Psychological Medicine

cambridge.org/psm

Editorial

Cite this article: Ormel J, Spinhoven P, de Vries YA, Cramer AOJ, Siegle GJ, Bockting CLH, Hollon SD (2020). The antidepressant standoff: why it continues and how to resolve it. *Psychological Medicine* **50**, 177–186. https://doi.org/10.1017/S0033291719003295

Received: 17 December 2018
Revised: 8 August 2019
Accepted: 28 October 2019
First published online: 29 November 2019

Key words:
Antidepressants; long-term effects; major depressive disorder; psychotherapy; recurrence; relapse

Author for correspondence:
Johan Ormel,
E-mail: J.ormel@umcg.nl

© Cambridge University Press 2019

CAMBRIDGE UNIVERSITY PRESS

The antidepressant standoff: why it continues and how to resolve it

Johan Ormel[1,2], Philip Spinhoven[2,3], Ymkje Anna de Vries[4], Angélique O. J. Cramer[2,5], Greg J. Siegle[6], Claudi L. H. Bockting[2,7] and Steven D. Hollon[8]

[1]Department of Psychiatry, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; [2]Netherlands Institute for Advanced Study KNAW, Amsterdam, The Netherlands; [3]Department of Psychiatry, Leiden University, Institute of Psychology, Leiden, The Netherlands; [4]Department of Developmental Psychology, University of Groningen, Groningen, The Netherlands; [5]Department of Methodology and Statistics, Tilburg School of Social and Behavioral Sciences, Tilburg University, Tilburg, The Netherlands; [6]University of Pittsburgh, School of Medicine, Pittsburgh, PA, USA; [7]Department of Psychiatry, University of Amsterdam, Amsterdam University Medical Centres, AMC, Amsterdam, The Netherlands and [8]Department of Psychology, Vanderbilt University, Nashville, Tennessee, USA

Abstract

Background. Antidepressant medications (ADMs) are widely used and long-term use is increasing. Given this extensive use and recommendation of ADMs in guidelines, one would expect ADMs to be universally considered effective. Surprisingly, that is not the case; fierce debate on their benefits and harms continues. This editorial seeks to understand why the controversy continues and how consensus can be achieved.

Methods. 'Position' paper. Critical analysis and synthesis of relevant literature.

Results. Advocates point at ADMs impressive effect size (number needed to treat, NNT = 6–8) in acute phase treatment and continuation/maintenance ADM treatment prevention relapse/recurrence in acute phase ADM responders (NNT = 3–4). Critics point at the limited clinically significant surplus value of ADMs relative to placebo and argue that effectiveness is overstated. We identified multiple factors that fuel the controversy: certainty of evidence is low to moderate; modest efficacy on top of strong placebo effects allows critics to focus on small net efficacy and advocates on large gross efficacy; ADM withdrawal symptoms masquerade as relapse/recurrence; lack of association between ADM treatment and long-term outcome in observational databases. Similar problems affect psychological treatments as well, but less so. We recommend four approaches to resolve the controversy: (1) placebo-controlled trials with relevant long-term outcome assessments, (2) inventive analyses of observational databases, (3) patient cohort studies including effect moderators to improve personalized treatment, and (4) psychological treatments as universal first-line treatment step.

Conclusions. Given the public health significance of depression and increased long-term ADM usage, new approaches are needed to resolve the controversy.

Introduction

Antidepressant medications (ADMs) are widely prescribed and maintenance of ADM treatment is increasing. In the USA, 12% of the population 12 years and older utilize ADMs, and the percentage using ADMs ⩾24 months more than doubled from 3% to 7% between 1999 and 2010 (Mojtabai & Olfson, 2014; Pratt, Brody, & Gu, 2017). Similar trends have been observed in the UK (Lockhart & Guthrie, 2011) and the Netherlands (Wildeboer, van der Hoek, & Verhaak, 2016). Given this utilization, the hundreds of double-blind placebo-controlled AD trials (Cipriani et al., 2018), and the prominent recommendation of ADMs in clinical guidelines, one would expect that ADMs are universally considered effective and worth their side effects. Surprisingly, that is not the case: fierce debate on their effectiveness continues in the scientific literature, media, and Internet (Moncrieff, 2018a). Here we explore the positions of the advocates and critics and examine why the debate continues and how it is best resolved. In addition, we examine the efficacy of evidence-based psychological treatments and their vulnerability to criticism.

Advocates point at the significant effect size of acute phase ADM treatment (aADM) of depression [number needed to treat (NNT) = 6–8], the ability of continuation and maintenance antidepressants (cmADs) to prevent relapse/recurrence in those who remit on aADM (NNT = 3–4), and the reduced risk of suicide (attempts) associated with ADM treatment (Bridge, Barbe, Birmaher, Kolko, & Brent, 2005; Cooney et al., 2013; Geddes & Cipriani, 2015; Leucht, Hierl, Kissling, Dold, & Davis, 2012; Nutt, Goodwin, Bhugra, Fazel, & Lawrie, 2014; Quitkin, Rabkin, Gerald, Davis, & Klein, 2000). Advocates acknowledge individual

differences in response and that many do not respond to ADMs as well as significant side effects and dropouts. However, they argue that most patients on cmADM choose to stay on ADMs, suggesting that the benefits prevail, and that good clinical management can keep dropouts and adverse side effects low.

Critics doubt whether antidepressants have clinically significant surplus value over and above placebo (PLA) (Antonuccio, Danton, DeNelsky, Greenberg, & Gordon, 1999; Fava, 2003; Gotschke, 2013; Kirsch, 2014; Kirsch et al., 2008; Moncrieff, 2002, 2018b; Moncrieff & Kirsch, 2015; Turner, Matthews, Linardatos, Tell, & Rosenthal, 2008; Whitaker, 2015). The standardized mean difference (SMD) observed in recent meta-analyses is ~0.3 (Cipriani et al., 2018). On average, this translates to 2–3 points more on the Hamilton Depression Rating Scale (HAMD) than the reduction seen with placebo, a group difference that critics have argued is clinically insignificant (Moncrieff & Kirsch, 2015). Critics argue that trials (and thus meta-analyses) overestimate effectiveness due to a variety of biases, owing to unblinding, high dropout rates, selective patient recruitment, publication bias, etc. (Turner et al., 2008). They point at the heavy marketing and manipulation of research data by the pharmaceutical industry and the many conflicts of interest of ADM advocates (Gotschke, 2013; Whitaker, 2015).

Unfortunately, both advocates and critics selectively cite studies in support of their positions but hardly address and analyze the evidence cited by the other side, with some exceptions (Quitkin et al., 2000). A nice illustration is provided by the citations on efficacy and suicide risk provided by Nutt et al. (2014) and Gotzsche (2014) in their 2014 letters in the Lancet on benefits and harms of ADMs. Such discussions remain sterile and do not advance knowledge. They benefit neither patients nor providers.

Methods

Our editorial is not a systematic review or meta-analysis but rather a logical conceptual and methodological analysis. We address two important issues: why have antidepressants become so widely prescribed and how valuable are antidepressants in the treatment of depression. Both have multiple facets. It is virtually impossible (and largely irrelevant to what we set out to do) to address each facet in a full systematic way or by meta-analysis. A fully systematic approach would have required many literature searches, and would probably result in multiple papers. What we chose to do was to think rather than to count. We think that there is value in a logical analysis. While we have done our best to use systematic reviews and meta-analyses to inform our thinking as much as possible, what we set out to do was to describe where we thought the field currently is and where it ought to go. We have tried to be balanced in our selection of literature and to describe studies systematically.

Factors that keep fueling the dispute

Acute phase ADM trials: efficacy and limitations

Response rates (symptom reduction of 50% or more from baseline) in aADM arms exceed those in PLA arms, typically by 10–15%. The largest recent meta-analysis of aAD efficacy by Cipriani et al. included 522 trials, published and unpublished, comprising 1 16 477 participants and found that ADs had better response rates than placebo, with ORs ranging between 1.37 and 2.13 and an averaged SMD = 0.30, similar to NNT = 8 (Meadows et al., 2019), comparable with other recent meta-analyses (Cipriani et al., 2018). Although treatment settings were often unclear, most trials did not take place in primary care settings.

Cipriani et al.'s meta-analysis may even underestimate antidepressants efficacy as post-hoc analyses have found consistent superiority and dose-dependency for selective serotonin reuptake inhibitors (SSRIs) relative to PLA when using depressed mood as the major outcome (Hieronymus, Emilsson, Nilsson, & Eriksson, 2016a; Hieronymus, Nilsson, & Eriksson, 2016b) or the six prototypical depression items from the HAMD [depressed mood, guilt feeling, loss of interests, psychomotor retardation, psychic anxiety, and tiredness (Bech, 2010)]. The dose-dependency figures suggest that inclusion of trials using suboptimal doses in Cipriani et al.'s meta-analyses may have led to an underestimation of aADM efficacy. Likewise, the use of more narrow depression symptoms might have yielded higher efficacies, although caution is needed given the post-hoc nature of these findings. Because it is not fully clear whether findings from specialty settings can be generalized to primary care as only a minority of cases are referred to specialty settings, Linde et al. meta-analyzed 66 randomized controlled trials (RCTs) with 15.161 patients treated in primary care (Linde et al., 2015). ADs were found to be significantly superior to PLA, with estimated ORs between 1.69 and 2.03, rather similar to the Cipriano results. There were no significant differences between drug classes. Compared to general medicine drugs (e.g. ACE inhibitors and statins for prevention of cardiovascular events and stroke) ADM are not in general less efficacious (Leucht et al., 2012), although the clinical significance of a given effect size is context-dependent and may differ between psychiatry and general medicine.

Such data should settle the debate, at least for the acute phase treatment. However, a few issues still produce uncertainties. First, the authors caution that the 'certainty of evidence was moderate to very low' with '46 (9%) of 522 trials rated as high risk of bias, 380 (73%) as moderate, and 96 (18%) as low' (p. 1), leaving space for speculation about bias. Second, the analysis did not address full remission efficacy, only response, i.e. 50% or more improvement relative to baseline. That is, patients were better but not necessarily well. Third, the modest SMD, corresponding to a few HAMD points, has been argued to be of limited clinical significance (Jakobsen et al., 2017; Kirsch, 2014; Moncrieff & Kirsch, 2015; Turner et al., 2008). Fourth, it remains unclear how large the risk of breaking the blind is in SSRI trials because they have no 'active' placebo. Side effects of TCAs can be mimicked and a meta-analysis comparing TCA and active placebos, though limited by methodological problems found no significant difference (Moncrieff, 2003). These limitations keep providing fuel for critics.

Prolonged acute phase trials: 6–8 months outcomes

Some acute phase placebo-controlled trials have prolonged follow-up assessments of response and remission up to 8 months post-baseline randomization. Deshauer et al.'s meta-analysis included six studies that randomized a total of 1299 patients to either ADM or PLA and followed both groups for 6–8 months (Deshauer et al., 2008). Response rates averaged 61% in the ADM and 48% in the PLA arm, a difference of 13%; rates for remission, assessed in only four trials, were 45% and 38%, a difference of 7%, yielding a non-significant trend ($p < 0.1$). It is important to note that at least half of the studies (Davidson et al., 2002; Detke et al., 2004; Murray et al., 2005) included in Deshauer et al.'s meta-analysis excluded patients from follow-up

https://doi.org/10.1017/S0033291719003295 Published online by Cambridge University Press

if they did not respond to the acute phase blinded treatment, even if their short-term data were used in the intention-to-treat analysis. It is unclear whether similar exclusions occurred in the other three studies. Differences in treatment dropouts and study completers were not significant between the arms but many participants did not complete the final outcome assessment, ranging from 23% to 73% with mean 52% (Deshauer et al., 2008). The studies evidenced moderate risk of bias and most failed to report key methodologic issues, for instance on what happened to those who did not improve during the acute phase. Hence it is unclear how valid and generalizable the results of these prolonged aADM trials are. Compared to remission rates in the 8–12 week trials, Deshauer et al.'s findings suggest that the placebo group tends to 'catch up', possibly due to the natural history of depression.

Placebo effects and spontaneous remission; net and gross efficacy

Response rates (⩾50% improvement) in PLA arms of RCTs of depression are large, often between 35% and 40% in the PLA arm (Furukawa et al., 2016; Levkovitz, Tedeschini, & Papakostas, 2011; Rutherford & Roose, 2013). The response in PLA arms is due to multiple factors: (i) *spontaneous remission* referring to the phenomenon that some people get better in the absence of any treatment as part of the natural course of their depression; (ii) *regression to the mean*, in which individuals who are far from the mean at the baseline measurement tend to be closer to the mean at the next measurement, in as far as their extreme baseline scores were inflated by measurement error; and (iii) *non-specific treatment* effects, which consist of improvement due to being treated by a professional including explanation, attention, support, positive expectations, amongst others. It is possible that the placebo response in ADM trials is larger than with real-world patients because patients in regulatory trials are not fully representative of real-world patients (Rush et al., 2006; Van der Lem, van der Wee, Van Veen, & Zitman, 2012). The large placebo effect indicates that many respond not because of the specific ADM treatment, but because of other effects. This makes it difficult for trials to establish the excess value of specific treatments and provides critics with ammunition to minimize the benefits of ADMs by focusing on net efficacy and advocates to maximize benefits by focusing on gross efficacy.

Continuation and maintenance ADM trials: efficacy and limitations

The most recent meta-analysis evaluated 72 cADM and 37 mADM trials with 14.450 and 7253 participants, respectively (Sim, Lau, Sim, Sum, & Baldessarini, 2016). In the 72 cADM trials, lasting 33 *weeks* on average, the pooled advantage of cADM over discontinuation to placebo in terms of the relative response rate (RR) was 1.90, CI 1.73–2.08, NNT 4.4; in the 37 mADM trials lasting 27 *months* on average, it was 2.03, CI 1.80–2.28; NNT = 3.8; with minor differences among drug types (Sim et al., 2016). The authors did not distinguish between the two designs that have been used to investigate the efficacy of continuation and maintenance treatment with ADMs (cmADMs): the discontinuation and the extension design (Fig. 1).

In the *discontinuation* (or placebo-substitution) design, patients who respond to aADM are randomized to cmADM *v.* PLA. Meta-analyses document substantially higher relapse/recurrence rates during the post-randomization 6–12 months in placebo-substitution arms (~42%) compared to the cmAD arms (~22%) (Geddes et al., 2003; Glue, Donovan, Kolluri, & Emir, 2010; Hansen et al., 2008; Kaymaz, van Os, Loonen, & Nolen, 2008). The major problem of the discontinuation design, however, is that AD withdrawal symptoms may masquerade as relapse/recurrence (Fava, Gatti, Belaise, Guidi, & Offidani, 2015; Greenhouse & Meyer, 1991). Its bias potential is unclear because of uncertainty about the risk period of withdrawal phenomena (Borges et al., 2014; El-Mallakh, Waltrip, & Peters, 1999; Kaymaz et al., 2008). Excess relapse/recurrence rates in the placebo-substitution arms compared to the cmADM arms are strongest in the first month after discontinuation and gradually drop to approximately zero in the next 3–6 months (Borges et al., 2014; El-Mallakh et al., 1999; Kaymaz et al., 2008).

Extension trials start as regular double-blind, placebo-controlled aADM trials that continue blinded treatment in responders to ADMs (~45%) and responders to PLA (~32%) for another 5–12 months (Zimmerman, Posternak, & Ruggero, 2007). In five double-blind extension trials including 901 participants and lasting on average 35 weeks, relapse/recurrence rates in those who initially remitted averaged 25% in the PLA arm *v.* 8% in the ADM arm, showing a strong benefit for cADM relative to cPLA (RR = 2.0) (Zimmerman et al., 2007). The extension design lacks the withdrawal issues of discontinuation trials, but still has the limitation that acute phase non-responders are not included and followed. This is a serious limitation as about half do not remit on acute phase treatment (Khin, Chen, Yang, Yang, & Laughren, 2011; Levkovitz et al., 2011; Pigott, Leventhal, Alter, & Boren, 2010; Rush et al., 2006). Treatment completion rates were often <50%.

In conclusion, compared to discontinuation trials, extension trials provide stronger evidence that cmADM in responders to aADM treatment reduce relapse/recurrence risk relative to cmPLA in aPLA responders because aADM responders were not switched to PLA, excluding confounding effects of ADM withdrawal. However, extension trials are relatively rare and patients on cmPLA, although they initially responded, may slowly realize that they are using placebos, with possibly negative psychological consequences.

Naturalistic long-term outcome studies

Many naturalistic outcome studies (i.e. non-randomized observational cohort or longitudinal studies) have charted the long-term course and outcome of depression (⩾2 years). Reviews typically report worse long-term outcomes for depressed patients who are managed in specialty settings compared to those managed in primary care (e.g. Brodaty, Luscombe, Peisah, Anstey, & Andrews, 2001; Hardeveld, Spijker, De Graaf, Nolen, & Beekman, 2010; Piccinelli & Wilkinson, 1994; Steinert, Hofmann, Kruse, & Leichsenring, 2014; van Weel-Baumgarten, Schers, van den Bosch, van den Hoogen, & Zitman, 2000). Piccinelli and Wilkinson, reviewing 50 outcome studies, reported weighted averages of 43% sustained recovery and 15% persistent depression at 1-year follow-up, and 24% sustained recovery and 12% persistent depression at >10-year follow-up. Other reviews report recurrence rates of up to 85% in 15 years for depression treated in specialty settings (Hardeveld et al., 2010; Mulder & Frampton, 2014). In contrast, two major reviews of adult depression in *general practice and the community* (20 studies of almost 6000 participants, most with 3–7 years follow-ups) found that although 10–17% had a chronic course, up to 85% recovered for some time and that 35–60% experienced

https://doi.org/10.1017/S0033291719003295 Published online by Cambridge University Press

https://doi.org/10.1017/S0033291719003295 Published online by Cambridge University Press

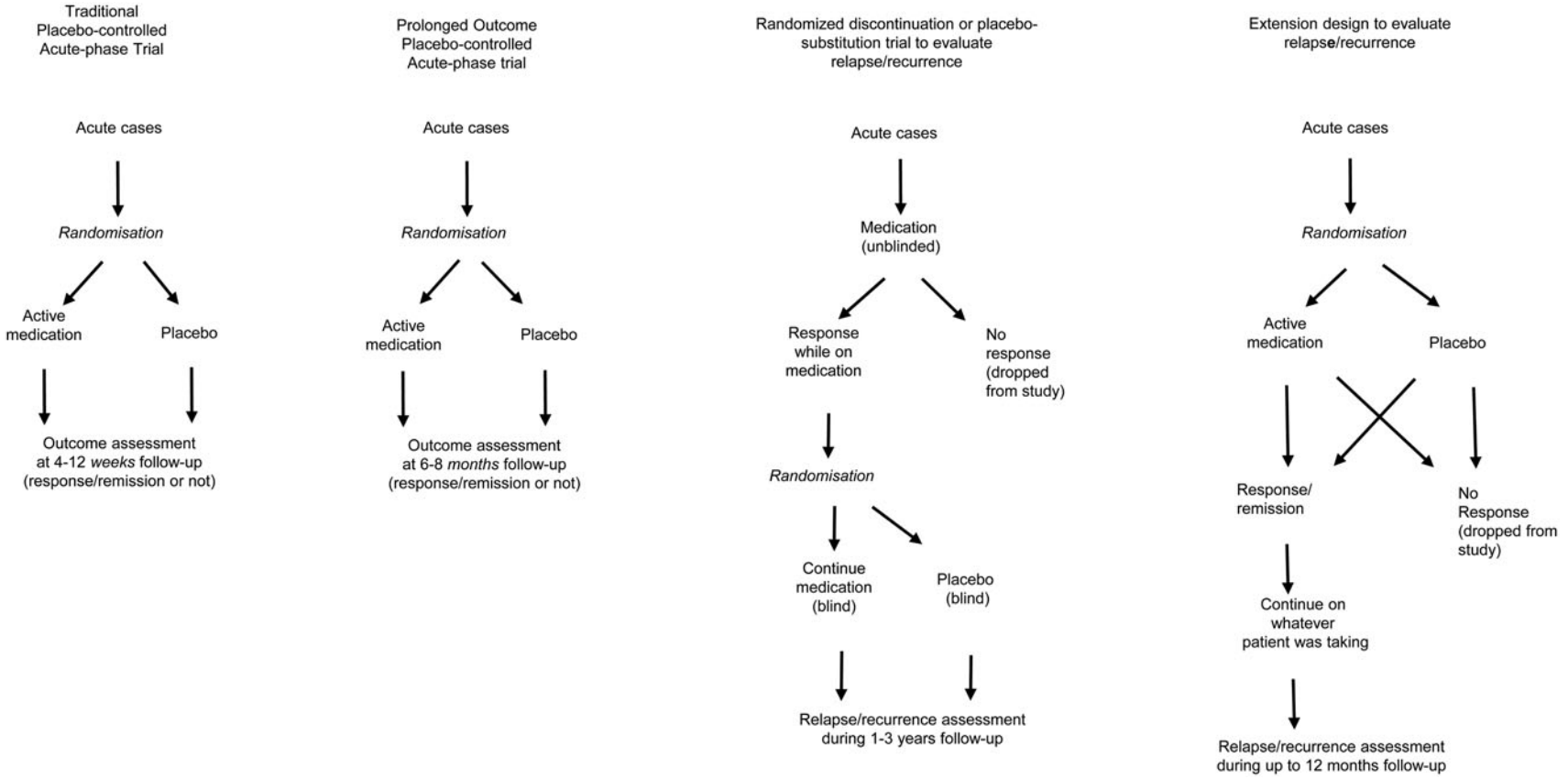


Fig. 1. Illustration of designs.

stable recovery (Steinert et al., 2014; van Weel-Baumgarten et al., 2000). Regarding outcome of non-treated community and general practice cases, multiple studies suggest that many depressive episodes (80–50%) are self-limiting as they remit within 3–12 months (Goldberg, Privett, Ustun, Simon, & Linden, 1998; Regier et al., 1998; Sareen et al., 2013; Spijker et al., 2002; Wang et al., 2017; Whiteford et al., 2013). For instance, Whiteford's (2013) meta-analysis of 19 studies of untreated depression in general practice settings estimated that 23% remit within 3 months, 32% within 6 months, and 53% within 12 months.

Observational studies typically report similar or worse outcomes for treated *v.* untreated depressive episodes in primary care and community samples (Boerema et al., 2017; Goldberg et al., 1998; Hughes & Cohen, 2009; Wang et al., 2017). Hughes and Cohen's review compared AD-treated cases in 12 naturalistic cohorts of treated patients ($n = 3901$ at final follow-up) and three non-ADM-treated samples of 1160 patients. Most participants were white females with one inpatient stay. Frequency, duration, and severity of episodes varied substantially. Outcomes were unrelated to treatment status, and not better in ADM-treated cohorts relative to non-ADM-treated samples. Heterogeneity of study designs and outcome definitions were large and hampered statistical analysis. Recent naturalistic long-term outcome studies comparing ADM-treated and untreated individuals with diagnosed depression consistently corroborate the findings of Hughes and Cohen's review (Bockting, Hollon, Jarrett, Kuyken, & Dobson, 2015; Hengartner, Angst, & Roessler, 2018; Nuijen, ten Have, Tuithof, van Dorsselaer, & van Bon-Martens, 2014; Verduijn et al., 2017; Vittengl, 2017; Wang et al., 2017). Statistical adjustment for confounders (e.g. educational level, marital status, baseline severity, family history) did not attenuate the association of ADM treatment with outcome (Hengartner et al., 2018; Vittengl, 2017). However, statistical adjustment cannot adjust for unmeasured differences between ADM *v.* non-ADM (known as confounding by indication) and it is very likely that untreated cases with established major depressive disorder are less vulnerable and more resilient than treated cases (e.g. Goldberg et al., 1998; Ormel, Oldehinkel, Brilman, & van den Brink, 1993).

Conclusion naturalistic studies

We want to emphasize two points. First, RCTs are the gold standard of treatment evaluation and naturalistic outcome studies can never prove treatment (non-) effectiveness relative to no treatment. However, because naturalistic outcome studies typically do not find an association between ADM-treatment and long-term outcome, their findings offer ammunition to critics to express doubts regarding long-term ADM effects, in particular in combination with (i) the limitations of and (ii) the weaker evidence of prolonged aADM trials. Second, naturalistic long-term outcome studies typically follow prevalent cases without distinguishing between first-ever onset cases *v.* recurrent cases. As the latter are overrepresented in patient series and population samples, they may not reflect the prognosis of first-ever onsets, which is reasonably positive, with 50–60% achieving stable recovery, 35–40% experiencing at least one recurrence in the next 15 years, and 15% becoming chronic (Eaton et al., 2008; Mattisson, Bogren, Horstmann, Munk-Jorgensen, & Nettelbladt, 2007).

Efficacy and limitations of psychological treatments

Critics of ADMs often advocate psychological treatments, especially cognitive behavioral therapy (CBT), interpersonal therapy (IPT), mindfulness-based cognitive therapy (MBCT), and problem-solving therapy (PST). But how efficacious are these treatments and do they offer a serious alternative to cmADM or are they vulnerable to similar criticisms as the ADM? Compared to PLA, care-as-usual (CAU), and waiting-list controls, meta-analyses indicate that *short-term efficacy* of psychological treatments is modest in magnitude and comparable to aADM (e.g. Barth et al., 2013; Cuijpers et al., 2014; Cuijpers, Cristea, Karyotaki, Reijnders, & Huibers, 2016). For instance, Barth's meta-analysis identified 198 studies, including 15 118 adult patients with depression. All major types of psychotherapeutic interventions were superior to waitlist control condition with moderate-to-large effects (range $d = 0.62–0.92$). Studies with larger samples and higher quality (blinded observers, self-report measures, adequate concealed allocation, etc.) reported substantially smaller effects of about $d = 0.30$. Cuijpers et al.'s meta-analysis of CBT for MDD, the most studied psychological treatment, including 63 studies and approximately 4000 participants, reported a pooled effect size of 0.75 (95% CI 0.64–0.87; NNT = 3.9). High-quality studies using CAU as a control had significantly smaller effect sizes (0.43, NNT 7.3) and so did trials using PLA as the control (0.55, NNT = 5.5). Direct comparisons between ADM and major psychological treatments indicate that they are about equally efficacious in the short-term treatment of major depression (Cuijpers et al., 2013a, 2013b), although an earlier meta-analysis found a slight excess benefit for SSRIs (Cuijpers, van Straten, van Oppen, & Andersson, 2008). Dropout rates are typically smaller in psychological treatments (Cuijpers et al., 2008). Similar to ADM trials, accumulating biases have also resulted in overestimation of the short-term efficacy of psychological treatments'(Cuijpers et al., 2016; Driessen, Hollon, Bockting, Cuijpers, & Turner, 2015).

Three meta-analyses examined relapse/recurrence in responders to psychological treatments (largely CBT, BA, and IPT) (Cuijpers et al., 2013a; Sim et al., 2016; Vittengl, Clark, Dunn, & Jarrett, 2007). Vittengl et al.'s meta-analysis of 28 studies including 1880 adults found substantial levels of symptom return after discontinuation of aCBT in CBT responders (29% within 1 year and 54% within 2 years). These rates appeared comparable to those associated with other depression-specific psychotherapies but lower than those associated with aADM. Seven trials compared aCBT and aADM. Averaging separately by treatment type yielded relapse/recurrence rates of 39% for aCBT and 61% for aADM over an average of 68 weeks.

Cuijpers et al.'s (2013b) meta-analysis examined the long-term (6–18 months) effects of aCBT *v.* aADM with and without cmADM (nine studies, 506 patients) (Cuijpers et al., 2013b). Short-term outcomes were comparable although dropout was lower in CBT. Patients who received acute phase CBT were significantly less likely to relapse after discontinuation than patients who were withdrawn from ADM (OR = 2.61, 95% CI 1.58–4.31, NNT = 5). There was a non-significant trend favoring prior CBT over acADM (five studies) ($p < 0.1$; OR = 1.62, 95% CI 0.97–2.72; NNT = 10) ($p < 0.1$; OR = 1.62, 95% CI 0.97–2.72; NNT = 10); suggesting that aCBT might be even more efficacious than continuation ADM in preventing relapse. Finally, Sim et al.'s meta-analysis included 22 psychosocial treatment trials [CBT, IPT, MBCT, psychoeducation that followed 1969 (mostly) remitted patients with recurrent depression across 24 months (Sim et al., 2016)]. Treatments were slightly more effective than controls in preventing relapse/recurrence: pooled RR = 1.39 (1.13–1.70); substantially less effective than in Vittengl et al.'s meta-analysis.

https://doi.org/10.1017/S0033291719003295 Published online by Cambridge University Press

Some trials have recently examined whether *preventive variants* of CBT/MBCT have better long-term outcomes than cmADM. These trials randomized aADM responders into two or three arms: continuing ADM *v.* preventive CBT/MBCT with or without continuing ADM (Bockting et al., 2018; Huijbers et al., 2016; Kuyken et al., 2015; Segal et al., 2010). A recent meta-analysis of four preventive MBCT trials (637 participants, 266 relapses during the 60-week follow-up period) found a statistically significant advantage for MBCT compared to cmADM, suggesting that MBCT appears efficacious as a treatment for relapse prevention for those with recurrent depression in (partial) remission (Kuyken et al., 2016). Additionally, in a recent trial not included in that review, cmADM was not superior to preventive CT administered while tapering off ADM [relapse/recurrence risks over 15–24 months (60% *v.* 63%) (Bockting et al., 2018)]. However, tapering was often unsuccessful as many patients continued or resumed cmADM, which may have decreased symptom return in the taper PCT/MBCT arm. On the other hand, misclassified withdrawal symptoms may have increased symptom return in this arm.

Conclusions

Although these findings suggest that acute phase and preventive psychotherapy represent viable alternatives to ADM treatment, the evidence is still somewhat uncertain due to the relatively small number of high-quality trials and the inherent problem of blinding in psychotherapy evaluation. Currently, the evidence base of cmADM in acute phase ADM responders is substantially larger than that for recurrence-prevention using CBT/MBCT, which may reflect registration requirements and funding opportunities. Similar to ADM trials, accumulating biases have also resulted in overestimation of the short- and long-term efficacy of psychological treatments (Cuijpers et al., 2016; Driessen et al., 2015). Effective blinding remains a problem and waitlist controls are often used even though this is known to inflate contrasts (Cuijpers & Cristea, 2015). Furthermore, in all trials comparing psychological treatments with discontinued ADM, misclassified withdrawal symptoms in the ADM arm may have inflated the difference. Hence, there is a clear need for replication by high-quality head-to-head trials.

How to resolve the controversy

New data are urgently needed to push the (partly political) antidepressant debate toward a balanced and evidence-based discussion of benefits and harms of treatment modalities. We recommend four approaches: (1) placebo-controlled RCTs with relevant long-term outcomes, (2) analysis of non-randomized treatment outcome data using instrumental variable analysis and propensity score analysis, (3) patient cohort studies including effect moderators to enhance personalized treatment, and (4) psychological interventions as a universal first-line treatment step. The first two approaches will inform regarding treatment effects on long-term outcome whereas the latter two will reduce the relevance of the controversy. Negotiations between advocates and critics of AD to achieve consensus on the exact data needed to resolve the controversy should precede concrete research steps.

Placebo-controlled trials with relevant long-term outcomes

Logistically extremely difficult but in our opinion highly desirable are randomized placebo-controlled long-term outcome studies of ADM and psychotherapy. Long-term means at least 1 year follow-up, preferably 2 years. If sufficiently powered, an additional arm of combined treatment (COM) could be added. An example of a top priority trial is randomization of currently ADM-free depressed individuals, stratified by history of ADM treatment (never *v.* ever), to ADM *v.* CBT *v.* PLA. Acute phase non-responders, irrespective of treatment, will be offered the treatment they prefer, but kept in the study and administered the very same follow-up assessments, thereby addressing the major deficiency of previous trials. Alternatively, non-responders could be offered the other treatment modality in addition to the allotted treatment (COM) and provide the same for the PLA non-responders. Primary outcomes should be psychopathology, role function, adverse events, and acceptability. Machine learning can be used to generate selection algorithms that identify the optimal intervention for each individual who does not remit on COM and use those algorithms in the patient by mechanism interactions to improve the power of the tests of mediation (Cohen & DeRubeis, 2018). If putative effect moderators are included at baseline, much can be learned about interactions between treatment modality and effect moderators. Finally, it is important to perform cost-effectiveness analyses as well because what drives shifts in practice are often shifts in reimbursement and what drives shifts in reimbursement is evidence of cost-effectiveness. As an anonymous program officer at NIMH once said: 'You cannot herd cats but you can move their food.'

Multiple problems threaten the feasibility and validity of the proposed study: dropouts, protocol violations, imbalance between the arms at the entry of the continuation phase, complex logistics, and ethical concerns. Since 2000, it has been difficult to keep patients on PLA for more than a few months because it is considered unethical to withhold effective treatment. However, the modest excess efficacy of ADM and CBT relative to PLA implies that the majority of patients who respond to specific treatments do so for other reasons; that is, due to non-specific treatment effects and spontaneous remission, and not a 'specific' ADM or CBT effect. Furthermore, ethical concerns are manageable using patient-level stopping rules. The modest net efficacy of ADM or CBT (specificity) reduces the threat of imbalance as well, as many will respond to PLA, and balance can be managed with propensity analyses. Logistical problems necessitate the cooperation of many primary care providers from large catchment areas.

Inventive analyses of observational databases

The unsurpassed advantage of randomization is that the resulting groups are comparable (apart from differences due to chance). This allows the unambiguous attribution of differences in outcomes between groups to differences in treatment. Observational, non-randomized research can never rule out that the groups differ in characteristics (confounders) that affect both treatment and outcome. Notorious 'confounders' are severity and history. However, statisticians have recently developed methods that allow, under certain conditions and with reasonable certainty, causal statements about treatment effects based on observational data, even if not all potential 'confounders' are measured. The methods are instrumental variable analysis (Martens, Pestman, de Boer, Belitser, & Klungel, 2006) and propensity score analysis (Austin, 2011). Potentially relevant observational databases can be generated from routine outcome measurement (ROM), patient cohort data, and longitudinal population-based studies, provided they assess long-term outcome.

https://doi.org/10.1017/S0033291719003295 Published online by Cambridge University Press

Cohort studies including effect moderators

Predicting who will benefit from which treatment is difficult. As a result, personalized treatment ('precision medicine') is not yet possible with depression although there are interesting developments in this direction (Cohen & DeRubeis, 2018; DeRubeis et al., 2014; Fournier et al., 2009; Kessler, 2018). Cohort studies of patients with the same diagnosis can provide insight if putative effect moderators are included at baseline. If treatments become more precise and personalized, the relevance of the ADM controversy will decline. Relatively easy to measure potential effect moderators are clinical characteristics, personality, biomarkers to be determined from bodily fluids including DNA, social network, and support characteristics.

Psychological help as a first-line treatment

An approach that may also reduce the relevance of the controversy as well is to offer short-term psychological treatment as the first-line treatment step for all patients with depression unless they refuse or cannot receive psychological treatment or have characteristics shown to predict a better response to ADM. ADM then remains reserved for patients who do not want, get, or respond to this first step. It is of great importance which of the non-responders to psychological treatment improves on ADM and whether psychological treatment can subsequently contribute to the phasing out of their ADM use. Severely ill patients could be started on combination of ADM and psychotherapy. Some guidelines actually do advise psychological treatment for adolescents with depression as the first-line treatment step, and other guidelines, such as the UK NICE guideline for depression and the Dutch multidisciplinary guideline for depression, also recommend psychological treatment as the default first step for adults with mild-to-moderate depression(Clark, 2018; NICE, 2011; Spijker et al., 2019).

Concluding comments

The ADM debate lingers because of mixed messages and differences between ADM critics and advocates in their interpretations of the research findings, the weaknesses and strengths of the various studies, and the lack of high-quality RCTs with long-term follow-ups; not a single large trial investigated head-on the long-term outcomes of acute phase CBT *v.* ADM *v.* PLA. The problems of generalizability to real-world patients, unblinding, dropout, withdrawal symptoms, exclusion of acute phase non-responders, selective reporting, and publication bias make it easy for critics and advocates to minimize the relevance of study findings that do not fit their interests and beliefs and emphasize the relevance of findings that do fit. It is worth noting that these problems affect both psychotherapy and ADM. What also plays a role is that we do not fully understand what depression is, or what the precise mechanism of action of ADM is. The combination of multiple theories with little substrate is not helping to resolve the controversy.

Overall, relatively strong evidence (in terms of low bias risk) in favor of ADM relative to PLA comes from the 'classical' trials with relatively long follow-ups (6–8 months) and the extension studies targeting relapse in responders to acute phase ADM *v.* PLA. However, the extension studies are limited to acute phase responders and the 'classical' trials are few in number and show statistically significant associations only for response and not remission. This suggests that long-term follow-up effects of ADM may be smaller than 8-week post-treatment effects; control groups tend to 'catch up' with the treatment group across time.

It is likely to prove especially important to conduct cost-effectiveness analyses of these studies. With the advent of the SSRIs, the pendulum has swung toward a wholesale reliance on ADM with the vast majority of the prescriptions now written in general practice. The UK has bucked this trend by bypassing the providers and going straight to the funders in the government with relevant cost-effectiveness data (Clark, 2018). The National Health Service invested £700 million pounds to train psychotherapists to do interventions judged to be efficacious in clinical practice guidelines developed by the National Institute for Health Care Evaluation (NICE). Patient outcomes are monitored at every session and anonymized aggregate data posted on a publically accessible website. That recovery rates have risen from 35% to over 50% across the last decade speaks to the power of shifting reimbursement to the most cost-efficient interventions and monitoring outcomes in a wholly transparent fashion speaks to the utility of this English Improving Access to Psychological Therapies (IAPT) program (Clark, 2018). Although the overall outcomes achieved by IAPT until 2017 are encouraging, it is unclear to what extent gains are maintained as systematic long-term FU are lacking. Hence, it is important to wait for the final evaluation of the IAPT program.

Despite countless studies on the short-term effectiveness of ADMs and hundreds of billions of revenue, the ADM debate continues. The lack of long-term outcome information hampers a balanced consideration. RCTs comparing the long-term outcome of ADM *v.* PLA and psychological treatments can provide the necessary data. Research consortia with researchers from varied fields and different allegiances are best situated to perform such 'mega-trials' so as to guarantee balanced composition (Quitkin et al., 2000) in accordance with the principle of 'adversarial collaboration' (Mellers, Hertwig, & Kahneman, 2001). Given the substantial prevalence of depression and cmADM treatment and the availability of alternative treatments, it is crucial that the lacking data become available.

Author contributions. Jo conceived and managed the study, collected and interpreted the literature, developed the argument, and wrote various drafts of the report. SH added important material. All co-authors contributed to the interpretation of relevant literature, provided critical feedback on earlier drafts, and approved the final version of the manuscript.

Conflict of interest. All authors declare to have no competing interests. JO is an epidemiologist and sociologist with a record of accomplishment in psychiatric epidemiology, and has been involved in a range of studies including a few AMD and psychotherapy trials. PS is a clinical psychologist who has been strongly involved in psychotherapy research. CLHB and SDH are clinical psychologists involved in both AMD and psychotherapy trials. YAdeV is an epidemiologist, not involved in trials, and GJS and AOJC are psychologists who have not been involved in randomized trials of ADM or psychotherapy.

References

Antonuccio, D., Danton, W., DeNelsky, G., Greenberg, R., & Gordon, J. (1999). Raising questions about antidepressants. *Psychotherapy and Psychosomatics*, *68*(1), 3–14.

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research*, *46*(3), 399–424.

Barth, J., Munder, T., Gerger, H., Nueesch, E., Trelle, S., Znoj, H., *et al.* (2013). Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. *PLoS Medicine*, *10*(5), e1001454.

https://doi.org/10.1017/S0033291719003295 Published online by Cambridge University Press

Bech, P. (2010). Is the antidepressive effect of second-generation antidepressants a myth? *Psychological Medicine*, *40*(2), 181–186.

Bockting, C. L., Hollon, S. D., Jarrett, R. B., Kuyken, W., & Dobson, K. (2015). A lifetime approach to major depressive disorder: The contributions of psychological interventions in preventing relapse and recurrence. *Clinical Psychology Review*, *41*, 16–26.

Bockting, C. L. H., Klein, N. S., Elgersma, H. J., van Rijsbergen, G. D., Slofstra, C., Ormel, J., *et al.* (2018). Effectiveness of preventive cognitive therapy while tapering antidepressants versus maintenance antidepressant treatment versus their combination in prevention of depressive relapse or recurrence (DRD study): A three-group, multicentre, randomised controlled trial. *The Lancet. Psychiatry*, *5*(5), 401–410.

Boerema, A. M., ten Have, M., Kleiboer, A., de Graaf, R., Nuyen, J., Cuijpers, P., & Beekman, A. T. F. (2017). Demographic and need factors of early, delayed and no mental health care use in major depression: A prospective study. *BMC Psychiatry*, *17*, 367.

Borges, S., Chen, Y., Laughren, T. P., Temple, R., Patel, H. D., David, P. A., … Khin, N. A. (2014). Review of maintenance trials for major depressive disorder: A 25-year perspective from the US food and drug administration. *Journal of Clinical Psychiatry*, *75*(3), 205–214.

Bridge, J., Barbe, R., Birmaher, B., Kolko, D., & Brent, D. (2005). Emergent suicidality in a clinical psychotherapy trial for adolescent depression. *American Journal of Psychiatry*, *162*(11), 2173–21U2.

Brodaty, H., Luscombe, G., Peisah, C., Anstey, K., & Andrews, G. (2001). A 25-year longitudinal, comparison study of the outcome of depression. *Psychological Medicine*, *31*(8), 1347–1359.

Cipriani, A., Furukawa, T. A., Salanti, G., Chaimani, A., Atkinson, L. Z., Ogawa, Y., … Geddes, J. R. (2018). Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: A systematic review and network meta-analysis. *Lancet*, *391*(10128), 1357–1366.

Clark, D. M. (2018). Realizing the mass public benefit of evidence-based psychological therapies: The IAPT program. *Annual Review of Clinical Psychology*, *14*, 159–183.

Cohen, Z. D., & DeRubeis, R. J. (2018). Treatment selection in depression. *Annual Review of Clinical Psychology*, *14*, 209–236.

Cooney, G. M., Dwan, K., Greig, C. A., Lawlor, D. A., Rimer, J., Waugh, F. R., *et al.* (2013). Exercise for depression. *Cochrane Database of Systematic Reviews* (9), CD004366.

Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., & Dobson, K. S. (2013a). A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie*, *58*(7), 376–385.

Cuijpers, P., & Cristea, I. A. (2015). What if a placebo effect explained all the activity of depression treatments? *World Psychiatry*, *14*(3), 310–311.

Cuijpers, P., Cristea, I. A., Karyotaki, E., Reijnders, M., & Huibers, M. J. H. (2016). How effective are cognitive behavior therapies for major depression and anxiety disorders? A meta-analytic update of the evidence. *World Psychiatry*, *15*(3), 245–258.

Cuijpers, P., Hollon, S. D., van Straten, A., Bockting, C., Berking, M., & Andersson, G. (2013b). Does cognitive behaviour therapy have an enduring effect that is superior to keeping patients on continuation pharmacotherapy? A meta-analysis. *BMJ Open*, *3*(4), e002542.

Cuijpers, P., Karyotaki, E., Weitz, E., Andersson, G., Hollon, S. D., & van Straten, A. (2014). The effects of psychotherapies for major depression in adults on remission, recovery and improvement: A meta-analysis. *Journal of Affective Disorders*, *159*, 118–126.

Cuijpers, P., van Straten, A., van Oppen, P., & Andersson, G. (2008). Are psychological and pharmacologic interventions equally effective in the treatment of adult depressive disorders? A meta-analysis of comparative studies. *Journal of Clinical Psychiatry*, *69*(11), 1675–1685.

Davidson, J., Gadde, K., Fairbank, J., Krishnan, R., Califf, R., Binanay, C., *et al.* (2002). Effect of *Hypericum perforatum* (St John's wort) in major depressive disorder – a randomized controlled trial. *JAMA: Journal of the American Medical Association*, *287*(14), 1807–1814.

DeRubeis, R. J., Cohen, Z. D., Forand, N. R., Fournier, J. C., Gelfand, L. A., & Lorenzo-Luaces, L. (2014). The personalized advantage index: Translating research on prediction into individualized treatment recommendations. A demonstration. *PLoS ONE*, *9*(1), e83875.

Deshauer, D., Moher, D., Fergusson, D., Moher, E., Sampson, M., & Grimshaw, J. (2008). Selective serotonin reuptake inhibitors for unipolar depression: a systematic review of classic long-term randomized controlled trials. *Canadian Medical Association Journal*, *178*(10), 1293–1301.

Detke, M., Wiltse, C., Mallinckrodt, C., McNamara, R., Demitrack, M., & Bitter, I. (2004). Duloxetine in the acute and long-term treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. *European Neuropsychopharmacology*, *14*(6), 457–470.

Driessen, E., Hollon, S. D., Bockting, C. L. H., Cuijpers, P., & Turner, E. H. (2015). Does publication bias inflate the apparent efficacy of psychological treatment for major depressive disorder? A systematic review and meta-analysis of US national institutes of health-funded trials. *PLoS ONE*, *10*(9), e0137864.

Eaton, W. W., Shao, H., Nestadt, G., Lee, B. H., Bienvenu, O. J., & Zandi, P. (2008). Population-based study of first onset and chronicity in major depressive disorder. *Archives of General Psychiatry*, *65*(5), 513–520.

El-Mallakh, R. S., Waltrip, C., & Peters, C. (1999). Can long-term antidepressant use be depressogenic? *Journal of Clinical Psychiatry*, *60*(4), 263–263.

Fava, G. A. (2003). Can long-term treatment with antidepressant drugs worsen the course of depression? *Journal of Clinical Psychiatry*, *64*, 123–133.

Fava, G. A., Gatti, A., Belaise, C., Guidi, J., & Offidani, E. (2015). Withdrawal symptoms after selective serotonin reuptake inhibitor discontinuation: A systematic review. *Psychotherapy and Psychosomatics*, *84*(2), 72–81.

Fournier, J. C., DeRubeis, R. J., Shelton, R. C., Hollon, S. D., Amsterdam, J. D., & Gallop, R. (2009). Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression. *Journal of Consulting and Clinical Psychology*, *77*(4), 775–787.

Furukawa, T. A., Cipriani, A., Atkinson, L. Z., Leucht, S., Ogawa, Y., Takeshima, N., … Salanti, G. (2016). Placebo response rates in antidepressant trials: A systematic review of published and unpublished double-blind randomised controlled studies. *The Lancet. Psychiatry*, *3*(11), 1059–1066.

Geddes, J. R., Carney, S. M., Davies, C., Furukawa, T. A., Kupfer, D. J., Frank, E., & Goodwin, G. M. (2003). Relapse prevention with antidepressant drug treatment in depressive disorders: A systematic review. *Lancet*, *361*(9358), 653–661.

Geddes, J. R., & Cipriani, A. (2015). Time to abandon placebo control in pivotal phase III trials? *World Psychiatry*, *14*(3), 306–307.

Glue, P., Donovan, M. R., Kolluri, S., & Emir, B. (2010). Meta-analysis of relapse prevention antidepressant trials in depressive disorders. *Australian and New Zealand Journal of Psychiatry*, *44*(8), 697–705.

Goldberg, D., Privett, M., Ustun, B., Simon, G., & Linden, M. (1998). The effects of detection and treatment on the outcome of major depression in primary care: a naturalistic study in 15 cities. *British Journal of General Practice*, *48*(437), 1840–1844.

Gotschke, P. (2013). *Deadly medicines and organised crime*. London: Radcliffe Publishing.

Gotzsche, P. C. (2014). Why I think antidepressants cause more harm than good. *The Lancet. Psychiatry*, *1*(2), 104–106.

Greenhouse, J., & Meyer, M. (1991). A note on randomization and selection bias in maintenance therapy clinical-trials. *Psychopharmacology Bulletin*, *27*(3), 225–229.

Hansen, R., Gaynes, B., Thieda, P., Gartlehner, G., Deveaugh-Geiss, A., Krebs, E., & Lohr, K. (2008). Meta-analysis of major depressive disorder relapse and recurrence with second-generation antidepressants. *Psychiatric Services*, *59*(10), 1121–1130.

Hardeveld, F., Spijker, J., De Graaf, R., Nolen, W. A., & Beekman, A. T. F. (2010). Prevalence and predictors of recurrence of major depressive disorder in the adult population. *Acta Psychiatrica Scandinavica*, *122*(3), 184–191.

Hengartner, M. P., Angst, J., & Roessler, W. (2018). Antidepressant use prospectively relates to a poorer long-term outcome of depression: Results from a prospective community cohort study over 30 years. *Psychotherapy and Psychosomatics*, *87*(3), 181–183.

Hieronymus, F., Emilsson, J. F., Nilsson, S., & Eriksson, E. (2016a). Consistent superiority of selective serotonin reuptake inhibitors over placebo in reducing depressed mood in patients with major depression. *Molecular Psychiatry*, *21*(4), 523–530.

https://doi.org/10.1017/S0033291719003295 Published online by Cambridge University Press

Hieronymus, F., Nilsson, S., & Eriksson, E. (2016b). A mega-analysis of fixed-dose trials reveals dose-dependency and a rapid onset of action for the antidepressant effect of three selective serotonin reuptake inhibitors. *Translational Psychiatry*, *6*, e834.

Hughes, S., & Cohen, D. (2009). A systematic review of long-term studies of drug treated and non-drug treated depression. *Journal of Affective Disorders*, *118*(1–3), 9–18.

Huijbers, M. J., Spinhoven, P., Spijker, J., Ruhe, H. G., van Schaik, D. J. F., van Oppen, P., *et al.* (2016). Discontinuation of antidepressant medication after mindfulness-based cognitive therapy for recurrent depression: randomised controlled non-inferiority trial. *British Journal of Psychiatry*, *208*(4), 366, -+.

Jakobsen, J. C., Katakam, K. K., Schou, A., Hellmuth, S. G., Stallknecht, S. E., Leth-Moller, K., *et al.* (2017). Selective serotonin reuptake inhibitors versus placebo in patients with major depressive disorder. A systematic review with meta-analysis and Trial Sequential Analysis. *BMC Psychiatry*, *17*, 58.

Kaymaz, N., van Os, J., Loonen, A. J. M., & Nolen, W. A. (2008). Evidence that patients with single versus recurrent depressive episodes are differentially sensitive to treatment discontinuation: A meta-analysis of placebo-controlled randomized trials. *Journal of Clinical Psychiatry*, *69*(9), 1423, -+.

Kessler, R. C. (2018). The potential of predictive analytics to provide clinical decision support in depression treatment planning. *Current Opinion in Psychiatry*, *31*(1), 32–39.

Khin, N. A., Chen, Y., Yang, Y., Yang, P., & Laughren, T. P. (2011). Exploratory analyses of efficacy data from major depressive disorder trials submitted to the US food and drug administration in support of new drug applications. *Journal of Clinical Psychiatry*, *72*(4), 464–472.

Kirsch, I. (2014). Antidepressants and the placebo effect. *Zeitschrift Fur Psychologie – Journal of Psychology*, *222*(3), 128–134.

Kirsch, I., Deacon, B. J., Huedo-Medina, T. B., Scoboria, A., Moore, T. J., & Johnson, B. T. (2008). Initial severity and antidepressant benefits: a meta-analysis of data submitted to the food and drug administration. *PLoS Medicine*, *5*(2), 260–268.

Kuyken, W., Hayes, R., Barrett, B., Byng, R., Dalgleish, T., Kessler, D., … Byford, S. (2015). Effectiveness and cost-effectiveness of mindfulness-based cognitive therapy compared with maintenance antidepressant treatment in the prevention of depressive relapse or recurrence (PREVENT): a randomised controlled trial. *Lancet*, *386*(9988), 63–73.

Kuyken, W., Warren, F. C., Taylor, R. S., Whalley, B., Crane, C., Bondolfi, G., *et al.* (2016). Efficacy of mindfulness-based cognitive therapy in prevention of depressive relapse an individual patient data meta-analysis from randomized trials. *JAMA Psychiatry*, *73*(6), 565–574.

Leucht, S., Hierl, S., Kissling, W., Dold, M., & Davis, J. M. (2012). Putting the efficacy of psychiatric and general medicine medication into perspective: review of meta-analyses. *British Journal of Psychiatry*, *200*(2), 97–106.

Levkovitz, Y., Tedeschini, E., & Papakostas, G. I. (2011). Efficacy of antidepressants for dysthymia: a meta-analysis of placebo-controlled randomized trials. *Journal of Clinical Psychiatry*, *72*(4), 509–514.

Linde, K., Kriston, L., Ruecker, G. *et al.* (2015). Efficacy and acceptability of pharmacological treatments for depressive disorders in primary care: systematic review and network meta-analysis. *Annals of Family Medicine*, *13*, 69–79.

Lockhart, P., & Guthrie, B. (2011). Trends in primary care antidepressant prescribing 1995–2007: a longitudinal population database analysis. *British Journal of General Practice*, *61*, 590.

Martens, E. P., Pestman, W. R., de Boer, A., Belitser, S. V., & Klungel, O. H. (2006). Instrumental variables application and limitations. *Epidemiology*, *17*(3), 260–267.

Mattisson, C., Bogren, M., Horstmann, V., Munk-Jorgensen, P., & Nettelbladt, P. (2007). The long-term course of depressive disorders in the Lundby Study. *Psychological Medicine*, *37*(6), 883–891.

Meadows, G.N., Prodan, A., Patten, S., Shawyer, F., Francis, S., Enticott, J., *et al.* (2019). Resolving the paradox of increased mental health expenditure and stable prevalence. *Australian and New Zealand Journal of Medicine*, *53*(9), 844–850.

Mellers, B., Hertwig, R., & Kahneman, D. (2001). Do frequency representations eliminate conjunction effects? An exercise in adversarial collaboration. *Psychological Science*, *12*(4), 269–275.

Mojtabai, R., & Olfson, M. (2014). National trends in long-term use of antidepressant medications: results from the US national health and nutrition examination survey. *Journal of Clinical Psychiatry*, *75*(2), 169–177.

Moncrieff, J. (2002). The antidepressant debate continues – reply. *British Journal of Psychiatry*, *181*, 531–531.

Moncrieff, J. (2003). A comparison of antidepressant trials using active and inert placebos. *International Journal of Methods in Psychiatric Research*, *12*(3), 117–127.

Moncrieff, J. (2018a). Research on a 'drug-centred' approach to psychiatric drug treatment: assessing the impact of mental and behavioural alterations produced by psychiatric drugs. *Epidemiology and Psychiatric Sciences*, *27*(2), 133–140.

Moncrieff, J. (2018b). What does the latest meta-analysis really tell us about antidepressants? *Epidemiology and Psychiatric Sciences*, *27*(5), 430–432.

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

Mulder, R. T., & Frampton, C. M. A. (2014). Outcome of mood disorders before psychopharmacology: a systematic review. *Australian and New Zealand Journal of Psychiatry*, *48*(3), 224–236.

Murray, V., von Arbin, M., Bartfai, A., Berggren, A., Landtblom, A., Lundmark, J., *et al.* (2005). Double-blind comparison of sertraline and placebo in stroke patients with minor depression and less severe major depression. *Journal of Clinical Psychiatry*, *66*(6), 708–716.

NICE. (2011). Common mental health disorders: identification and pathways to care. Clin. Guidel. 123, Natl. Inst. Clin. Excell., London.

Nuijen, J., ten Have, M., Tuithof, M., van Dorsselaer, S., & van Bon-Martens, M. (2014). *Geen gebruik van professionele zorg bij stemmings-en angststoornissen: zorgelijk of niet? Resultaten van de 'Netherlands Mental Health survey and incidence stuy-2 (NEMESIS-2)'.* [*No professional treatment use with anxiety and mood-disorders: a problem or not? Results of the 'Netherlands Mental Health survey and incidence stuy-2 (NEMESIS-2)'.*, Netherlands Institute of Mental Health and Addiction (Trimbos institute), Utrecht, The Netherlands.

Nutt, D. J., Goodwin, G. M., Bhugra, D., Fazel, S., & Lawrie, S. (2014). Attacks on antidepressants: signs of deep-seated stigma? *The Lancet. Psychiatry*, *1*(2), 102–104.

Ormel, J., Oldehinkel, A. J., Brilman, E. I., & van den Brink, W. (1993). Outcome of depression and anxiety in primary care: a three-wave 3 ½-year study of psychopathology and disability. *Archives of General Psychiatry*, *50*(10), 759–766.

Piccinelli, M., & Wilkinson, G. (1994). Outcome of depression in psychiatric settings. *British Journal of Psychiatry*, *164*, 297–304.

Pigott, H. E., Leventhal, A. M., Alter, G. S., & Boren, J. J. (2010). Efficacy and effectiveness of antidepressants: current status of research. *Psychotherapy and Psychosomatics*, *79*(5), 267–279.

Pratt, L. A., Brody, D. J., & Gu, Q. (2017). Antidepressant use among persons aged 12 and over: United States, 2011–2014. NCHS Data Brief, No. 283, 1–8.

Quitkin, F. M., Rabkin, J. G., Gerald, J., Davis, J. M., & Klein, D. F. (2000). Validity of clinical trials of antidepressants. *American Journal of Psychiatry*, *157*(3), 327–337.

Regier, D. A., Kaelber, C. T., Rae, D. S., Farmer, M. E., Knauper, B., Kessler, R. C., & Norquist, G. S. (1998). Limitations of diagnostic criteria and assessment instruments for mental disorders. Implications for research and policy. *Archives of General Psychiatry*, *55*, 109–115.

Rush, A. J., Trivedi, M. H., Wisniewski, S. R., Nierenberg, A. A., Stewart, J. W., Warden, D., *et al.* (2006). Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *American Journal of Psychiatry*, *163*(11), 1905–1917.

Rutherford, B. R., & Roose, S. P. (2013). A model of placebo response in antidepressant clinical trials. *American Journal of Psychiatry*, *170*(7), 723–733.

Sareen, J., Henriksen, C. A., Stein, M. B., Afifi, T. O., Lix, L. M., & Enns, M. W. (2013). Common mental disorder diagnosis and need for treatment are not the same: Findings from a population-based longitudinal survey. *Psychological Medicine*, *43*(9), 1941–1951.

Segal, Z. V., Bieling, P., Young, T., MacQueen, G., Cooke, R., Martin, L., … Levitan, R. D. (2010). Antidepressant monotherapy vs sequential

https://doi.org/10.1017/S0033291719003295 Published online by Cambridge University Press

pharmacotherapy and mindfulness-based cognitive therapy, or placebo, for relapse prophylaxis in recurrent depression. *Archives of General Psychiatry*, *67*(12), 1256–1264.

Sim, K., Lau, W. K., Sim, J., Sum, M. Y., & Baldessarini, R. J. (2016). Prevention of relapse and recurrence in adults with major depressive disorder: systematic review and meta-analyses of controlled trials. *International Journal of Neuropsychopharmacology*, *19*(2), 1–13.

Spijker, J., de Graaf, R., Bijl, R. V., Beekman, A. T. F., Ormel, J., & Nolen, W. A. (2002). Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *British Journal of Psychiatry*, *181*, 208–213.

Spijker, J., Meeuwissen, J., Aalbers, S., van Avendonk, M., van Bon, M., Huson, A., *et al.* (2019). De zorgstandaard Depressieve stoornissen. *Tijdschrift voor Psychiatrie*, *61*(2), 112–120.

Steinert, C., Hofmann, M., Kruse, J., & Leichsenring, F. (2014). The prospective long-term course of adult depression in general practice and the community. A systematic literature review. *Journal of Affective Disorders*, *152*, 65–75.

Turner, E. H., Matthews, A. M., Linardatos, E., Tell, R. A., & Rosenthal, R. (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine*, *358*(3), 252–260.

Van der Lem, R., van der Wee, N. J. A., Van Veen, T., & Zitman, F. G. (2012). Ëfficacy versus effectiveness: a direct comparison of the outcome of treatment for mild to moderate depression in randomized controlled trials and daily practice. *Psychotherapy and Psychosomatics*, *81*, 226–234.

van Weel-Baumgarten, E. M., Schers, H. J., van den Bosch, W. J., van den Hoogen, H. J., & Zitman, F. G. (2000). Long-term follow-up of depression among patients in the community and in family practice settings – a systematic review. *Journal of Family Practice*, *49*(12), 1113–1120.

Verduijn, J., Verhoeven, J. E., Milaneschi, Y., Schoevers, R. A., van Hemert, A. M., Beekman, A. T. F., & Penninx, B. W. J. H (2017). Reconsidering the prognosis of major depressive disorder across diagnostic boundaries: Full recovery is the exception rather than the rule. *BMC Medicine*, *15*, 215.

Vittengl, J. R. (2017). Poorer long-term outcomes among persons with major depressive disorder treated with medication. *Psychotherapy and Psychosomatics*, *86*(5), 302–304.

Vittengl, J. R., Clark, L. A., Dunn, T. W., & Jarrett, R. B. (2007). Reducing relapse and recurrence in unipolar depression: a comparative meta-analysis of cognitive-behavioral therapy's effects. *Journal of Consulting and Clinical Psychology*, *75*(3), 475–488.

Wang, Y., Henriksen, C. A., ten Have, M., de Graaf, R., Stein, M. B., Enns, M. W., & Sareen, J. (2017). Common mental disorder diagnosis and need for treatment are not the same: Findings from the NEMESIS study. *Administration and Policy in Mental Health and Mental Health Services Research*, *44*(4), 572–581.

Whitaker, R. (2015). *Anatomy of an epidemic*. New York: Broadway Books.

Whiteford, H. A., Harris, M. G., McKeon, G., Baxter, A., Pennell, C., Barendregt, J. J., & Wang, J. (2013). Estimating remission from untreated major depression: A systematic review and meta-analysis. *Psychological Medicine*, *43*(8), 1569–1585.

Wildeboer, R. M., van der Hoek, L., & Verhaak, P. F. M. (2016). Use of general practice services 5 years after an episode of mental illness: Case-control study using electronic records. *British Journal of General Practice*, *66* (646), E347–E353.

Zimmerman, M., Posternak, M. A., & Ruggero, C. J. (2007). Impact of study design on the results of continuation studies of antidepressants. *Journal of Clinical Psychopharmacology*, *27*(2), 177–181.

https://doi.org/10.1017/S0033291719003295 Published online by Cambridge University Press