of severity, as proposed by Akiskal and Mc Kinney.[14] In 1980, the Diagnostic and Statistical Manual of Mental Disorders (DSM) III[15] retained the latter view by combining the two entities under the label of major depressive disorder (MDD).

Non-scientific reasons have probably also contributed to the wide acceptation of the concepts of antidepressant and MDD.[4] Concerning the ideological conflict of interest, these concepts were not in favour of the psychiatric profession's desire to integrate with general medicine and to counter attacks from the antipsychiatry movement. Concerning the financial conflict of interest, the pharmaceutical industry also had an interest in promoting these concepts.[4]

3. RCTs became inescapable in the evaluation of antidepressants

Alongside these conceptual changes randomized controlled trials (RCTs) developed in the evaluation of medication. The Medical Research Council (MRC) ran the first RCT *versus* placebo in 1948 to explore the efficacy of streptomycin in tuberculosis.[16] Previous non-randomized studies had established that streptomycin worked in the short term treatment of tuberculosis, but an *a posteriori* interpretation of this trial is that it probably proved the "efficacy of RCTs" rather than the efficacy of streptomycin.[17] In the years following this trial, many RCTs were funded by national public bodies, for example the MRC evaluation of imipramine versus phenelzine, electroconvulsive therapy and placebo in the relief of depressive illness.[18] These trials were often concerned with broad questions regarding classes of treatments, rather than specific compounds.[19]

After the thalidomide crisis in 1962, the Kefauver-Harris drug amendments were passed to ensure drug efficacy and greater drug safety. It was because medications entailed a risk that evidence of efficacy was sought and, for the first time, drug manufacturers were required to prove to the Food and Drug Administration (FDA) the efficacy of their products before marketing them. Gradually, the situation changed, public funding declined and the vast majority of clinical trials on drug treatments in psychiatry

were sponsored and conducted by the pharmaceutical industry, the number of trials increased dramatically, trials concerned single patented compounds and were designed to meet the requirements of the regulatory bodies.[19] While for a large proportion of medical interventions, few or no clinical trials are ever conducted, for antidepressants there are probably now well over a thousand.[20]

4. The mutual influence of psychopharmacology and trial methodology

Nonetheless, it should be noted that antidepressant efficacy, if not proved, was accepted before antidepressant RCTs were run, and that no antidepressant in the RCT era was proved to be superior to imipramine in terms of efficacy.[21] Thus, being thought-provocative, one can say that antidepressants have made advances in methodology possible, rather than stating that methodology has enabled major advances in psychopharmacology for MDD. Indeed, when RCTs became a gold standard, it became necessary for them to take into account the particular features of psychopharmacology, and especially those relating to MDD, for instance paying particular attention to inclusion criteria and outcomes. Concerning inclusion criteria, as it became necessary to accept a common definition of MDD, the DSM viewpoint was reinforced as a standard. It also became necessary to adopt measurable, relevant and consensual outcomes providing a sensitive and accurate estimate of change occurring with antidepressants.[22] The Hamilton Depression Rating Scale (HDRS), developed in 1960,[23] was progressively imposed as a standard, and was subsequently challenged by the Montgomery and Åsberg Depression Rating Scale (MADRS),[22] a scale developed to be particularly sensitive to treatment effects. It is nonetheless interesting that a scale that is to be used to assess the difference between a treatment and a placebo was developed to be particularly sensitive to specific changes occurring under treatment. The Clinical Global Impression[24] (CGI) which rates severity on a scale of 1 to 7, was retained as a reference for global assessment and some self-administered questionnaires like the Beck Depression Inventory (BDI) among others were popularised by the wide development of RCTs in MDD.[25] Binary outcomes also had to be adopted, such as response and remission, which have meaning for clinicians. Despite the fact that they are intuitive, their definition is not straightforward and a consensus emerged to derive these outcomes from continuous rating scales by calculating the proportion of people who fall below predefined threshold scores, which tend to be validated merely by convention and tradition.[26] Since 1991[27] remission is defined as a score $\leqslant 7$ on the 17 items of the Hamilton Depression Rating Scale (HDRS-17) and response is usually defined as a reduction of 50% on the HDRS-17.

5. Statistically significant versus clinically meaningful results

While these methodological tools enable the measurement of outcomes and test whether antidepressants provide statistically significant benefits as compared with a placebo, there is a considerable debate concerning the real meaning of the difference in term of its clinical significance. Indeed, the identification of a minimal clinically relevant difference on a scale is not straightforward. In 2004, the National Institute of Clinical Excellence[28] stated that a Hamilton score difference of three points across groups could be considered as clinically significant. This threshold was consistent with previous research[29] but a recent linking analysis provided new insight by suggesting that a slight reduction on the HAMD-17 of up to 3 points corresponds to a rating of "no change" as measured with the CGI. A change close to 10 points was linked to the "much improved" category defined by the CGI.[30] But these considerations on an individual level are not totally transposable to group level. On the other hand, this study also suggested that the commonly used measures for response (1) and remission (2) in MDD trials could reasonably be considered valid because they were coherent with the CGI definitions "much improved" (1) and "not at all" or "borderline mentally ill" (2), respectively. Bearing in mind that the CGI is not a perfect gold standard, these results are very interesting.

6. RCTs and the dilution of efficacy

To cope with the questions of variability and randomness, randomised controlled trials (RCT) "tell stories" about average patients, and the statistical inferences underpinning RCT conclusions concern expected values of random variables.[31] This type of paradigm implies that sufficient efficacy in a subgroup of patients can induce an impression of efficacy for the whole group, providing the study is adequately powered. This "dilution" of efficacy can occur especially in the case of heterogeneous categories such as MDD. Recent meta-analyses have indeed shed new light on this debate. Meta-analyses on aggregated data by Khan *et al.*[32] and Kirsh *et al.*[33] suggested that the baseline severity of depressive symptoms is related to clinical trial outcomes. These two meta-analyses were based on FDA data (i.e. an exhaustive set of studies) but were prone to an ecological fallacy[20] since they were based on aggregated data. Nonetheless, their results were reproduced by Fournier *et al.* within the framework of an individual data meta-analysis.[34] This study addressed the limitations

of aggregated data meta-analyses, but since personal data are difficult to collect, it was prone to publication bias. Nevertheless, these three meta-analyses concluded consistently that the distinction between antidepressants and placebo is clinically meaningful (using the National Institute for Clinical Excellence threshold for clinical significance) only for severe and very severe patients.

Interestingly, Gibbons *et al.*[35] addressed the limitations of the preceding studies by reanalysing all intent-to-treat individual longitudinal data during the first 6 weeks of treatment for major depressive disorder from all sponsored randomized controlled trials on fluoxetine and venlafaxine. In this meta-analysis, average differences at 6 weeks were small and not clinically meaningful (2.5 HAM-D units) and baseline severity was not shown to affect symptom reduction. But these small overall mean differences translated into clinically significant differences in response rates (estimated response rates were 58.4% for drug *versus* 39.9% for placebo) and remission rates (59.1% for drug *versus* 41.9% for placebo, relative risk = 1.5, number needed to treat = 5). This finding seems surprising. Intuitively, the two methods of assessing outcome should produce similar conclusions, since they are derived from the same data. However, this result can be explained by an artefact inherent in the transformation of continuous data into categorical data, which can magnify small differences.[36] But on the other hand, transformation of continuous outcomes into categorical outcomes implies a misclassification bias, and measures of association such as relative risk are likely to be biased towards 1.[37] An alternative explanation is that "efficacy dilution" is at play here.

7. Antidepressant alibis

In all events, beyond any fuzziness concerning the interpretation of antidepressant efficacy in MDD, a large number of RCTs turn out negative. It is frequently suggested that this is due to a marked placebo response in antidepressant trials, which could result from many different factors, such as spontaneous improvement,[38] statistical regression to the mean, low level of severity at inclusion, co-interventions, and other biases in addition to the so-called placebo effect. For example, spontaneous improvement is common in clinical practice,[38,39] and the number of follow-up assessments[40] is related to a significant therapeutic effect.[41] From a naïve point of view, one might have expected that in MDD, since it is a "mental disorder", the placebo effect (with its psychological component) might be greater than in other conditions and, as a consequence, the resulting true "pharmacological" effect would be weaker than in general medicine. However the distinction is probably more subtle.[42] In a meta-analysis,[43] Hrobjartsson *et al*. identified no statistically significant effect of placebo interventions in depression, while a meta-analysis by Kirsh *et al*. suggested that placebo effects were considerable.[44] But the RCTs included in Hrobjartsson's meta-analysis were not designed (they were underpowered) to study the placebo effect adequately. Similarly, in Kirsch's meta-analysis, which comprises no "untreated group" or waiting list, we cannot determine the size of the placebo effect. There is thus considerable debate about the size, the nature and the mechanism of the placebo effect in depression.[42] For example, it has been proposed that the apparent antidepressant effect could be in part an active placebo effect, or result from bias, since side effects like sexual effects[45] of antidepressants could reveal the identity of the medication to participants or investigators.[46]

Nonetheless, while some general medical drugs have very high effect sizes, the effect sizes obtained by psychiatric drugs are in the same range as most general medical pharmaceuticals.[47] Although it is difficult to compare effect sizes of drugs in different conditions, indications and outcomes, this finding puts the small effect sizes observed with antidepressants into perspective.

8. Overestimation and distortion of efficacy

Antidepressants efficacy is nonetheless certainly overestimated in the published literature by selective publication and selective outcome reporting. To explore this phenomenon, Turner *et al*. performed an analysis of 74 studies that were submitted to FDA for the approval of 12 antidepressant drugs. Among these studies, the FDA considered that 38 (51%) were "positive" (with a statistically significant result on the principal outcome), 12 (16%) "indeterminate" and 24 (33%) "negative" (with no statistically significant result on the principal outcome). Among the "positive" studies, 37 (97%) were published and only one (3%) was not published. Among the "indeterminate" studies, 6 (50%) were published as positive and 6 (50%) were unpublished. Finally, of the "negative" studies, 3 (12%) were published as "negative", in agreement with the opinion of the FDA, 5 (21%) were published as "positive", in disagreement with the opinion of FDA and 16 (67%) were not published. The effect size measured by performing a meta-analysis on the basis of published results is 0.41 with a 95% confidence interval of [0.36-0.45], whereas it is estimated to be 0.31 with a 95% confidence interval of [0.27-0.35] based on all studies reported to FDA.

The best-documented case of selective outcome reporting is probably study 329.[48–50] It was a large study of 275 depressed adolescents conducted by SmithKline Beecham in the US from 1993-1996. Its results failed to show any statistically significant difference between paroxetine and placebo for the two primary outcomes. A GSK internal document stated that the results of study 329 indicated paroxetine was no more effective than

placebo, and provided guidance on how to manage these disappointing results by recommending they should "effectively manage the dissemination of these data in order to minimize any potential negative commercial impact." It also stated that "it would be commercially unacceptable to include a statement that efficacy had not been demonstrated, as this would undermine the profile of paroxetine."[51] Subsequently, an article was written (or more precisely ghostwritten) with positive results concerning new secondary outcome measures that had been introduced. It was concluded that paroxetine is "generally well tolerated and effective for major depression in adolescents."[52]

9. Paradoxes in comparative effectiveness assessments

As a result of selective outcome reporting of this type, meta-analyses are likely to give misleading impressions about efficacy and comparative effectiveness of antidepressants.[53,54] There is the case of reboxetine, a selective norepinephrine reuptake inhibitor used in the treatment of depression. The previously favorable risk-benefit profile of reboxetine shown in published trials[55] was reversed by the addition of unpublished data.[56] In a network meta-analysis performed by Cipriani *et al*., reboxetine was consistently shown to be worse than 11 other antidepressants,[57] including paroxetine which was however found in another meta-analysis by the same team not to have any superiority over placebo.[58] All in all, these meta-analyses appear paradoxical, giving the impression that paroxetine is not superior to placebo, while it does better than reboxetine, which has itself been shown not to be superior to placebo. Additionally, although the Cipriani study found differences between antidepressants, this was not the case for another network meta-analysis performed by Gartlehner *et al*.[59]

Another paradox has been shown in a recent paper comparing citalopram with its "me-too", escitalopram, which found an inconsistency between direct evidence (showing a superiority of escitalopram) and indirect evidence (which did not find any significant difference).[60]

10. Poor transposability of RCT results

Beyond these issues RCTs are often criticised for their lack of external validity. Indeed, the vast majority of patients with clinical depression are catered for in primary care, and most RCTs have involved secondary care patients.[61] These patients probably differ from primary care patients.[62,63] in terms of severity (primary care patients are less severely depressed, milder course of illness) and in terms of complaints (fatigue and somatic symptoms).[64] Additionally, antidepressant RCTs use numerous non-inclusion criteria (for example suicidal ideations)[65–67] and excluded patients are a more chronically ill group with more numerous previous episodes, greater psychosocial impairment, and more frequent personality disorders. Finally, the vast majority of RCTs last no more than 8 weeks, whereas it is recommended that an antidepressant treatment be continued for at least 6 months after remission of the episode.[68]

There is debate as to whether these issues can be translated into different outcomes between RCTs and a "real life" setting.[69–72]

11. Conclusion

While meta-analyses should be reproducible, in 2013, a meta-analysis of published and unpublished studies on agomelatine found "evidence suggesting that a clinically important difference between agomelatine and placebo in patients with unipolar major depression was unlikely";[73] in 2014 a meta-analysis of published and unpublished studies on agomelatine found that it "was an effective antidepressant with similar efficacy to standard antidepressants".[74] This particular paradox sums up the fuzziness of antidepressant literature. We suggest that, instead of providing easy answers, the application of the RCT paradigm to MDD raises many questions. This is probably in the nature of all scientific studies, but it can be in contradiction with clinicians' expectations: what they want is to be sure that the treatment will work for individual patients (or to know if it will not). At the same time, their clinicial experience is biased by many other parameters, including placebo response. This is precisely where the debate arises.

Acknowledgements. We thank Angela Swaine Verdier for revising the English.

Details of contributors. Wrote the first draft of the paper: N.F.
Revised the paper critically for important intellectual content: B.R, P.C., G.L., R. J.M., F. B.
Final approval of the version to be published: N.F. B.R, P.C., G.L., R. J.M., F.B.

Conflict of interests. There are no conflicts of interest regarding this paper. All authors have completed the Unified Competing Interest form at `http://www.icmje.org/coi_disclosure.pdf` (available on request from the corresponding author) and declare that (1) No authors have support from any company for the submitted work; (2) N.F. has relationships (board membership or travel/accommodations expenses covered/reimbursed) with Servier, BMS, Lundbeck and Janssen who might have an interest in the work submitted in the previous 3 years ; B.R., P.C. G.L. and R.J.M. have no relationships with any company that might

have an interest in the submitted work in the previous 3 years; F.B has relationships (board membership or consultancy or payment for manuscript preparation or Travel/accommodations expenses covered/reimbursed) with Sanofi-Aventis, Servier, Pierre-Fabre, MSD, Lilly, Janssen-Cilag, Otsuka, Lundbeck, Genzime, Roche, BMS who might have an interest in the work submitted in the previous 3 years (3) none of the authors' spouses, partners, or children have any financial relationships that may be relevant to the submitted work; and (4) none of the authors has any non-financial interests that may be relevant to the submitted work.

Abbreviations. BDI: Beck Depression Inventory; CGI: Clinical Global Impression; DSM: Diagnostic And Statistical Manual Of Mental Disorders; FDA: Food and Drug Administration; HDRS: Hamilton depression rating scale; MDD: major depressive disorder; MRC: Medical Research Council; RCTs: randomised controlled trials; SRIs: serotonin reuptake inhibitors.

References

1. Gøtzsche PC. Why I think antidepressants cause more harm than good. Lancet Psychiatry 2014; 1(2): 104-6 doi:10.1016/S2215-0366(14)70280-9
2. Nutt DJ, Goodwin GM, Bhugra D, *et al*. Attacks on antidepressants: signs of deep-seated stigma? Lancet Psychiatry 2014; 1(2): 102-4 doi:10.1016/S2215-0366(14)70232-9
3. Naudet F, Falissard B. Antipsychiatry and the antidepressants debate. Lancet Psychiatry 2014; 1(3): 173-4 doi:10.1016/S2215-0366(14)70296-2
4. Moncrieff J. The creation of the concept of an antidepressant: an historical analysis. Soc Sci Med 2008; 66(11): 2346-55
5. Kuhn R. The treatment of depressive states with G 22355 (imipramine hydrochloride). Am J Psychiatry 1958; 115(5): 459-64
6. Stanford SC. Selective serotonin reuptake inhibitors: SSRIS past, present and future: R G Landes Company, 1999
7. Sigg EB. Pharmacological studies with tofranil. Can Psychiatr Assoc J 1959; 4(Suppl): 75-85
8. Burn JH. Tyramine and other amines as noradrenaline-releasing substances. Ciba Foundation Symposium - Aetiology of diabetes mellitus and its complications (Colloquia on Endocrinology). John Wiley & Sons, Ltd., 2008: 326-36
9. Marshall EF, Stirling GS, Tait AC, *et al*. The effect of iproniazid and imipramine on the blood platelet 5-hydroxytrptamine level in man. Br J Pharmacol Chemother 1960; 15: 35-41
10. Axelrod J, Whitby LG, Hertting G. Effect of psychotropic drugs on the uptake of H3-norepinephrine by tissues. Science 1961; 133(3450): 383-4
11. Dengler HJ, Spiegel HE, Titus EO. Uptake of tritium-labeled norepinephrine in brain and other tissues of cat *in vitro*. Science 1961; 133(3458): 1072-3
12. Carlsson A. Some current problems related to the mode of action of antidepressant drugs. Acta Psychiatr Scand Suppl 1981; 290: 63-6
13. Healy D. Prozac, les bonnes fortunes d'une molécule : l'enfance turbulente de la star des antidépresseurs. La recherche 2004; (Hors série numéro 16): 12-22
http://davidhealy.org/wp-content/uploads/2012/05/2004-Les-Bonnes-Fortunes-12-13_HealyProzac.pdf
Accessed December 8th, 2014 (2 pages)
14. Akiskal HS, McKinney WT Jr. Depressive disorders: toward a unified hypothesis. Science 1973; 182(4107): 20-9
15. Association AP, nomenclature APATFo, statistics, nomenclature APACo, DSM-III. APAWGtR. Diagnostic and statistical manual of mental disorders: American Psychiatric Association, 1980
16. Medical Research Council Streptomycin in Tuberculosis Trials Committee. Streptomycin treatment of pulmonary tuberculosis: a medical research council investigation. Br Med J 1948 Oct 30; 2: 769-82
17. Healy D. Clinical care and translational epidemiology: lessons from the antidepressants trials.
http://davidhealy.org/articles/#journalpublications
Accessed December 8th, 2014
18. Thiery M. Clinical trial of the treatment of depressive illness. Report to the medical research council by its clinical psychiatry committee. Br Med J 1965; 1(5439): 881-6
19. Geddes JR. Large simple trials in psychiatry: providing reliable answers to important clinical questions. Epidemiol Psichiatr Soc 2005; 14(3): 122-6
20. Ioannidis JP. Effectiveness of antidepressants: an evidence myth constructed from a thousand randomized trials? Philos Ethics Humanit Med 2008; 3: 14
21. Anderson IM. Selective serotonin reuptake inhibitors *versus* tricyclic antidepressants: a meta-analysis of efficacy and tolerability. J Affect Disord 2000; 58(1): 19-36
22. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979; 134: 382-9
23. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry 1960; 23: 56-62
24. Guy W. ECDEU Assessment manual for psychopharmacology. In: U.S. Department of Health EaW, Public Health Service, Alcohol, Drug Abuse and Mental Health administration, (Ed.). Rockville, Maryland,1976
25. Beck AT, Ward CH, Mendelson M, *et al*. An inventory for measuring depression. Arch Gen Psychiatry 1961; 4: 561-71
26. Mulder RT, Joyce PR, Frampton C. Relationships among measures of treatment outcome in depressed patients. J Affect Disord 2003; 76(1-3): 127-35
27. Frank E, Prien RF, Jarrett RB, *et al*. Conceptualization and rationale for consensus definitions of terms in major depressive disorder. Remission, recovery, relapse, and recurrence. Arch Gen Psychiatry 1991; 48(9): 851-5
28. National Institute for Clinical Excellence. Depression: management of depression in primary and secondary care. Clinical practice guideline No 23. London: NICE; 2004
29. Falissard B, Lukasiewicz M, Corruble E. The MDP75: a new approach in the determination of the minimal clinically meaningful difference in a scale or a questionnaire. J Clin Epidemiol 2003; 56(7): 618-21
30. Leucht S, Fennema H, Engel R, *et al*. What does the HAMD mean? J Affect Disord 2013; 148(2-3): 243-8
31. Naudet F, Falissard B. Does reductio ad absurdum have a place in evidence-based medicine? BMC Med 2014; 12: 106
32. Khan A, Leventhal RM, Khan SR, *et al*. Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration database. J Clin Psychopharmacol 2002; 22(1): 40-5
33. Kirsch I, Deacon BJ, Huedo-Medina TB, *et al*. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. PLoS Med 2008; 5(2): e45
34. Fournier JC, DeRubeis RJ, Hollon SD, *et al*. Antidepressant drug effects and depression severity: a patient-level meta-analysis. JAMA 2010; 303(1): 47-53
35. Gibbons RD, Hur K, Brown CH, *et al*. Benefits from antidepressants: synthesis of 6-week patient-level outcomes from double-blind placebo-controlled randomized trials of fluoxetine and venlafaxine. Arch Gen Psychiatry 2012 Jun; 69(6): 572-9
36. Moncrieff J, Kirsch I. Efficacy of antidepressants in adults. BMJ 2005; 331(7509): 155-7
37. Naudet F, Millet B, Michel Reymann J, *et al*. The fallacy of thresholds used in defining response and remission in depression rating scales. Int J Methods Psychiatr Res 2013 Sep 4 doi: 10.1002/mpr.1393

38. Posternak MA, Zimmerman M. Short-term spontaneous improvement rates in depressed outpatients. J Nerv Ment Dis 2000; 188(12): 799-804
39. Posternak MA, Solomon DA, Leon AC, *et al*. The naturalistic course of unipolar major depression in the absence of somatic therapy. J Nerv Ment Dis 2006; 194(5): 324-9
40. Posternak MA, Zimmerman M. Therapeutic effect of follow-up assessments on antidepressant and placebo response rates in antidepressant efficacy trials: meta-analysis. Br J Psychiatry 2007; 190: 287-92
41. Naudet F, Millet B, Reymann JM, *et al*. Improving study design for antidepressant effectiveness assessment. Int J Methods Psychiatr Res 2013 Aug 30 doi: 10.1002/mpr.1391
42. Seemuller F, Moller HJ, Dittmann S, *et al*. Is the efficacy of psychopharmacological drugs comparable to the efficacy of general medicine medication? BMC Med 2012; 10: 17
43. Hrobjartsson A, Gotzsche PC. Placebo interventions for all clinical conditions. Cochrane Database Syst Rev. 2010; (1): CD003974
44. Kirsch I, Sapirstein G. Listening to Prozac but hearing placebo: A meta-analysis of antidepressant medication. Prevention & Treatment. 1998; 1(2): Article 0002a, posted June 26, 1998 `http://psychrights.org/research/Digest/CriticalThinkRxCites/KirschandSapirstein1998.pdf` Accessed December 8th, 2014 (16 pages)
45. Antonuccio D, Healy D. Relabeling the medications we call antidepressants. Scientifica (Cairo) 2012; 2012: 965908
46. Moncrieff J, Wessely S, Hardy R. Active placebos *versus* antidepressants for depression. Cochrane Database Syst Rev 2004; (1): CD003012
47. Leucht S, Hierl S, Kissling W, *et al*. Putting the efficacy of psychiatric and general medicine medication into perspective: review of meta-analyses. Br J Psychiatry 2012; 200(2): 97-106
48. Jureidini JN, McHenry LB, Mansfield PR. Clinical trials and drug promotion: selective reporting of study 329. Int J Risk Safety Med 20: 73-81
49. Doshi P. Putting GlaxoSmithKline to the test over paroxetine. BMJ 2013 Nov 12; 347: f6754
50. Newman M. The rules of retraction. BMJ 2010 Dec 7;341:c6985 doi: 10.1136/bmj.c6985
51. Kondro W, Sibbald B. Drug company experts advised staff to withhold data about SSRI use in children. CMAJ 2004 Mar 2;170(5):783
52. Keller MB, Ryan ND, Strober M, *et al*. Efficacy of paroxetine in the treatment of adolescent major depression: a randomized, controlled trial. J Am Acad Child Adolesc Psychiatry 2001; 40(7): 762-72
53. Trinquart L, Abbe A, Ravaud P. Impact of reporting bias in network meta-analysis of antidepressant placebo-controlled trials. PLoS One 2012; 7(4): e35219 doi: 10.1371/journal.pone.0035219
54. Naudet F, Millet B, Charlier P, *et al*. Which placebo to cure depression? A thought-provoking network meta-analysis. BMC Med 2013; 11(1): 230
55. Ferguson JM, Mendels J, Schwart GE. Effects of reboxetine on Hamilton Depression Rating Scale factors from randomized, placebo-controlled trials in major depression. Int Clin Psychopharmacol 2002; 17(2): 45-51
56. Eyding D, Lelgemann M, Grouven U, *et al*. Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials. BMJ 2010; 341: c4737
57. Cipriani A, Furukawa TA, Salanti G, *et al*. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. Lancet 2009; 373(9665): 746-58
58. Barbui C, Furukawa TA, Cipriani A. Effectiveness of paroxetine in the treatment of acute major depression in adults: a systematic re-examination of published and unpublished data from randomized trials. CMAJ 2008; 178(3): 296-305
59. Gartlehner G, Hansen RA, Morgan LC, *et al*. Comparative benefits and harms of second-generation antidepressants for treating major depressive disorder: an updated meta-analysis. Ann Intern Med 2011; 155(11): 772-85
60. Alkhafaji AA, Trinquart L, Baron G, *et al*. Impact of evergreening on patients and health insurance: a meta analysis and reimbursement cost analysis of citalopram/escitalopram antidepressants. BMC Med 2012 Nov 20;10:142
61. Arroll B, Elley CR, Fishman T, *et al*. Antidepressants *versus* placebo for depression in primary care. Cochrane Database Syst Rev 2009; (3): CD007954
62. Araya R. The management of depression in primary health care. Curr Opin Psychiatry 1999; 12(1): 103-7
63. Suh T, Gallo JJ. Symptom profiles of depression among general medical service users compared with specialty mental health service users. Psychol Med 1997; 27(5): 1051-63
64. Linde K, Schumann I, Meissner K, *et al*. Treatment of depressive disorders in primary care–protocol of a multiple treatment systematic review of randomized controlled trials. BMC Fam Pract 2011; 12: 127
65. Posternak MA, Zimmerman M, Keitner GI, *et al*. A reevaluation of the exclusion criteria used in antidepressant efficacy trials. Am J Psychiatry 2002; 159(2): 191-200
66. Zimmerman M, Chelminski I, Posternak MA. Generalizability of antidepressant efficacy trials: differences between depressed psychiatric outpatients who would or would not qualify for an efficacy trial. Am J Psychiatry 2005; 162(7): 1370-2
67. Zimmerman M, Mattia JI, Posternak MA. Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? Am J Psychiatry 2002; 159(3): 469-73
68. National Institute for Clinical Excellence. The treatment and management of depression in adults (updated edition) National Clinical Practice Guideline 90. The British Psychological Society and The Royal College of Psychiatrists, 2010 `http://www.nice.org.uk/guidance/cg90/resources/cg90-depression-in-adults-full-guidance2` Accessed December 8th, 2014 (707 pages)
69. Seemuller F, Moller HJ, Obermeier M, *et al*. Do efficacy and effectiveness samples differ in antidepressant treatment outcome? An analysis of eligibility criteria in randomized controlled trials. J Clin Psychiatry 2010; 71(11): 1425-33
70. Wisniewski SR, Rush AJ, Nierenberg AA, *et al*. Can phase III trial results of antidepressant medications be generalized to clinical practice? A STAR*D report. Am J Psychiatry 2009; 166(5): 599-607
71. Naudet F, Maria AS, Falissard B. Antidepressant response in major depressive disorder: a meta-regression comparison of randomized controlled trials and observational studies. PLoS One 2011; 6(6): e20811
72. van der Lem R, van der Wee NJ, van Veen T, *et al*. Efficacy *versus* effectiveness: a direct comparison of the outcome of treatment for mild to moderate depression in randomized controlled trials and daily practice. Psychother Psychosom 2012; 81(4): 226-34
73. Koesters M, Guaiana G, Cipriani A, *et al*. Agomelatine efficacy and acceptability revisited: systematic review and meta-analysis of published and unpublished randomised trials. Br J Psychiatry 2013; 203(3): 179-87
74. Taylor D, Sparshatt A, Varma S, Olofinjana O. Antidepressant efficacy of agomelatine: meta-analysis of published and unpublished studies. BMJ 2014; 348: g1888

Correspondence and offprints: *Florian Naudet*, INSERM U669, Maison de Solenn, 97 boulevard de Port Royal, 75679 Paris cedex 14, France.
E-mail: floriannaudet@gmail.com

