



HAL
open science

Comprendre le débat sur les antidépresseurs dans le traitement de l'épisode dépressif majeur [Understanding the Antidepressant Debate in the Treatment of Major Depressive Disorder]

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

► To cite this version:

Florian Naudet, Rémy Bousageon, Clément Palpacuer, Laurent Gallet, Jean-Michel Reymann, et al.. Comprendre le débat sur les antidépresseurs dans le traitement de l'épisode dépressif majeur [Understanding the Antidepressant Debate in the Treatment of Major Depressive Disorder]. *Thérapie*, 2015, 9 (19), pp.140028. 10.2515/therapie/2014228 . hal-01147040

HAL Id: hal-01147040

<https://univ-rennes.hal.science/hal-01147040v1>

Submitted on 23 Sep 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Understanding the Antidepressant Debate in the Treatment of Major Depressive Disorder

Florian Naudet^{1,2,3}, Rémy Boussageon⁴, Clément Palpacuer², Laurent Gallet³, 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 zimelidone, a new treatment blocking the uptake of serotonin without blocking the uptake of catecholamines.^[12] While zimelidone 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 acceptance 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 anti-psychiatry 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

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

7 **4. The mutual influence** 8 **of psychopharmacology and trial** 9 **methodology**

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

Scale (HDRS-17) and response is usually defined as a reduction
 of 50% on the HDRS-17. 46 47

5. Statistically significant versus clinically meaningful results 48 49

50 While these methodological tools enable the measurement
 51 of outcomes and test whether antidepressants provide statistically
 52 significant benefits as compared with a placebo, there is a considerable
 53 debate concerning the real meaning of the difference in term of its
 54 clinical significance. Indeed, the identification of a minimal clinically
 55 relevant difference on a scale is not straightforward. In 2004, the
 56 National Institute of Clinical Excellence^[28] stated that a Hamilton
 57 score difference of three points across groups could be considered
 58 as clinically significant. This threshold was consistent with previous
 59 research^[29] but a recent linking analysis provided new insight
 60 by suggesting that a slight reduction on the HAM-D-17 of up to
 61 3 points corresponds to a rating of “no change” as measured with
 62 the CGI. A change close to 10 points was linked to the “much
 63 improved” category defined by the CGI.^[30] But these considerations
 64 on an individual level are not totally transposable to group level.
 65 On the other hand, this study also suggested that the commonly
 66 used measures for response (1) and remission (2) in MDD trials
 67 could reasonably be considered valid because they were coherent
 68 with the CGI definitions “much improved” (1) and “not at all” or
 69 “borderline mentally ill” (2), respectively. Bearing in mind that
 70 the CGI is not a perfect gold standard, these results are very
 71 interesting.

6. RCTs and the dilution of efficacy 72

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

1 of aggregated data meta-analyses, but since personal data are dif- 47
 2 ficult to collect, it was prone to publication bias. Nevertheless, 48
 3 these three meta-analyses concluded consistently that the distinc- 49
 4 tion between antidepressants and placebo is clinically meaningful 50
 5 (using the National Institute for Clinical Excellence threshold for 51
 6 clinical significance) only for severe and very severe patients. 52

7 Interestingly, Gibbons *et al.*^[35] addressed the limitations of 53
 8 the preceding studies by reanalysing all intent-to-treat individual 54
 9 longitudinal data during the first 6 weeks of treatment for major 55
 10 depressive disorder from all sponsored randomized controlled trials 56
 11 on fluoxetine and venlafaxine. In this meta-analysis, average 57
 12 differences at 6 weeks were small and not clinically meaningful 58
 13 (2.5 HAM-D units) and baseline severity was not shown to af- 59
 14 fect symptom reduction. But these small overall mean differences 60
 15 translated into clinically significant differences in response rates 61
 16 (estimated response rates were 58.4% for drug *versus* 39.9% for 62
 17 placebo) and remission rates (59.1% for drug *versus* 41.9% for 63
 18 placebo, relative risk = 1.5, number needed to treat = 5). This 64
 19 finding seems surprising. Intuitively, the two methods of assess- 65
 20 ing outcome should produce similar conclusions, since they are
 21 derived from the same data. However, this result can be explained
 22 by an artefact inherent in the transformation of continuous data
 23 into categorical data, which can magnify small differences.^[36] But
 24 on the other hand, transformation of continuous outcomes into
 25 categorical outcomes implies a misclassification bias, and mea-
 26 sures of association such as relative risk are likely to be biased to-
 27 wards 1.^[37] An alternative explanation is that “efficacy dilution”
 28 is at play here.

29 7. Antidepressant alibis

30 In all events, beyond any fuzziness concerning the interpreta- 67
 31 tion of antidepressant efficacy in MDD, a large number of RCTs 68
 32 turn out negative. It is frequently suggested that this is due to a 69
 33 marked placebo response in antidepressant trials, which could re- 70
 34 sult from many different factors, such as spontaneous improve- 71
 35 ment,^[38] statistical regression to the mean, low level of severity 72
 36 at inclusion, co-interventions, and other biases in addition to the 73
 37 so-called placebo effect. For example, spontaneous improvement 74
 38 is common in clinical practice,^[38,39] and the number of follow- 75
 39 up assessments^[40] is related to a significant therapeutic effect.^[41] 76
 40 From a naïve point of view, one might have expected that in 77
 41 MDD, since it is a “mental disorder”, the placebo effect (with 78
 42 its psychological component) might be greater than in other con- 79
 43 ditions and, as a consequence, the resulting true “pharmacologi- 80
 44 cal” effect would be weaker than in general medicine. However 81
 45 the distinction is probably more subtle.^[42] In a meta-analysis,^[43] 82
 46 Hrobjartsson *et al.* identified no statistically significant effect of 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93

placebo interventions in depression, while a meta-analysis by 47
 Kirsh *et al.* suggested that placebo effects were considerable.^[44] 48
 But the RCTs included in Hrobjartsson’s meta-analysis were not 49
 designed (they were underpowered) to study the placebo effect 50
 adequately. Similarly, in Kirsch’s meta-analysis, which comprises 51
 no “untreated group” or waiting list, we cannot determine the size 52
 of the placebo effect. There is thus considerable debate about the 53
 size, the nature and the mechanism of the placebo effect in de- 54
 pression.^[42] For example, it has been proposed that the apparent 55
 antidepressant effect could be in part an active placebo effect, or 56
 result from bias, since side effects like sexual effects^[45] of an- 57
 tidepressants could reveal the identity of the medication to partic- 58
 ipants or investigators.^[46] 59

60 Nonetheless, while some general medical drugs have very 60
 high effect sizes, the effect sizes obtained by psychiatric drugs are 61
 in the same range as most general medical pharmaceuticals.^[47] 62
 Although it is difficult to compare effect sizes of drugs in different 63
 conditions, indications and outcomes, this finding puts the small 64
 effect sizes observed with antidepressants into perspective. 65

66 8. Overestimation and distortion of efficacy

67 Antidepressants efficacy is nonetheless certainly overesti- 67
 68 mated in the published literature by selective publication and se- 68
 69 lective outcome reporting. To explore this phenomenon, Turner 69
 70 *et al.* performed an analysis of 74 studies that were submitted 70
 71 to FDA for the approval of 12 antidepressant drugs. Among these 71
 72 studies, the FDA considered that 38 (51%) were “positive” (with a 72
 73 statistically significant result on the principal outcome), 12 (16%) 73
 74 “indeterminate” and 24 (33%) “negative” (with no statistically 74
 75 significant result on the principal outcome). Among the “posi- 75
 76 tive” studies, 37 (97%) were published and only one (3%) was 76
 77 not published. Among the “indeterminate” studies, 6 (50%) were 77
 78 published as positive and 6 (50%) were unpublished. Finally, of 78
 79 the “negative” studies, 3 (12%) were published as “negative”, in 79
 80 agreement with the opinion of the FDA, 5 (21%) were published 80
 81 as “positive”, in disagreement with the opinion of FDA and 16 81
 82 (67%) were not published. The effect size measured by perform- 82
 83 ing a meta-analysis on the basis of published results is 0.41 with 83
 84 a 95% confidence interval of [0.36-0.45], whereas it is estimated 84
 85 to be 0.31 with a 95% confidence interval of [0.27-0.35] based on 85
 86 all studies reported to FDA. 86

87 The best-documented case of selective outcome reporting 87
 88 is probably study 329.^[48-50] It was a large study of 275 de- 88
 89 pressed adolescents conducted by SmithKline Beecham in the US 89
 90 from 1993-1996. Its results failed to show any statistically sig- 90
 91 nificant difference between paroxetine and placebo for the two 91
 92 primary outcomes. A GSK internal document stated that the re- 92
 93 sults of study 329 indicated paroxetine was no more effective than 93

1 placebo, and provided guidance on how to manage these disap-
 2 pointing results by recommending they should “effectively man-
 3 age the dissemination of these data in order to minimize any po-
 4 tential negative commercial impact.” It also stated that “it would
 5 be commercially unacceptable to include a statement that efficacy
 6 had not been demonstrated, as this would undermine the profile
 7 of paroxetine.”^[51] Subsequently, an article was written (or more
 8 precisely ghostwritten) with positive results concerning new sec-
 9 ondary outcome measures that had been introduced. It was con-
 10 cluded that paroxetine is “generally well tolerated and effective
 11 for major depression in adolescents.”^[52]

9. Paradoxes in comparative effectiveness assessments

14 As a result of selective outcome reporting of this type, meta-
 15 analyses are likely to give misleading impressions about efficacy
 16 and comparative effectiveness of antidepressants.^[53,54] There is
 17 the case of reboxetine, a selective norepinephrine reuptake in-
 18 hibitor used in the treatment of depression. The previously favor-
 19 able risk-benefit profile of reboxetine shown in published trials^[55]
 20 was reversed by the addition of unpublished data.^[56] In a network
 21 meta-analysis performed by Cipriani *et al.*, reboxetine was consis-
 22 tently shown to be worse than 11 other antidepressants,^[57] includ-
 23 ing paroxetine which was however found in another meta-analysis
 24 by the same team not to have any superiority over placebo.^[58] All
 25 in all, these meta-analyses appear paradoxical, giving the impres-
 26 sion that paroxetine is not superior to placebo, while it does better
 27 than reboxetine, which has itself been shown not to be superior
 28 to placebo. Additionally, although the Cipriani study found differ-
 29 ences between antidepressants, this was not the case for another
 30 network meta-analysis performed by Gartlehner *et al.*^[59]

31 Another paradox has been shown in a recent paper compar-
 32 ing citalopram with its “me-too”, escitalopram, which found an
 33 inconsistency between direct evidence (showing a superiority of
 34 escitalopram) and indirect evidence (which did not find any sig-
 35 nificant difference).^[60]

10. Poor transposability of RCT results

37 Beyond these issues RCTs are often criticised for their lack of
 38 external validity. Indeed, the vast majority of patients with clinical
 39 depression are catered for in primary care, and most RCTs have
 40 involved secondary care patients.^[61] These patients probably dif-
 41 fer from primary care patients.^[62,63] in terms of severity (primary
 42 care patients are less severely depressed, milder course of illness)
 43 and in terms of complaints (fatigue and somatic symptoms).^[64]
 44 Additionally, antidepressant RCTs use numerous non-inclusion

criteria (for example suicidal ideations)^[65–67] and excluded pa- 45
 46 tients are a more chronically ill group with more numerous previ-
 47 ous episodes, greater psychosocial impairment, and more frequent
 48 personality disorders. Finally, the vast majority of RCTs last no
 49 more than 8 weeks, whereas it is recommended that an antidepres-
 50 sant treatment be continued for at least 6 months after remission
 51 of the episode.^[68]

52 There is debate as to whether these issues can be trans-
 53 lated into different outcomes between RCTs and a “real life”
 54 setting.^[69–72]

11. Conclusion

56 While meta-analyses should be reproducible, in 2013, a meta-
 57 analysis of published and unpublished studies on agomelatine
 58 found “evidence suggesting that a clinically important difference
 59 between agomelatine and placebo in patients with unipolar major
 60 depression was unlikely”;^[73] in 2014 a meta-analysis of pub-
 61 lished and unpublished studies on agomelatine found that it “was
 62 an effective antidepressant with similar efficacy to standard an-
 63 tidepressants”.^[74] This particular paradox sums up the fuzziness
 64 of antidepressant literature. We suggest that, instead of providing
 65 easy answers, the application of the RCT paradigm to MDD raises
 66 many questions. This is probably in the nature of all scientific
 67 studies, but it can be in contradiction with clinicians’ expectations:
 68 what they want is to be sure that the treatment will work for indi-
 69 vidual patients (or to know if it will not). At the same time, their
 70 clinical experience is biased by many other parameters, including
 71 placebo response. This is precisely where the debate arises.

72 *Acknowledgements.* We thank Angela Swaine Verdier for revising the
 73 English.

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

79 **Conflict of interests.** There are no conflicts of interest regarding
 80 this paper. All authors have completed the Unified Competing In-
 81 terest form at http://www.icmje.org/coi_disclosure.pdf
 82 (available on request from the corresponding author) and de-
 83 clare that (1) No authors have support from any company for the
 84 submitted work; (2) N.F. has relationships (board membership
 85 or travel/accommodations expenses covered/reimbursed) with
 86 Servier, BMS, Lundbeck and Janssen who might have an inter-
 87 est in the work submitted in the previous 3 years ; B.R., P.C. G.L.
 88 and R.J.M. have no relationships with any company that might

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

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

17 **References**

18 1. Göttsche PC. Why I think antidepressants cause more harm than good. *Lancet*
 19 *Psychiatry* 2014; 1(2): 104-6 doi:10.1016/S2215-0366(14)70280-9
 20 2. Nutt DJ, Goodwin GM, Bhugra D, et al. Attacks on antidepressants:
 21 signs of deep-seated stigma? *Lancet Psychiatry* 2014; 1(2): 102-4
 22 doi:10.1016/S2215-0366(14)70232-9
 23 3. Naudet F, Falissard B. Antipsychiatry and the antidepressants debate. *Lancet*
 24 *Psychiatry* 2014; 1(3): 173-4 doi:10.1016/S2215-0366(14)70296-2
 25 4. Moncrieff J. The creation of the concept of an antidepressant: an historical
 26 analysis. *Soc Sci Med* 2008; 66(11): 2346-55
 27 5. Kuhn R. The treatment of depressive states with G 22355 (imipramine hy-
 28 drochloride). *Am J Psychiatry* 1958; 115(5): 459-64
 29 6. Stanford SC. Selective serotonin reuptake inhibitors: SSRIS past, present and
 30 future: R G Landes Company, 1999
 31 7. Sigg EB. Pharmacological studies with tofranil. *Can Psychiatr Assoc J* 1959;
 32 4(Suppl): 75-85
 33 8. Burn JH. Tyramine and other amines as noradrenaline-releasing substances.
 34 Ciba Foundation Symposium - Aetiology of diabetes mellitus and its com-
 35 plications (Colloquia on Endocrinology). John Wiley & Sons, Ltd., 2008:
 36 326-36
 37 9. Marshall EF, Stirling GS, Tait AC, et al. The effect of iproniazid and
 38 imipramine on the blood platelet 5-hydroxytryptamine level in man. *Br J*
 39 *Pharmacol Chemother* 1960; 15: 35-41
 40 10. Axelrod J, Whitby LG, Hertting G. Effect of psychotropic drugs on the uptake
 41 of H3-norepinephrine by tissues. *Science* 1961; 133(3450): 383-4
 42 11. Dengler HJ, Spiegel HE, Titus EO. Uptake of tritium-labeled norepinephrine
 43 in brain and other tissues of cat *in vitro*. *Science* 1961; 133(3458): 1072-3
 44 12. Carlsson A. Some current problems related to the mode of action of antide-
 45 pressant drugs. *Acta Psychiatr Scand Suppl* 1981; 290: 63-6
 46 13. Healy D. Prozac, les bonnes fortunes d'une molécule : l'enfance turbulente
 47 de la star des antidépresseurs. *La recherche* 2004; (Hors série numéro 16):
 48 12-22
 49 [http://davidhealy.org/wp-content/uploads/2012/05/
 50 2004-Les-Bonnes-Fortunes-12-13_HealyProzac.pdf](http://davidhealy.org/wp-content/uploads/2012/05/2004-Les-Bonnes-Fortunes-12-13_HealyProzac.pdf)
 51 Accessed December 8th, 2014 (2 pages)
 52 14. Akiskal HS, McKinney WT Jr. Depressive disorders: toward a unified hypoth-
 53 esis. *Science* 1973; 182(4107): 20-9

15. Association AP, nomenclature APATFo, statistics, nomenclature APACo, 54
 DSM-III. APAWGtR. Diagnostic and statistical manual of mental disor- 55
 ders: American Psychiatric Association, 1980 56
 16. Medical Research Council Streptomycin in Tuberculosis Trials Committee. 57
 Streptomycin treatment of pulmonary tuberculosis: a medical research 58
 council investigation. *Br Med J* 1948 Oct 30; 2: 769-82 59
 17. Healy D. Clinical care and translational epidemiology: lessons from the an- 60
 tidepressants trials. 61
<http://davidhealy.org/articles/#journalpublications> 62
 Accessed December 8th, 2014 63
 18. Thiery M. Clinical trial of the treatment of depressive illness. Report to the 64
 medical research council by its clinical psychiatry committee. *Br Med J* 65
 1965; 1(5439): 881-6 66
 19. Geddes JR. Large simple trials in psychiatry: providing reliable answers to 67
 important clinical questions. *Epidemiol Psychiatr Soc* 2005; 14(3): 122-6 68
 20. Ioannidis JP. Effectiveness of antidepressants: an evidence myth constructed 69
 from a thousand randomized trials? *Philos Ethics Humanit Med* 2008; 3: 70
 14 71
 21. Anderson IM. Selective serotonin reuptake inhibitors versus tricyclic antide- 72
 pressants: a meta-analysis of efficacy and tolerability. *J Affect Disord* 2000; 73
 58(1): 19-36 74
 22. Montgomery SA, Asberg M. A new depression scale designed to be sensitive 75
 to change. *Br J Psychiatry* 1979; 134: 382-9 76
 23. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 77
 1960; 23: 56-62 78
 24. Guy W. ECDEU Assessment manual for psychopharmacology. In: U.S. 79
 Department of Health EaW, Public Health Service, Alcohol, Drug Abuse 80
 and Mental Health administration, (Ed.). Rockville, Maryland, 1976 81
 25. Beck AT, Ward CH, Mendelson M, et al. An inventory for measuring depres- 82
 sion. *Arch Gen Psychiatry* 1961; 4: 561-71 83
 26. Mulder RT, Joyce PR, Frampton C. Relationships among measures of treat- 84
 ment outcome in depressed patients. *J Affect Disord* 2003; 76(1-3): 127-35 85
 27. Frank E, Prien RF, Jarrett RB, et al. Conceptualization and rationale for con- 86
 sensus definitions of terms in major depressive disorder. Remission, recov- 87
 ery, relapse, and recurrence. *Arch Gen Psychiatry* 1991; 48(9): 851-5 88
 28. National Institute for Clinical Excellence. Depression: management of de- 89
 pression in primary and secondary care. Clinical practice guideline No 23. 90
 London: NICE; 2004 91
 29. Falissard B, Lukasiewicz M, Corruble E. The MDP75: a new approach in the 92
 determination of the minimal clinically meaningful difference in a scale or 93
 a questionnaire. *J Clin Epidemiol* 2003; 56(7): 618-21 94
 30. Leucht S, Fennema H, Engel R, et al. What does the HAMD mean? *J Affect* 95
Disord 2013; 148(2-3): 243-8 96
 31. Naudet F, Falissard B. Does reductio ad absurdum have a place in evidence- 97
 based medicine? *BMC Med* 2014; 12: 106 98
 32. Khan A, Leventhal RM, Khan SR, et al. Severity of depression and re- 99
 sponse to antidepressants and placebo: an analysis of the Food and Drug 100
 Administration database. *J Clin Psychopharmacol* 2002; 22(1): 40-5 101
 33. Kirsch I, Deacon BJ, Huedo-Medina TB, et al. Initial severity and antide- 102
 pressant benefits: a meta-analysis of data submitted to the Food and Drug 103
 Administration. *PLoS Med* 2008; 5(2): e45 104
 34. Fournier JC, DeRubeis RJ, Hollon SD, et al. Antidepressant drug effects and 105
 depression severity: a patient-level meta-analysis. *JAMA* 2010; 303(1): 106
 47-53 107
 35. Gibbons RD, Hur K, Brown CH, et al. Benefits from antidepressants: synthe- 108
 sis of 6-week patient-level outcomes from double-blind placebo-controlled 109
 randomized trials of fluoxetine and venlafaxine. *Arch Gen Psychiatry* 2012 110
 Jun; 69(6): 572-9 111
 36. Moncrieff J, Kirsch I. Efficacy of antidepressants in adults. *BMJ* 2005; 112
 331(7509): 155-7 113
 37. Naudet F, Millet B, Michel Reymann J, et al. The fallacy of thresholds used in 114
 defining response and remission in depression rating scales. *Int J Methods* 115
Psychiatr Res 2013 Sep 4 doi: 10.1002/mpr.1393 116

- 1 38. Posternak MA, Zimmerman M. Short-term spontaneous improvement rates in
2 depressed outpatients. *J Nerv Ment Dis* 2000; 188(12): 799-804
- 3 39. Posternak MA, Solomon DA, Leon AC, *et al.* The naturalistic course of unipo-
4 lar major depression in the absence of somatic therapy. *J Nerv Ment Dis*
5 2006; 194(5): 324-9
- 6 40. Posternak MA, Zimmerman M. Therapeutic effect of follow-up assessments
7 on antidepressant and placebo response rates in antidepressant efficacy tri-
8 als: meta-analysis. *Br J Psychiatry* 2007; 190: 287-92
- 9 41. Naudet F, Millet B, Reymann JM, *et al.* Improving study design for antide-
10 pressant effectiveness assessment. *Int J Methods Psychiatr Res* 2013 Aug
11 30 doi: 10.1002/mpr.1391
- 12 42. Seemuller F, Moller HJ, Dittmann S, *et al.* Is the efficacy of psychopharmacolo-
13 gical drugs comparable to the efficacy of general medicine medication?
14 *BMC Med* 2012; 10: 17
- 15 43. Hrobjartsson A, Gotzsche PC. Placebo interventions for all clinical condi-
16 tions. *Cochrane Database Syst Rev.* 2010; (1): CD003974
- 17 44. Kirsch I, Sapirstein G. Listening to Prozac but hearing placebo: A meta-
18 analysis of antidepressant medication. *Prevention & Treatment.* 1998;
19 1(2): Article 0002a, posted June 26, 1998
20 [http://psychrights.org/research/Digest/](http://psychrights.org/research/Digest/CriticalThinkRxCites/KirschandSapirstein1998.pdf)
21 [CriticalThinkRxCites/KirschandSapirstein1998.pdf](http://psychrights.org/research/Digest/CriticalThinkRxCites/KirschandSapirstein1998.pdf)
22 Accessed December 8th, 2014 (16 pages)
- 23 45. Antonuccio D, Healy D. Relabeling the medications we call antidepressants.
24 *Scientifica (Cairo)* 2012; 2012: 965908
- 25 46. Moncrieff J, Wessely S, Hardy R. Active placebos *versus* antidepressants for
26 depression. *Cochrane Database Syst Rev* 2004; (1): CD003012
- 27 47. Leucht S, Hierl S, Kissling W, *et al.* Putting the efficacy of psychiatric and
28 general medicine medication into perspective: review of meta-analyses. *Br*
29 *J Psychiatry* 2012; 200(2): 97-106
- 30 48. Jureidini JN, McHenry LB, Mansfield PR. Clinical trials and drug promotion:
31 selective reporting of study 329. *Int J Risk Safety Med* 20: 73-81
- 32 49. Doshi P. Putting GlaxoSmithKline to the test over paroxetine. *BMJ* 2013 Nov
33 12; 347: f6754
- 34 50. Newman M. The rules of retraction. *BMJ* 2010 Dec 7;341:c6985 doi:
35 10.1136/bmj.c6985
- 36 51. Kondro W, Sibbald B. Drug company experts advised staff to withhold data
37 about SSRI use in children. *CMAJ* 2004 Mar 2;170(5):783
- 38 52. Keller MB, Ryan ND, Strober M, *et al.* Efficacy of paroxetine in the treatment
39 of adolescent major depression: a randomized, controlled trial. *J Am Acad*
40 *Child Adolesc Psychiatry* 2001; 40(7): 762-72
- 41 53. Trinquart L, Abbe A, Ravaud P. Impact of reporting bias in network meta-
42 analysis of antidepressant placebo-controlled trials. *PLoS One* 2012; 7(4):
43 e35219 doi: 10.1371/journal.pone.0035219
- 44 54. Naudet F, Millet B, Charlier P, *et al.* Which placebo to cure depression? A
45 thought-provoking network meta-analysis. *BMC Med* 2013; 11(1): 230
- 46 55. Ferguson JM, Mendels J, Schwart GE. Effects of reboxetine on Hamilton
47 Depression Rating Scale factors from randomized, placebo-controlled tri-
48 als in major depression. *Int Clin Psychopharmacol* 2002; 17(2): 45-51
- 49 56. Eyding D, Lelgemann M, Grouven U, *et al.* Reboxetine for acute treatment
50 of major depression: systematic review and meta-analysis of published and
51 unpublished placebo and selective serotonin reuptake inhibitor controlled
52 trials. *BMJ* 2010; 341: c4737
- 53 57. Cipriani A, Furukawa TA, Salanti G, *et al.* Comparative efficacy and accept-
54 ability of 12 new-generation antidepressants: a multiple-treatments meta-
55 analysis. *Lancet* 2009; 373(9665): 746-58
- 56 58. Barbui C, Furukawa TA, Cipriani A. Effectiveness of paroxetine in the treat-
57 ment of acute major depression in adults: a systematic re-examination
58 of published and unpublished data from randomized trials. *CMAJ* 2008;
59 178(3): 296-305
- 60 59. Gartlehner G, Hansen RA, Morgan LC, *et al.* Comparative benefits and harms
61 of second-generation antidepressants for treating major depressive disor-
62 der: 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
63 and health insurance: a meta analysis and reimbursement cost analysis of
64 citalopram/escitalopram antidepressants. *BMC Med* 2012 Nov 20;10:142
65
61. Arroll B, Elley CR, Fishman T, *et al.* Antidepressants *versus* placebo for
66 depression in primary care. *Cochrane Database Syst Rev* 2009; (3):
67 CD007954
68
62. Araya R. The management of depression in primary health care. *Curr Opin*
69 *Psychiatry* 1999; 12(1): 103-7
70
63. Suh T, Gallo JJ. Symptom profiles of depression among general medical ser-
71 vice users compared with specialty mental health service users. *Psychol*
72 *Med* 1997; 27(5): 1051-63
73
64. Linde K, Schumann I, Meissner K, *et al.* Treatment of depressive disorders
74 in primary care—protocol of a multiple treatment systematic review of ran-
75 domized controlled trials. *BMC Fam Pract* 2011; 12: 127
76
65. Posternak MA, Zimmerman M, Keitner GI, *et al.* A reevaluation of the exclu-
77 sion criteria used in antidepressant efficacy trials. *Am J Psychiatry* 2002;
78 159(2): 191-200
79
66. Zimmerman M, Chelminski I, Posternak MA. Generalizability of antidepres-
80 sant efficacy trials: differences between depressed psychiatric outpatients
81 who would or would not qualify for an efficacy trial. *Am J Psychiatry* 2005;
82 162(7): 1370-2
83
67. Zimmerman M, Mattia JI, Posternak MA. Are subjects in pharmacological
84 treatment trials of depression representative of patients in routine clinical
85 practice? *Am J Psychiatry* 2002; 159(3): 469-73
86
68. National Institute for Clinical Excellence. The treatment and management
87 of depression in adults (updated edition) National Clinical Practice
88 Guideline 90. The British Psychological Society and The Royal College of
89 Psychiatrists, 2010
90 [http://www.nice.org.uk/guidance/cg90/resources/](http://www.nice.org.uk/guidance/cg90/resources/cg90-depression-in-adults-full-guidance2)
91 [cg90-depression-in-adults-full-guidance2](http://www.nice.org.uk/guidance/cg90/resources/cg90-depression-in-adults-full-guidance2)
92 Accessed December 8th, 2014 (707 pages)
93
69. Seemuller F, Moller HJ, Obermeier M, *et al.* Do efficacy and effectiveness
94 samples differ in antidepressant treatment outcome? An analysis of eligibil-
95 ity criteria in randomized controlled trials. *J Clin Psychiatry* 2010; 71(11):
96 1425-33
97
70. Wisniewski SR, Rush AJ, Nierenberg AA, *et al.* Can phase III trial results of
98 antidepressant medications be generalized to clinical practice? A STAR*D
99 report. *Am J Psychiatry* 2009; 166(5): 599-607
100
71. Naudet F, Maria AS, Falissard B. Antidepressant response in major depressive
101 disorder: a meta-regression comparison of randomized controlled
102 trials and observational studies. *PLoS One* 2011; 6(6): e20811
103
104
72. van der Lem R, van der Wee NJ, van Veen T, *et al.* Efficacy *versus* effective-
105 ness: a direct comparison of the outcome of treatment for mild to moderate
106 depression in randomized controlled trials and daily practice. *Psychother*
107 *Psychosom* 2012; 81(4): 226-34
108
73. Koesters M, Guaiana G, Cipriani A, *et al.* Agomelatine efficacy and accept-
109 ability revisited: systematic review and meta-analysis of published and un-
110 published randomised trials. *Br J Psychiatry* 2013; 203(3): 179-87
111
74. Taylor D, Sparshatt A, Varma S, Olofinjana O. Antidepressant efficacy of
112 agomelatine: meta-analysis of published and unpublished studies. *BMJ*
113 2014; 348: g1888
114

Correspondence and offprints: Florian Naudet, INSERM U669, Maison de 115
Solenn, 97 boulevard de Port Royal, 75679 Paris cedex 14, France. 116
E-mail: floriannaudet@gmail.com 117