More treatment but no less depression

Ormel, Johan; Hollon, Steven D; Kessler, Ronald C; Cuijpers, Pim; Monroe, Scott M

Published in:
Clinical Psychology Review

DOI:
10.1016/j.cpr.2021.102111

Publication date:
2022

Clinical Psychology Review 91 (2022) 102111

Contents lists available at ScienceDirect

Clinical Psychology Review

journal homepage: www.elsevier.com/locate/clinpsychrev

Review

More treatment but no less depression: The treatment-prevalence paradox

Johan Ormel [a,*], Steven D. Hollon [b], Ronald C. Kessler [c], Pim Cuijpers [d], Scott M. Monroe [e]

[a] University of Groningen, University Medical Center Groningen, Department of Psychiatry, Groningen, the Netherlands
[b] Department of Psychology, Vanderbilt University, Nashville, TN, USA
[c] Department of Health Care Policy, Harvard Medical School, Boston, MA, USA
[d] Department of Psychology, Free University, Amsterdam, the Netherlands
[e] Department of Psychology, University of Notre Dame, IN, USA

ARTICLE INFO

Keywords:
Depression
Treatment
Prevalence
More treatment but not less depression
Explanations treatment-prevalence paradox

ABSTRACT

Treatments for depression have improved, and their availability has markedly increased since the 1980s. Mysteriously the general population prevalence of depression has not decreased. This "treatment-prevalence paradox" (TPP) raises fundamental questions about the diagnosis and treatment of depression. We propose and evaluate seven explanations for the TPP. First, two explanations assume that improved and more widely available treatments have reduced prevalence, but that the reduction has been offset by an increase in: 1) misdiagnosing distress as depression, yielding more "false positive" diagnoses; or 2) an actual increase in depression incidence. Second, the remaining five explanations assume prevalence has not decreased, but suggest that: 3) treatments are less efficacious and 4) less enduring than the literature suggests; 5) trial efficacy doesn't generalize to real-world settings; 6) population-level treatment impact differs for chronic-recurrent versus non-recurrent cases; and 7) treatments have some iatrogenic consequences. Any of these seven explanations could undermine treatment impact on prevalence, thereby helping to explain the TPP. Our analysis reveals that there is little evidence that incidence or prevalence have increased as a result of error or fact (Explanations 1 and 2), and strong evidence that (a) the published literature overestimates short- and long-term treatment efficacy, (b) treatments are considerably less effective as deployed in "real world" settings, and (c) treatment impact differs substantially for chronic-recurrent cases relative to non-recurrent cases. Collectively, these a-c explanations likely account for most of the TPP. Lastly, little research exists on iatrogenic effects of current treatments (Explanation 7), but further exploration is critical.

It is widely believed that treatments for major depression have improved since the 1980s, and that these treatments have become more widely available for helping depressed people. Surprisingly, there has not been a commensurate reduction in the population prevalence of depression. We refer to the increasing availability of better treatments, juxtaposed with the absence of a corresponding decrease in depression's prevalence, as the *treatment-prevalence paradox* (TPP).

The present article combines conceptual analysis, along with evidence based virtually entirely on recent meta-analyses, to address the TPP. Unless otherwise indicated, throughout the manuscript depression refers to major depressive disorder (MDD), and prevalence refers to *point*-prevalence (i.e., the percentage of people who meet diagnostic criteria for depression at a particular point in time, typically the 30 days preceding the examination). (In contrast, lifetime prevalence refers to the percentage of people who develop depression at any point in their life.) The TPP requires critical attention and analysis, as it signifies unforeseen shortcomings in understanding of depression and its treatments (e.g., see also (Jorm, Patten, Brugha, & Mojtabai, 2017; Meadows et al., 2019). We propose and analyze seven possible explanations and evaluate the evidence (or lack thereof) bearing on each.

Depression is a primary public health concern worldwide, with prevalence of 4.7% (95% CI 4.4–5.0%) (Bromet et al., 2011; Ferrari et al., 2013). Depression contributes to lowered work productivity, family dysfunction, substance misuse, suicide, and reduced life expectancy (Murray et al., 2015; J. Ormel et al., 1994; Vos et al., 2017). High rates of depression in the general population were documented during the first three post-WWII decades (Hagnell, Lanke, Rorsman, & Öjesjö, 1982; Klerman, 1988; J. M. Murphy, Laird, Monson, Sobol, & Leighton,

* Corresponding author.
E-mail addresses: j.ormel@umcg.nl (J. Ormel), steven.d.hollon@Vanderbilt.Edu (S.D. Hollon), kessler@hcp.med.harvard.edu (R.C. Kessler), p.cuijpers@vu.nl (P. Cuijpers), smonroe1@nd.edu (S.M. Monroe).

https://doi.org/10.1016/j.cpr.2021.102111
Received 1 July 2021; Received in revised form 28 October 2021; Accepted 6 December 2021
Available online 11 December 2021

Table 1
Increased trends in mental health expenditure and treatment over recent decades in five Western countries.[#]

Country	Expenditure/ Workforce	Medications (Antidepressants)	Psychological treatments
Netherlands (Bongers, 2009; Niaounakis, 2013)	2000–2010 110% increase in costs. Number of jobs doubled.	1987–2001 PCP prescribing medications for depression from 44% to 80%. 1996–2013 yearly increase with 4%. Stabilizing since 2018.	No differentiated data on increase psychological treatment, only indirect (increase of mental health care costs 110%).
Australia (Jorm et al., 2017)	1992–2011 178% increase in expenditure. Workforce 35% increase per capita.	1990–2002: 352% increase 2000–2011: 95% increase (2nd highest in OECD).	2006–2012 substantial increase in CBT and e-therapy.
Canada (Patten, Williams, Lavorato, Bulloch et al., 2016)	2002–2013 indirect evidence: increase of 40%.	1994–2012: 3-fold increase; stabilized since (3rd highest in OECD).	2002–2012 indirect evidence: 28% to 40% increase (6+ visits).
England (Brugha et al., 2004; Spiers et al., 2016)	NPMS 1993–2007: Small increase in contact with GP for MH problems.	1993–2000-2007: 3-fold increase until 2000; small increase since.	Limited evidence. Since 2005, some increase due to the IAPT program.
United States (Kessler et al., 2005; Marcus & Olfson, 2010a; Mojtabai & Olfson, 2014; Mojtabai, Olfson, & Han, 2016; Olfson & Marcus, 2010)	1987–1996 23% increase in treatment. 1991–2002 65% increase in treatment.	Strong increase in those diagnosed with MDD from 37% in 1987 to 74% in 1997 to 82% in 2007.	Gradual decrease in those diagnosed with MDD from 71% in 1987 to 53% in 1998 to 43% in 2007.

Note. MDD, Major depressive disorder. PCP, Primary care provider. MH, Mental health. NPMS, National psychiatric morbidity study. APMS, Adult psychiatric morbidity study. IAPT, Improving access to psychological therapies.
[#] See for references the text as well.

Table 2
Changes in prevalence of depression over recent decades in five Western countries.[1] [#]

Country	Research diagnosis (major depressive disorder, MDD)[2]
Netherlands (R. de Graaf, Ten Have, van Gool, & van Dorsselaer, 2012; Meertens, Scheepers, & Tax, 2003; Schoemaker, Ten Have, Sytema, & Verhaak, 2007)	PSE 1983–1998 Depression: No change. GHQ-30[3] 1983–1998: increase from 3.1 to 4.6. SCP=CBS Depressive symptoms: 1975–1996: Stable with fluctuation between 1981 and 1991. CIDI in NEMESIS 1996–2008: Mood disorder: 12-month prevalence 7.4 in 1996 and 6.1 in 2008 (no difference after adjustment for sociodemographic differences).
Australia (Jorm, 2014)	CIDI 1997–2007: no reduction (18% to 20%).
Canada (Patten, Williams, Lavorato, Bulloch et al., 2016; Patten, Williams, Lavorato, Wang et al., 2016)	CIDI in NHPS and CCHS: MDE 1994–2012: No time trend. Annual prevalence ~5%. CCHS 2001–2013: No evidence of time trend in episode duration. NPHS 1994–2010: No evidence of change over time in new episodes. Pooled 1.8%.
England (Spiers et al., 2016)	CIS-R in NPMS 1993–2000-2007: Small increase CMDs 14.3% to 16.0 to 16.0%.
United States (Compton, Conway, Stinson, & Grant, 2006; Kessler, Berglund, Borges, Nock, & Wang, 2005; Mojtabai & Jorm, 2015)	CIDI in NCS 1991–2003: CMDs no change; MDD 10.1% to 8.7% (drop due to demographic change). AUDADIS in NESARC 1991–2002: increase in MDD 3.3% to 7.0%.
Global Burden of Disease (Vos et al., 2015; Vos et al., 2017)	Increase 1990–2013 by 42% (anxiety) and 53% depression, largely due to population growth and aging.

Note. CMD, Common mental disorders, typically include affective and anxiety disorders and inconsistently substance abuse/dependence. AUDADIS, Alcohol Use Disorder and Associated Disabilities; Interview Schedule. CIDI, Composite International Diagnostic Interview. CIS-R, Clinical Interview Schedule-Revised. PSE, Present State Examination. PHQ, Present Health Questionnaire. GHQ, General Health Questionnaire. NS, neurotic symptoms. K10: Kessler 10. NHS, National Health Surveys. NCS, National Comorbidity Studies. NESARC, National Epidemiologic Survey on Alcohol and Related Conditions/ NPMS, National Psychiatric Morbidity Survey. NEMESIS, Netherlands mental health survey and incidence study. MHI, Mental Health Inventory. SCP, Sociaal en Cultureel Planbureau-depressive symptoms. NHANES, National health and Nutrition study. NHIS, National Health Interview Survey.
[1] Most prevalence data concern 12-month prevalence but some refer to 1-month prevalence.
[2] Research diagnostic data refer to MDD although there are (small) differences in diagnostic criteria between instruments and classifications. Occasionally trend data refer to common mental disorders (CMDs) which include also anxiety disorders.
[3] Symptoms refer typically to symptoms of depression, but some measures are broader, tagging psychological distress (e.g., the GHQ).
[#] See for references the text as well.

2000). Since these alarming epidemiological publications, mental health care expenditures and MDD treatment have increased sharply, as shown in Table 1 (Jorm et al., 2017; Mark, Levit, Vandivort-Warren, Buck, & Coffey, 2011; Niaounakis, 2013; M. Olfson, Kroenke, Wang, and Blanco, 2014b). These increases are especially apparent for antidepressant medications (further abbreviated as "medication," including amongst others trycyclic drugs and SSRIs). Hence, many more depressed people have received treatment, especially in general medical settings over the period 1985–2010. For instance, in the US the rate of outpatient treatment for depression increased from 0.73/100 in 1987 to 2.88/100 in 2007, a *four-fold surge* over just two decades (Marcus & Olfson, 2010a; M. Olfson, Marcus, Druss, & Pincus, 2002; M. Olfson, Marcus, Wan, & Geissler, 2004).

The increased availability of effective treatments should shorten depressive episodes, reduce relapses, and curtail recurrences. Combined, these treatment advances unequivocally should result in lower point-prevalence estimates of depression. Have these reductions occurred? The empirical answer clearly is NO (Table 2). Recent meta-analyses of epidemiological surveys in the general population of Western countries since 1980 do not report decreasing prevalence rates of depression. In fact, there may have been a slight increase as two out of three meta-analyses published since 1978 found a small upward temporal trend, while the third meta-analysis reported an unchanged prevalence ((Ferrari et al., 2013): OR = 1.02 (95%CI 1.01–1.04); (Richter, Wall, Bruen, & Whittington, 2019): 1.29 (95%CI:1.06–1.58); (Baxter et al., 2014): unchanged 4.4% (4.2–4.7%)). The minimal increase probably is best explained by changes in demographic composition over time (Collaborators, 2018).

Major repeated cross-sectional studies also fail to find any decrease in the prevalence of depression. The US-based National Comorbidity Study (NCS, $N \sim 5000$) and the Dutch-based NEMESIS study ($N \sim 7000$), which both examined trends over periods of 10–12 years around 2000, each failed to find significant change in 12-month prevalence of MDD (de Graaf et al., 2012; Kessler et al., 2005). However, more recently two large repeated cross-sectional studies reported a small, but statistically significant, *increase* (Compton, Conway, Stinson, & Grant, 2006; Weinberger et al., 2018). We conclude that a prevalence decrease is highly unlikely, but that a small increase since the turn of the century

Table 3
Overview of explanations of the treatment-prevalence paradox (TPP), Available evidence and preliminary conclusions.

Possible Explanations for the TPP	Evidence and (Preliminary) Conclusions
1) Have prevalence estimates been spuriously inflated due to increasing societal recognition of depression and associated diagnostic practices?	People have probably become more willing to admit depressive symptoms and to present for treatment, where they may receive false-positive diagnoses of MDD. But since epidemiologic surveys are conducted by well-trained interviewers using structured interviews to generate well-standardized diagnoses, it is unlikely that systematic drift at the population level of 'caseness' has occurred. Thus, an increase in "false positives" diagnoses would not mask a treatment-driven drop in "true" epidemiological prevalence.
2) Have first incidence rates increased and offset a "true" treatment-driven reduction in point-prevalence?	Post-1980 first incidence studies and information on trends in causal risk factors are few, and too inconsistent to provide a conclusive answer. The handful of incidence studies, though, do not hint at any significant rise in incidence since the 1980's, and some even suggest a decrease. The evidence is sparse and uncertain, though, and ends around 2010. It seems unlikely that a true increase in depression incidence offsets any treatment-driven prevalence reduction.
3) Do RCTs overestimate Acute-Phase treatment efficacy? Might biases both within the trials and across the larger literature on medication and psychotherapy inflate these short-term benefits of treatment?	RCTs do yield inflated acute-phase efficacy estimates. Adjusted for bias, efficacy drops by a third to half to modest effect sizes at best (about 0.30 for medications vs. pill-placebo and psychotherapy vs. care-as-usual). It is unclear how long Acute-Phase treatment benefits persist. Given the biases and large heterogeneity, it is not surprising that there is significant disagreement about the clinical impact of treatments. It is not that treatments do not work, just that they do not work as well as the published literature suggests, or as is widely believed. Hence, the more accurate estimate of short-term efficacy is at best modest, which can explain in part the TPP.
4) Does research on maintenance of treatment gains and long-term efficacy over estimate beneficial effects? Are medication and psychotherapy interventions to prevent relapse-recurrence upwardly biased due to non-eligibility, insufficient response to Acute-Phase treatment, symptom return risk, and a variety of biases that need to be taken into account?	RCTs evaluating treatments aimed to reduce relapse-recurrence risk show substantial efficacy for preventive psychotherapy and for continued medication. However, these "effects" are rife with possible biases (misclassification, unblinding, allegiance effects, and differential mortality) complicating interpretation. In addition, many patients without sufficient response to acute-phase treatment are not eligible for relapse or recurrence prevention trials, and relapse-recurrence rates over two years after preventive treatment remain substantial, (though estimates vary greatly). Hence, limited overall long-term efficacy also may help to explain the TPP.
5) Do RCTs generalize to real-world settings? How large is the gap between RCT-based efficacy and real-world effectiveness?	RCT-based efficacy does not generalize all that well to real-world practice, both for medication and for psychotherapy. Reasons: Large gaps in treatment choice and implementation quality exist in real-world practice; compared to the typical RCT patient, the real-world patient is somewhat less treatable (suicidal ideation; addiction, severe comorbidities). What the gaps, along with naturalistic follow-up studies, tell us is that treatment is not as effective long-term as we would like it to be. This explanation appears to be one of the strongest candidates for understanding the TPP as it also amplifies the contribution of explanations 3 and 4.
6) Does treatment efficacy vary by different subtypes of depression? Specifically, could differential treatment benefits for chronic-recurrent versus non-recurrent cases dilute the potential beneficial effects of treatments for those most in clinical need? Further, chronic-recurrent cases are often very difficult to treat, or treatment-resistant.	The majority of people who initially become depressed have few if any recurrences, whereas recurrent and chronic cases become or remain depressed for much more time over the course of their lives. The availability of more and better treatments consequently has many more opportunities to benefit the smaller number of chronic-recurrent cases, while treatment effects at the population level for the many more non-recurrent cases most likely will be very limited. The resulting limited effects at the population level for the larger non-recurrent group could dilute more pronounced effects for the chronic-recurrent subgroup, obscuring a positive impact on prevalence for those in greatest need. However, it is unclear to what extent advances in preventive treatments specifically benefit the chronic-recurrent subgroup, or if these treatments are adequately transported into routine care for them (Explanations 3–5). Individually and combined, these subgroup considerations also provide potentially strong explanations for the TPP.
7) Can treatment sometimes also have counterproductive consequences?	Oppositional perturbation refers to a medication-induced state of built-up perturbation in homeostatic monoamine regulatory mechanisms that "bounces back" when medication is discontinued, and then overshoots the normal balance of monoamine storage and release, increasing the risk for symptom return compared to spontaneous remission. Loss of agency refers to the hypothesis that either medication or psychotherapy could be counterproductive if either or both reduce self-help activity and active coping and thereby interfere with natural recovery mechanisms. Although some indirect evidence exists for each possibility, both mechanisms are largely speculative. The explanatory potential of this concern remains to be demonstrated for understanding the TPP, but is worthy of further investigation.

remains possible, especially in adolescents (Daly, 2021).

How can the TPP be understood? Logically, there are two basic sets of scenarios: The first set assumes that increased treatment has reduced prevalence, but that the prevalence reduction has been offset by: 1) an increase in false positive diagnoses in recent prevalence studies, or 2) an actual increase in the incidence of depression. The second set of scenarios includes the remaining five explanations, all of which assume prevalence has not decreased, yet suggest instead that: 3) treatments are less efficacious, or 4) less enduring than the literature suggests; 5) trial efficacy doesn't generalize well to real-world settings; 6) population-level treatment impact differs substantially for chronic-recurrent cases versus non-recurrent ones; and 7) treatments can have both beneficial and iatrogenic consequences. If valid, these latter explanations would

Table 4
Bias in RCTs and meta-analyses.

Type of bias	Description	References
Publication bias	Non-publication of trial with nonsignificant or negative findings.	(Turner, Matthews, Linardatos, Tell, & Rosenthal, 2008)(Driessen, Hollon, Bockting, Cuijpers, & Turner, 2015)
Selective reporting or *Outcome reporting bias*	Failure to describe negative findings within a published report or switching the status of (nonsignificant) primary and (significant) secondary outcomes	(Cooney et al., 2013)
Outcome misclassification bias	Measures and assessors are imperfect. In studies that discontinue medication, withdrawal symptoms may masquerade as depressive symptoms, thereby conflating the two.	(Fava, Gatti, Belaise, Guidi, & Offidani, 2015)
Imperfect blinding	Patients, treatment providers or assessors know the true status of the randomized subjects: intervention or control condition. In medication trials, this may occur because of side effects.	(Kirsch, 2014; Moncrieff, Wessely, & Hardy, 1998)
Spin bias	Reporting strategies that often mislead readers	(Boutron et al., 2014; De Vries et al., 2018)
Citation bias	Trials with positive results receive more citations than negative studies, leading to a heightened visibility of positive findings and reduced discoverability of negative trials.	(Boutron et al., 2014; De Vries et al., 2018)

undermine any treatment impact on prevalence, thereby explaining the TPP. Below we elaborate these 7 alternative explanations and evaluate their credibility given relevant evidence. Table 3 summarizes the 7 explanations and – spoiler alert – our ultimate conclusions.

1. Method and material

Our objectives are to describe seven candidate explanations for the TPP, evaluate each in light of available evidence, and identify knowledge gaps that might inform future research aimed at resolving the paradox. Our method is best characterized as integrative narrative review that uses recent comprehensive meta-analyses, systematic reviews, silver bullet studies, and logic. We do not present a formal umbrella review of meta-analyses. According to APA rules and PRISMA guidelines, the number and diversity of questions that the candidate explanations raise are too numerous for systematic (umbrella) reviews to integrate the material in a concise fashion. The alternative would have been at least five umbrella reviews. Instead, our approach combines conceptual analysis with evidence, based on major population-based and epidemiological studies and recent comprehensive meta-analyses of randomized clinical trials (RCT). This allows us to evaluate each potential (sub)explanation in terms of availability of evidence, and, if the evidence is sufficient, its explanatory value for understanding the TPP. What we set out to do was to describe how the TPP can be explained, given where we thought the field currently is.

Current psychiatric conventions distinguish between three treatment phases (acute, continuation, maintenance), and differentiate treatment response (typically defined as a 50% reduction in symptoms over baseline) from remission (the full normalization of symptoms) (Frank et al., 1991). Both terms are further differentiated from recovery, the presumed end of the current underlying episode. The field further distinguishes between relapse (the return of symptoms associated with the current episode) versus recurrence (the onset of a wholly new episode). The term "symptom return" is often used to refer to either relapse or recurrence. For treatment to reduce prevalence, it must shorten the episode or prevent relapse or recurrence.

There are hundreds if not thousands of individual RCTs on depression treatments, far too many, and too heterogeneous in methodological quality and results for summarizing. Accordingly, we almost exclusively use comprehensive meta-analyses and systematic reviews to inform our thinking and turn quantitative syntheses into digestible conclusions (Hollon et al., 2014). We include individual patient data meta-analyses and recently developed network meta-analyses, which generate a web of interrelatedness that permit to estimate differences in effects between conditions that have never been directly compared. We draw on research largely from four English-speaking countries (Australia, Canada [bilingual], UK, and US) and the Netherlands (to represent western continental Europe). These countries are known for high-quality research on these pertinent topics, as well as for wide accessibility of resources.

A variety of biases threaten the validity of individual RCTs and meta-analyses and may inflate apparent efficacy. Best known of these are publication bias, outcome reporting bias, and a variety of other biases, usually assessed with the Cochrane 'Risk of Bias' Tool. Table 4 provides an overview. Only the most recent comprehensive meta-analyses adjust for quantifiable biases.

The cumulative effect of all biases is unclear, but publication bias alone has inflated apparent treatment efficacy by a third (32%) (De Vries et al., 2018; Turner et al., 2008). As Fig. 2 suggests, the cumulative impact of all biases could be considerable. Because pharmaceutical companies are required by law to preregister all trials they intend to use to obtain approval to market from the US Food and Drug Administration (FDA), these reviews could quantify publication and outcome reporting bias. Hence, nonsignificant results are still accessible, even if never published, and published results can be checked against the archive for accuracy.

This article is structured as follows for understanding the origins of the TPP. First, we address the two explanations that assume that treatment-driven prevalence reduction has occurred, but has been offset by an increase in false positives or true incidence. Then, we move to explanations targeting acute-phase treatment efficacy, followed by efficacy of interventions to prevent relapse (continuation phase) and recurrence (maintenance phase). Next, we investigate explanations involving the extent RCT's efficacy generalizes to real-world settings, and then the degree to which treatments might benefit some types of patients more than others (recurrent versus non-recurrent subgroups). Finally, we address explanations related to the possibility that treatments might have adverse consequences, which in turn dilute their impact on population prevalence.

2. Explanation 1: Have increased false positive diagnoses masked a treatment-driven prevalence drop?

Can the epidemiological evidence that the prevalence of depression has not decreased be trusted? Or has a drift towards greater willingness to admit depression to interviewers in surveys, or more lenient diagnostic criteria, spuriously increased the detection of prevalence, thereby masking any underlying treatment-driven prevalence drop? The lack of any definitive physiological criteria for depression (e.g., a depression "thermometer"), coinciding with imperfections in measurement and diagnostic systems (G. Andrews & Peters, 1998; Brugha, Jenkins, Taub, Meltzer, & Bebbington, 2001; Wittchen, Kessler, Zhao, & Abelson, 1995), makes 'diagnostic creep' all the more worthy of investigation (Haslam, 2016). As all additional explanations for the TPP critically hinge upon this basic question, a conservative approach in evaluating the matter is needed.

Over recent decades a significant public mental health movement has unfolded to legitimize depression and its treatment. Pharmaceutical

Table 5
Illustrative cross-classification of true and observed negative and positive cases.

	True status		
Observed status	Negative (N)	Positive (P)	Total
Negative (N)	(a) True N 82	(b) False N 4	86
Positive (P)	(c) False P 4	(d) True P 10	14
Total	86	14	100

Sensitivity = d/b + d; Specificity = a/a + c.
ADMIT scenario = One or more false N become true P.
EXXAGERATE scenario = One or more true N become false P.

advertising conveyed the message that depression is a biological condition, caused a chemical imbalance in the brain (not a weakness in character), that can be conveniently remedied with the 'right' medication (Donohue, Cevasco, & Rosenthal, 2007; J. Moncrieff, 2018). These developments fueled two major trends (Schomerus et al., 2012): (1) A greater emphasis on biological explanations (although psychosocial explanations, particularly stress-related, continue to enjoy high popularity); (2) More positive attitudes towards seeking professional help and greater openness about mental health problems. These trends reduced the stigma associated with depression and increased the acceptability of depression and its treatment, both for others and for oneself (Mojtabai, 2007; Phelan, Link, Stueve, & Pescosolido, 2000; Reavley & Jorm, 2014; Thornicroft et al., 2016).

In concert with other social-cultural developments, these trends may have fostered a greater sensitivity to emotional distress, increased willingness to disclose depressive symptoms, and more lenient use of diagnostic criteria in general health care (Horwitz, 2007; Mulder, 2008; Wakefield, 2002). Direct-to-consumer advertising of antidepressant medication, beginning in the early 1990s, along with the necessity of assigning DSM diagnosis to qualify for reimbursement, elevated the presentation and recognition of "normal" distress as possible depression to health care professionals. As a consequence, there was an increase in the diagnosis of 'depressive disorder' and 'depressive symptoms' in general medical settings, phenomena often denoted as the medicalization of distress (Donohue et al., 2007; Furer, 2001; Meadows et al., 2019; J. Moncrieff, 2018). For example, in Dutch general practice these diagnoses increased almost *four-fold*, rising steadily from 1.1/100 registered patients in 1990 (Velden, De Bakker, Claessens, & Schellevis, 1991) to 2.7/100 in 2000 (Van der Linden, Westert, Bakker, & Schellevis, 2004) and then to 4.0/100 in 2015 (Nuijen et al., 2018).

Have greater willingness to disclose symptoms and medicalization of distress influenced 'hard' data on population prevalence rates, thereby obscuring any treatment-driven prevalence drop? Two considerations suggest the likely answer is "NO." First, as shown in Table 5 the impact can go either or both ways. In the "ADMIT" scenario, symptoms that previously were ignored or unmentioned are now endorsed, flipping prior false-negative cases into true-positive ones, resulting in higher and now more valid prevalence rates (due to increased sensitivity). In contrast, in the "EXAGGERATE" scenario, symptoms previously viewed as subclinical distress, are now 'amplified' and endorsed as depressive symptoms, flipping prior true-negative cases into false-positive ones, resulting in spuriously higher, less valid, prevalence rates (due to reduced specificity). In both scenarios, the estimated prevalence will go up even if the true prevalence remains unchanged.

The second consideration involves the degree to which any potential increased willingness to disclose and medicalization of normal distress have worked their way into formal diagnostic practices in population-based epidemiological surveys. Clinical validity studies suggest that it is less likely that epidemiological research, which typically uses structured diagnostic interviews administered by well-trained lay-interviewers (e.g., DIS, CIDI), has become more lenient in applying the standardized decision rules for rendering research diagnoses as judged against the semi-structured clinical interviews administered by clinical experts (e.g., PSE, SCAN) (G. Andrews & Peters, 1998; Brugha et al., 2001; Haro et al., 2006; Kessler et al., 2009; Wittchen et al., 1995). Overall, it is doubtful that significant drift in case-definition and -ascertainment has occurred in epidemiological studies employing structured interviews, standardized classification, and well-trained interviewers.

The slight prevalence increase reported in Richter's and Ferrari's meta-analyses could be biased upward (inflated), due to counting instances previously considered to be false negatives as true positives (ADMIT scenario, true 'correction'), or 2) counting previous true negatives as false positives (EXAGGERATE scenario, inflated prevalence). *Either would suggest that there has been no substantive increase in true prevalence*. In contrast, Baxter's meta-analysis reported that depression's prevalence has remained stable. In this latter empirical context, both hypothetical scenarios may actually imply a reduction in true-prevalence, because the observed prevalence can only remain stable with fewer false negatives or fewer true-positives (Table 4).

We would need an objective depression marker to establish with certainty that depression's prevalence has not decreased. Such markers do not exist. Suicide is a very crude proxy of depression and it is worth noting that the official suicide rate has for the most part been flat or decreased in most Western countries (but not the USA with its OxyContin crisis) over recent decades. But official suicide rates have biases as well, and these might have changed over time. Furthermore, even though depression is the mental disorder most strongly associated with suicide, only a tiny proportion of depressed people die by suicide, making changes in the suicide rate an unreliable indicator of changes in depression prevalence. That said, the suicide rate has been stable or gone down whereas the prevalence rate has remained stable or gone up slightly.

2.1. Conclusion Explanation 1

Collectively, the available studies and scenarios suggest that the true prevalence of depression has not decreased systematically over the past 30–40 years. Indeed, if anything prevalence may have increased (especially amongst youth). Consequently, neither greater willingness to disclose symptoms nor greater medicalization of distress over the last few decades is likely to have masked any significant underlying treatment-driven decrease in MDD prevalence. Explanation 1 can be ruled out as a strong explanation for the TPP.

3. Explanation 2: Has an increase in MDD incidence offset any treatment-driven decrease in prevalence?

The epidemiological evidence just reviewed supports the conclusion that the prevalence of depression has not decreased. But what if first *incidence* has increased? If more people are becoming depressed in the first place, the rise in incidence could offset any treatment-driven decrease in prevalence, thereby at least partially explaining the TPP.

Because the unexpected lack of any decrease in prevalence refers to the post-1980 decades, we searched the literature for population-based incidence studies of depression covering post-1980 periods that met the criteria of random samples, modern diagnostic classifications, and standardized interviews administered by trained lay interviewers. A Web-of-Science search using the string (TI = (incidence) AND (depression)) yielded 618 hits, but virtually all referred to special populations (post-partum women, children, disabled elderly, chronic disease, etc.), used non-standardized methods, or presented total incidence in the interval and not first incidence (throughout the rest of the manuscript incidence refers to first incidence unless indicated otherwise).

Some epidemiological prevalence studies have provided lifetime prevalence rates for their samples (did you ever in your life experience …). If unbiased, lifetime prevalence is equal to first incidence. Unfortunately, lifetime estimates suffer from recall bias and strongly

Table 6
First Incidence Studies with Short-Term Follow-Up, Using Structured Diagnostic Interviews and Standardized Modern Classifications.

Country	N (age range); case-finding	Data collection: Duration follow-up	Person-years at risk	Annual incidence per 1000
ECA, USA (Eaton et al., 1989)	10,035 (18+); DIS, DSM-III	1981–1982; 12 months	10,035	15.9
Edmonton, Canada (Newman & Bland, 1998)	1964 (18–64); DIS, DSM-IV	1986–1989; 33 months	5499	27.9
Netherlands, NEMESIS-1 (Bijl, de Graaf, Ravelli, Smit, & Vollebergh, 2002)	5618 (18–64); CIDI, DSM-IIIR	1997–1999; 12 months	5618	*27.2*
ODIN, Finland (Lehtinen et al., 2005)	1412 (18–64); SCAN, ICD-10	1998–1999; 12 months	1412	20.5
USA, NESARC (Grant et al., 2009)	28,859 (18+); AUDADIS-IV	2004–2006; 12 months	28,614	15.1
Netherlands, NEMESIS-2 (R. de Graaf, ten Have, Tuithof, & van Dorsselaer, 2013)	4172 (18–64); CIDI, DSM-IV	2008–2011; 36 months	12,311	15.8

Note. AUDADIS, Alcohol Use Disorder and Associated Disabilities Interview Schedule. CIDI, Composite International Diagnostic Interview.
SCAN, successor of the Present State Examination (PSE). DSM, Diagnostic Statistical Manual.

underestimate first incidence (T. E. Moffitt et al., 2010; Patten, 2009; Simon & Von Korff, 1995).

3.1. Direct evidence on MDD first incidence

The 6 studies shown in Table 6 show little indication of an increase in annual incidence. Indeed, if anything they suggest that there has been a decrease between 1985 and 2011. The two large USA studies (the Epidemiological Catchment Area [ECA]) and the National Epidemiologic Survey of Alcoholism and Related Conditions [NESARC]) suggest stability of the annual incidence rate of about 15–16 per 1000, whereas the two Dutch studies (NEMESIS-1 and -2) suggest an incidence drop. The three studies that report substantially higher incidence rates were performed outside the USA: in Canada, Finland, and the Netherlands ([Nemesis-1]). It is unclear why these three studies report higher rates, but their samples were small. The second Dutch study (Nemesis-2) is more in line with the large USA studies in suggesting an incidence drop as well. To place the depression incidence rates in perspective, annual incidence rates per 1000 are 0.6 for lung cancer, 4.5 for stroke, 15 for cardiovascular disease, and 50 for hypertension (National Heart, Lung, and Blood Institute, 2006; Ries et al., 2006).

Table 6 requires two comments. First, the lower age range in Table 6 incidence studies is typically 18 while most first episodes occur during late adolescence and early adulthood (age 15–34) (Bijl et al., 2002; Bromet et al., 2011). Hence the studies may have missed individuals with episodes that remitted before age 18 and did not experience a recurrence. However, the 18 lower age limit applies to all Table 6 incidence studies and is unrelated to study year. Consequently, some underestimation of first incidence is likely but affects all studies since 1980.

Second, as mentioned previously, measurement and diagnostic systems are not perfect, and their performance may have altered over time as a result of changes in societal attitudes about mental illness and the surrounding stigma. If recent population-based incidence studies have become more sensitive and capable of correctly identifying previous false-negative cases (updated as true cases), the observed stable incidence rates now actually would suggest a *downward* trend in true depression incidence. This conclusion would be true as well if recent population surveys have produced the less likely, but still feasible, scenario of more false-positive cases. All of these points bolster the argument for no major upward changes in incidence over recent decades, and forces further recognition of the importance of understanding the TPP.

We conclude that the handful of studies do not hint at a significant rise in first incidence since the 1980's, and indeed some suggest a decrease. But it should be acknowledged that the evidence is sparse and uncertain, and ends around 2010.

3.2. Indirect evidence on incidence: trends in causal risk factors

If incidence had increased, there should be a commensurate increase in risk factors for the depression. We reiterate that, since the question of increased incidence is so foundational for interrogating the TPP, we also evaluate changes in known environmental and personality risk factors for depression.

MDD is understood to be very heterogeneous in terms of symptoms, course and complex multifactorial etiology, implicating both personal and environmental factors. From the perspective of population-attributable risk factors, salient person characteristics include genetic influences and personality traits. Prospective studies show that prominent individual-level environmental risk factors include childhood adversity, long-term difficulties such as poverty, and major stressful life events. Distal, societal-level risks include war, economic crises, inequality, and lack of social cohesion.

3.2.1. Person factors

Four decades is too brief a period for major changes to occur in the human genome. To some extent this also applies to personality traits, as personality change typically is temporary (returning to person-characteristic set-points after life events). But significant persistent change has been documented in response to long-term environmental change, such as entering a stable marriage or long-term unemployment (J. Ormel, VonKorff, & Riese, 2017; Roberts & Mroczek, 2008; Roberts, Hill, & Davis, 2017). Prospective studies show that depression is predicted, in terms of the five-factor personality model, primarily by high neuroticism, and secondarily low extraversion and low conscientiousness (Jeronimus, Kotov, Riese, & Ormel, 2016; Kotov, Gamez, Schmidt, & Watson, 2010; Malouff, Thorsteinsson, & Schutte, 2005; T. E. Moffitt et al., 2011).

Longitudinal studies typically report slight increases in extraversion, agreeableness, and conscientiousness, along with slight decreases in neuroticism (although the evidence on neuroticism is less consistent) (Mroczek & Spiro, 2003; Smits, Dolan, Vorst, Wicherts, & Timmerman, 2011; A. Terracciano, McCrae, Brant, & Costa, 2005; A. Terracciano, 2010; A. Terracciano, McCrae, & Costa, 2010; Trzesniewski & Donnellan, 2010). Meta-analyses of individual scores find rather meager personality changes, but these analyses are not exclusively based on standard personality scales (Trzesniewski & Donnellan, 2010). Analyses of personality scores from around the world (cross-national/cultural studies) suggest that social and cultural differences do not account for much variance in major dimensions of personality, but the studies did not test metric equivalence (A. Terracciano, 2010). In sum, the apparently modest changes in these traits are unlikely to account for major change in depression incidence. But even if they did, there would be a *decrease* in incidence.

3.2.2. Environmental factors

Unprecedented technological developments have seen the light since the 1990s, including the emergence of the Internet, smart phones, and social media. This holds even more for sociocultural change, driven by technological and economic developments (e.g., globalization), conflicts

and climate change, fueling migration. We collected information on trends in environmental factors monitored by statistical offices or addressed in national surveys (e.g., child maltreatment, persistently being bullied, chronic physical health problems, social isolation, divorce, unemployment, poverty, income inequality). Trends in more subjective indicators of environmental risk also were explored (e.g., unmet expectations, social isolation).

Collectively, the information from these diverse sources does not permit firm conclusions whether environmental risks have increased or decreased. Different data sources and substantial between-country differences yield inconsistent and contradictory findings. In addition, there is the problem of bidirectional causality, and many parameters do not show systematic trends but rather fluctuations (e.g., unemployment). The impact of technological, economic, and social-cultural developments on proximal determinants such as stressors and social supports also are difficult to detect. While conclusions are qualified by this lack of relevant statistics, what is clear is that there are no consistent or robust findings to suggest any noteworthy increase in risk factors.

3.3. Conclusion on direct and indirect evidence of incidence trends

Could an increase in first incidence have offset the expected treatment-driven decrease in prevalence? Probably not. However, the available information does not allow for a conclusive answer. Post-1980 incidence studies are few and information on trends in risk inconsistent. While some risks may have increased (e.g., socio-economic inequality, downward social mobility, terrorism, relative poverty), others may have decreased (e.g., child abuse, bullying, anchored poverty). Overall, the evidence does not support any major rise in incidence. Explanation 2 also can be ruled out as a strong explanation for the TPP.

4. Explanation 3: Is the efficacy of acute-phase treatment too small to matter?

Generally stated, if treatment efficacies are more modest than conventionally believed, then even with more people receiving purportedly gold standard interventions, any decrease in the prevalence would be smaller than expected. This could help to explain the TPP. There are, however, three related concerns about efficacy. Explanation 3 evaluates treatment efficacy with respect to the acute episode, whereas Explanation 4 (in the next section) evaluates efficacy with regard to maintaining treatment gains (e.g., preventing relapse and recurrence). Finally, we examine the extent that efficacious treatments are actually adequately implemented in real-world treatment settings, the transition of efficacy to effectiveness (see Explanation 5 to follow).

The leading Clinical Practice Guidelines on depression indicate that either medication or any of several empirically supported psychological treatments (hereafter denoted as psychotherapy, including a variety of evidence-based cognitive and behavioral treatments) are efficacious. Guidelines typically recommend prescribers to continue medication for at least 6–9 months after remission. Antidepressant medication remains the current standard of treatment and their presumed efficacy and relative safety have made them among the most widely prescribed medications in the world (although the evidence is increasing that psychotherapy may achieve more often sustained response (T. A. Furukawa et al., 2021)).

4.1. Acute-phase efficacy of medication

The most comprehensive recent meta-analysis, adjusted for publication bias and outliers, indicate modest acute-phase efficacy for medication relative to pill-placebo (Cipriani et al., 2018). They analyzed 474 trials (106,966 patients) that prescribed medication within their licensed range and covered 21 different medication. The primary outcomes were response rate and acceptability (defined as the inverse of discontinuation for any reason). The authors found that all of the medication examined are more efficacious than PLA at post-treatment (typically eight weeks after randomization) in terms of response rate. The summary Standardized Mean Difference (SMD) across all medication pooled was g = 0•30 (95% CI 0•26–0•34) and OR = 1.7), very similar to Turner's 2008 meta-analysis. The overall *response* rates (50% symptom reduction) were 37% in pill-placebo versus 50% for medication, with an OR = 1.67 (95%CI:1.60–1.74) and number needed to treat (NNT) 8.00 (95%CI:7.4 to 8.7) (Furukawa, personal communication). What NNT means is that one additional patient responds to medication for every eight patients treated on pill-placebo. The findings were robust across sensitivity analyses that eliminated outliers and studies that failed to prescribe within the recommended dosage ranges and were limited to low risk of bias. Unblinding remains a difficult to quantify validity threat as side effects of medication may reveal the identity of medication in trials using inert placebos and some work suggests that trials using 'active' placebos which mimic some of medication's side effects find smaller effect sizes (J. Moncrieff, 2003).

To understand the TPP, *remission* rates (HAMD-D < 7) are more important because response does not exclude the possibility that an individual still meets disorder criteria. Although not a meta-analysis, but rather a single study without controls for spontaneous remission, the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial is highly informative because of its enormous sample size (4000 patients across the full range of depression) and optimal ADM treatment including switching or augmenting medication every three months up to 12 months (Pigott, Leventhal, Alter, & Boren, 2010; Rush et al., 2006). A little over a third of the patients (37%) had remitted at three months on their initial medication, higher than the 23% spontaneous remission rate observed for untreated depression in six adult samples recruited from primary care settings (443 cases) (Whiteford et al., 2013). However, the official STAR*D rates have been challenged because they were based on the secondary outcome measure. According to the primary outcome measure (HRSD) a little over a quarter remitted within 3 months (Pigott et al., 2010).

4.2. Acute-phase efficacy of psychotherapy

Similar to medication, recent comprehensive meta-analyses of psychotherapy (P. Cuijpers et al., 2020; P. Cuijpers et al., 2021; P. Cuijpers, Karyotaki, Reijnders, & Ebert, 2019; Driessen et al., 2015) indicate that psychotherapy has some benefit relative to care-as-usual and pill-placebo but its benefit has been overestimated in earlier reviews (Barth et al., 2013; P. Cuijpers, 2017). Adjusted for publication bias, risk of bias, and excluding trials using wait-list controls, Cuijpers and colleagues found that the effect size of psychotherapy compared with control groups dropped by more than half, from g = 0.70 to g = 0.31 (95%CI 0.24–0.38)(P. Cuijpers et al., 2019), with a corresponding increase in NNT from 4.2 to 10.5. Wait-list controls do worse than other controls (the nocebo effect), likely because while awaiting treatment they do not do the things that they might otherwise do to feel better (T. A. Furukawa et al., 2014). Even keeping in mind that care-as-usual and pill-placebo likely outperform spontaneous remission, what this suggests is that, while psychotherapy is efficacious, that efficacy is not as great as the published literature would lead one to believe.

In the most recent comprehensive meta-analysis of 228 trials of psychotherapy, the overall response rate in psychotherapies at 2 (±1) months after baseline was 41% (95% CI: 38–43) versus 17% (15–20) for care-as-usual (P. Cuijpers et al., 2021). Limited to the 66 low-risk-of-bias trials and adjusted for publication bias response rate slightly dropped to 0.38. No significant differences between types of psychotherapy were found. A quarter to a third remitted after psychotherapy compared with 9–17% in control conditions. The NNTs for therapy versus care-as-usual were 5.3 (3.9–7.4) for response and 7.0 (3.4–20.8) for remission. Most sensitivity analyses supported the general findings. These meta-analytic findings suggest that psychotherapy is efficacious in absolute terms.

Cognitive behavior therapy (CBT) is by far the best studied type of

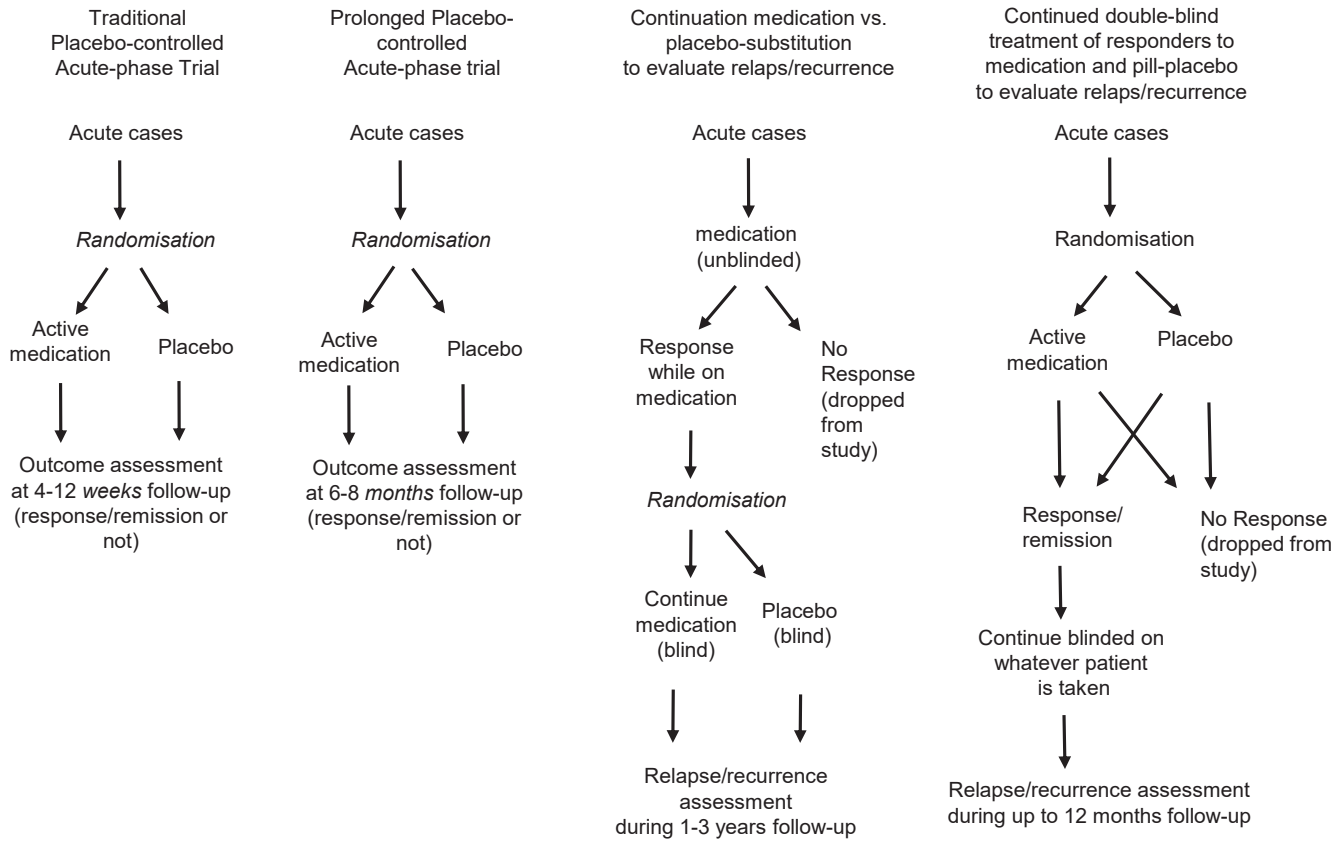

Fig. 1. Illustration of different RCT designs to evaluate acute-phase and continuation/maintenance-phase antidepressant medication.

psychotherapy but interpersonal psychotherapy, problem-solving therapy, and behavioral activation have comparable efficacies. These are exactly the types of psychotherapies that have held their own vis-à-vis medication in "silver bullet" trials in which both showed specificity relative to placebo (Elkin et al., 1989)(DeRubeis et al., 2005; Dimidjian et al., 2006; Elkin et al., 1989; Mynors-Wallis, Gath, Lloyd-Thomas, & Tomlinson, 1995).

4.3. ADM and PTX: compared and combined

According to the most recent and comprehensive meta-analysis comparing medication to psychotherapy, each alone and in combination, using 115 comparisons across 101 studies involving 11,910 patients (P. Cuijpers et al., 2020), differences in efficacy between medication and psychotherapy across response, remission, effect size are negligible and statistically non-significant. But psychotherapy outperforms medication in acceptability indexed as the inverse of attrition for any reason (RR = 1.17, 95% CI: 1.02–1.32). Combined treatment is typically more efficacious than either single modality alone. For instance, with respect to remission combined treatment was superior to medication (RR = 1.23, 1.09–1.39), and also more acceptable (RR = 1.23, 95% CI: 1.05–1.45), as well as superior to psychotherapy (RR = 1.22, 1.08–1.39) but not more acceptable. This translates into NNTs favoring combined treatment of about 10. The acute phase benefits of combined treatment relative to medication appear to hold during the maintenance phase as well (P. Cuijpers et al., 2014; T. A. Furukawa et al., 2021; Karyotaki et al., 2016).

The most recent network meta-analysis of the association between initial treatment and *sustained* response included 81 RCTs in 13,722 adult patients (T. A. Furukawa et al., 2021). Sustained response was defined as responding to the acute treatment and subsequently having no depressive relapse through the maintenance phase (mean duration: 42.2 ± 16.2 weeks, range 24–104 weeks). By design, acute phase treatment could be continued into the maintenance phase, switched to another treatment or followed by discretionary treatment. Psychotherapy kept patients well more often than medication, both when these treatments were continued into the maintenance phase (OR = 1.53, 95% CI: 1.00–2.35) and when they were followed by discretionary treatment (OR = 1.66, 95% CI: 1.13–2.44). The benefit of combined treatment relative to medication was even somewhat larger. Given the average sustained response rate of 29% on standard treatment, the advantages of psychotherapy or combined treatment over medication and care-as-usual translated into risk differences ranging from 12 to 16 percentage points.

4.4. Conclusions efficacy of acute-phase treatments

Medication and psychotherapy are the two most widely practiced depression treatments. Both appear to be comparably efficacious in terms of response, with adjusted ES's around 0.30 (NNT = 8); psychotherapy tends to be preferred by patients (McHugh, Whitton, Peckham, Welge, & Otto, 2013). Remission rates are substantially lower than response rates (about 15% percentage points). Combining the two is superior to either one alone. Neither psychotherapy nor medication works as well as the (older) literature suggest, as adjusting for biases reduces efficacy substantially (by a third to half). Heterogeneity is great as response-remission rates vary enormously across studies.

Given the biases and heterogeneity it is not surprising that there is significant disagreement about the clinical significance of treatments (Leucht, Hierl, Kissling, Dold, & Davis, 2012; J. Moncrieff & Kirsch, 2015). Here follows our view of the literature, realizing that the presented estimates should be interpreted with caution. In untreated samples from naturalistic studies, remission rates average 23% (Whiteford et al., 2013). For medication, the acute-phase remission rate typically lies between 25%–37%, for psychotherapy between 26%–43%, and for care-as-usual and pill-placebo around 17%. These rates suggest a

remission benefit of treatment between 2% and 23%. Remission rates double over the course of the subsequent year in a decelerating fashion, but also do so in untreated patients. In terms of sustained remission, defined as responding to the acute treatment and subsequently having no depressive relapse over the next 10–12 months, psychotherapy seems to outperform medication, with about 12 percentage points, a significant difference given the 29% sustained remission rate for standard treatment (T. A. Furukawa et al., 2021).

The rates indicate that Acute Phase efficacy exists, but to a much lesser degree than has been believed. The (sustained) remission rates imply that treatment should have some impact on prevalence. Nonetheless, on its own, Explanation 3 is not a strong candidate for helping to understand the TPP. However, Explanation 3 alone does not address the extent to which acute treatment effects endure over the long-term (Explanation 4 below), or the extent to which people have access to adequately implemented treatment in the real world (Explanation 5). As is becoming apparent, the intertwined nature of Explanations 3–5 may be especially important for understanding the TPP.

5. Explanation 4: Is the efficacy of interventions for preventing relapse and recurrence too small to matter?

It is well recognized that many patients who respond to *acute-phase* treatment do not maintain their clinical gains. More specifically, recent meta-analyses indicate that relapse/recurrence rates for responders to treatment are substantial. For medication responders who have been discontinued from medication, about a third relapse/recur within 33 weeks (Sim, Lau, Sim, Sum, & Baldessarini, 2016). For psychotherapy (CBT) responders, about a third relapse/recur at 1 year, and half at 2 years (C. L. Bockting, Hollon, Jarrett, Kuyken, & Dobson, 2015; C. L. H. Bockting et al., 2018; Vittengl, Clark, Dunn, & Jarrett, 2007). In addition, a significant number continue to struggle with residual symptoms (C. L. Bockting et al., 2015; S. M. Monroe & Harkness, 2011; J. Ormel, Oldehinkel, Brilman, & van den Brink, 1993). To prevent relapse/recurrence, interventions have been developed and evaluated.

5.1. Continuation and maintenance medication to prevent relapse and recurrence

Two types of RCTs have evaluated medication's efficacy for preventing relapse/recurrence: continuation of medication versus placebo-substitution, and double-blind extension of medication and pill-placebo responders (see Fig. 1). Both have advantages and limitations. Findings from the two kinds of designs converge, indicating medication about halves risk of relapse/recurrence relative to pill-placebo.

5.1.1. Continuation vs placebo discontinuation

In the discontinuation design, patients who respond to acute-phase medication are randomized to either continuation of medication or withdrawn onto pill-placebo and followed double blind. In the most recent meta-analysis of 72 trials (14,450 subjects) lasting up to 8 months, medication continuation was clearly more effective than placebo in preventing relapses (OR = 1.90, CI: 1.73–2.08; NNT = 4.4) (Sim et al., 2016). In 37 trials lasting up to 27 months, the advantage of medication continuation over pill-placebo was even greater (OR = 2.03, CI 1.80–2.28; NNT = 3.8) (with minor differences among specific drug types). These findings confirmed earlier meta-analyses reporting substantially higher relapse rates during the 6–12 months post-randomization period in placebo-substitution arms (~42%) compared to the medication continuation (~22%) (Geddes et al., 2003; Glue, Donovan, Kolluri, & Emir, 2010; Hansen et al., 2008; Kaymaz, van Os, Loonen, & Nolen, 2008). This design, though, is vulnerable to two types of bias (Fava et al., 2015). First, there is an increased risk of unblinding when switched to placebo substitution, which may reduce placebo effects (i.e., an expectation of a continued positive response). Second, withdrawal symptoms may occur and be misclassified as relapse/recurrence. Opinions on the magnitude of this misclassification bias differ greatly.

5.1.2. Double-blind extension in medication and pill-placebo responders

These types of studies begin as regular, double-blind, placebo-controlled acute-phase medication trials, but patients who respond to medication or to placebo then continue to receive the same blind treatment for 5–12 months. This design overcomes the expectation and misclassification concerns of placebo discontinuation trials but cannot entirely rule out unblinding. It also is susceptible to bias via differential mortality since patients at greater risk are more likely to be among the greater number of patients who respond to medication over pill-placebo. During 35 weeks follow-up of 901 patients (5 studies), on average 8% relapsed in the medication continuations arm compared to 23% in the pill-placebo arm (Zimmerman, Posternak, & Ruggero, 2007).

5.2. Preventive psychological treatments

Psychotherapy designed to reduce relapse/recurrence risk include mindfulness based cognitive therapy (MBCT), preventive cognitive therapy (PCT), and continuation of psychotherapy (CBT, IPT) in a much lower frequency. Their efficacy has usually been investigated in patients diagnosed with chronic-recurrent depression, who were at least in partial remission at randomization. Recent meta-analyses of these trials generally indicate efficacy compared to controls, but their magnitude depends on the nature of the control group (pill-placebo, care-as-usual, medication) (Biesheuvel-Leliefeld et al., 2015; P. Cuijpers et al., 2013; Kuyken et al., 2016; Sim et al., 2016; Vittengl et al., 2007). We focus on three 'silver bullet' meta-analyses.

Kuyken's 2016 meta-analysis of nine preventive MBCT trials examined long-term efficacy over a 60-week follow-up period for 1258 patients with at least two lifetime depressive episodes (Kuyken et al., 2016). A statistically significant advantage was found for MBCT compared to care-as-usual (9 studies, hazard ratio (0.69; 95%CI 0.58–0.82), any active treatment (5 studies, HR = 0.79; 95%CI, 0.64–0.97). This suggests that MBCT is efficacious for reducing risk of relapse/recurrence, particularly for those with more pronounced residual symptoms.

Biesheuvel-Leliefeld's 2015 meta-analysis of 25 trials that compared the effectiveness of preventive psychotherapy versus care-as-usual to reduce relapse/recurrence risk in 2055 patients in (partial) remission of depression. A total of 932 patients were randomized to an intervention condition: 529 received preventive CT, 142 IPT and 261 MBCT. The remaining 1123 patients were randomized to comparator conditions: 670 receiving care-as-usual and 453 receiving medication. Most follow-ups lasted 12–24 months. Preventive psychotherapy performed significantly better than care-as-usual (mostly routine clinical management) in reducing relapse/recurrence risk (RR = 0.64, 95%CI = 0.53–0.76, NNT = 5) and was also slightly more successful than medication continuation (RR = 0.83, 95%CI = 0.70–0.97, NNT = 13).

In contrast to the previous two meta-analyses, Cuijpers et al. (2013) meta-analysis of 9 studies targeted relapse/recurrence in 506 patients who responded to acute-phase CBT and were subsequently followed during 6–18 months. These patients were significantly less likely to relapse compared to patients who responded to, and then were *withdrawn* from acute-phase medication (OR = 2.61, 95% CI 1.58 to 4.31, NNT = 5) (P. Cuijpers et al., 2013). Interestingly, acute-phase CBT alone might be even more efficacious than continued medication (five studies), as there was a non-significant trend favoring acute-phase CBT over continued medication ($p < 0.1$; OR = 1.62, 95% CI 0.97 to 2.72; NNT = 10) (P. Cuijpers et al., 2013).

5.3. Conclusions on treatments for preventing relapse/recurrence

The meta-analyses suggest strong benefits for continued medication and preventive psychotherapy relative to controls for preventing

relapse/recurrence. However, methodological concerns remain, complicating interpretation, including misclassification of medication withdrawal symptoms, unblinding, heterogeneity of care-as-usual controls, and therapeutic allegiance problems. Two other issues also are relevant. Patients without response to acute-phase treatment typically were not eligible for relapse/recurrence prevention trials. Regarding acute-phase medication, STAR*D is an exception as it continued medication in all participants for a year. After remission, 37% relapsed and after response 64% (Rush et al., 2006). These rates are substantially higher than the relapse rates in continued medication arms of discontinuation trials, raising questions about the representativeness of the patients in the continued medication arms. Second, relapse/recurrence rates in patients who receive preventive treatment remain substantial, although estimates vary greatly: from ~25% during 6–12 months of continued medication (Sim et al., 2016) to 60% during two years of continued medication (Bockting Claudi, ten Doesschate, Spijker, Spinhoven, & Koeter, 2008), and 38% during 14 months of follow-up after MBCT (Kuyken et al., 2016) to 43% for psychotherapy plus continued medication during 2-year follow-up (C. L. H. Bockting et al., 2018).

All things considered, research on RCTs evaluating efficacy of treatments aiming to reduce relapse/recurrence risk reveals two faces: one optimistic, one more pessimistic. On balance, preventive interventions have sufficient efficacy to impact at the population level, although alone will not contribute much to understanding the TPP. But once again, the matter is dependent upon how widely and adequately these treatments are implemented in real-world care, to which we now turn.

6. Explanation 5: Do RCTs generalize to real-world settings?

RCT-based treatment efficacy may not generalize well to "real world" practice: the well-recognized distinction between ***efficacy*** as established in RCTs (i.e., under optimal conditions) versus ***effectiveness*** as realized in daily practice (i.e., under typical conditions) (Streiner, 2002). Large gaps in treatment quality between RCTs and real-world practice would result in less effective treatment, which would help to explain the TPP.

6.1. Differences between RCTs and real-world practice

The typical depressed patient in real-world practice may have two disadvantages compared to any RCT counterparts: poorer prognosis and less optimal treatment. In the past, RCTs of medication often excluded patients suffering from major medical or other psychiatric comorbidities that have poorer prognoses (van der Lem, de Wever, van der Wee, van Veen, Cuijpers, and Zitman, 2012a; van der Lem, van der Wee, van Veen, and Zitman, 2012b; Wells, 1999). More recent RCTs tend to recruit more representative samples, but patients with diagnosable disorders that require a different or immediate treatment or serious substance abuse typically still are excluded (DeRubeis et al., 2005; Stirman, DeRubeis, Crits-Christoph, & Rothman, 2005). Another difference is that treatment protocols in RCTs are explicitly specified, exactingly administered, and rigorously monitored. Therapists are trained extensively in the interventions, and treatment fidelity and patient improvement are closely monitored. Such care and detail often are not feasible in typical treatment venues. Another difference lies in the frequency of psychotherapy sessions. Most of the RCTs that have established the efficacy of CBT and behavioral activation begin with twice weekly sessions which does appear to enhance outcomes (and for IPT too) (Bruijniks et al., 2020). Unfortunately, this is rarely done in actual practice.

6.2. Treatment (quality) gaps

The gaps are considerable. The World Mental Health surveys in high-income countries found that 2468 (5.2%) adults met 12-month DSM-IV MDD criteria and 65% of those 2468 had a perceived need for treatment (Thornicroft et al., 2017). Of those, 78% made at least one visit to a service provider. Yet only 44% of those who received treatment obtained treatment that met minimal standards. Consequently, only 22% of all individuals needing treatment received minimally adequate treatment. Other sources indicate that early this century in high-income countries, a third of patients with *severe* disorders (i.e. suicide attempt, severe role impairment, or poor overall functioning) did not receive any care in the previous year, and many of the treatments that were provided did not meet clinical guidelines (Boerema et al., 2017; Demyttenaere et al., 2004; Fullerton, Busch, Normand, McGuire, & Epstein, 2011; Harris et al., 2014; Harris et al., 2015; Jorm et al., 2017; Marcus & Olfson, 2010b; M. Olfson, Blanco, Wang, Laje, and Correll, 2014a; M. Olfson, Kroenke, Wang, and Blanco, 2014b; Spiers et al., 2016; Wang et al., 2017; Young, Klap, Shoai, & Wells, 2008).

Given these sizable gaps, it is not surprising that the effectiveness of treatment in daily practice is lower than the efficacy results from RCTs. For instance, remission rates in Dutch routine practice are substantially lower than in meta-analyses for all treatment modalities, although differences were less explicit for medication than for psychotherapy (van der Lem, van der Wee, van Veen, and Zitman, 2012b).

The Texas Medication Algorithm Project (TEMAP) clearly illustrates the problem (Trivedi et al., 2004). It is considered best clinical practice to monitor outcomes on an ongoing basis and adjust the treatment regime accordingly. The other major principle is to "dose to remission". In TEMAP, psychiatrists in different secondary care settings were randomized to either algorithm-driven best clinical practices, or to continue to provide their routine care-as-usual. Over the ensuing year, patients treated by psychiatrists in the algorithm-driven condition exhibited twice as much change on observer-rated measures as patients in care-as-usual, and three times the change on self-report measures. That differences this dramatic could be obtained by simply adhering to best clinical practices was an indictment of the quality routine secondary care. Even worse, psychiatrists in the algorithm-driven clinics reverted to their usual treatment practices once the trial was over.

If the situation is dire in secondary care settings in the US, it is probably even worse in primary care. It is now the case in the US that primary care providers write 90% of the scripts for medication. Primary care providers are even less likely to follow best clinical practices for medication than are psychiatrist; adherence to measurement-based care is virtually nonexistent, and are far less likely to dose aggressively, or to switch, or augment as needed (most prescribe only SSRIs). Thus, while access to treatment has greatly increased in the US over the last quarter century, the quality of the treatment provided is anything but optimal. Most primary care providers likely settle for response rather than pushing for remission, largely because they lack the training. The STAR*D project previously described drew from both primary and secondary care settings, with the prescribers in each following the same algorithm-driven treatment guidelines. Under those conditions, the primary care providers in STAR*D generated outcomes comparable to the psychiatrists (Gaynes et al., 2008).

6.3. Conclusions on generalization to real-world settings

Substantial gaps exist in the dissemination and implementation of evidence-based treatments. Hence, RCT-based efficacy does not generalize well into real-world effectiveness, both for medication and psychotherapy. This compounds the observation that neither medication nor psychotherapy work as well as the (older) literature suggests. Once transported into the real world, characterized by tougher patients and less adequate implementation, the already modest treatment effects for both medication and psychotherapy diminish even further. Explanation 5 remains a strong candidate for helping to understand the TPP, in particular in combination with explanations 3 and 4.

7. Explanation 6: Most depressed people have few if any recurrences

While Explanation 5 targets the prevalence impact of the efficacy-effectiveness drop, Explanation 6 examines whether treatment availability is optimally targeted (i.e., does it reach the right people to affect prevalence). Clinical trials and epidemiological studies typically handle episodes of depression interchangeably across individuals irrespective of whether it is a first lifetime episode, a fifth recurrence, or a chronic depressive episode. It is assumed that it is the episode that matters, not potentially important between-person differences in liability to depression over the life course. It is becoming increasingly clear, however, that this silent assumption and research practice is unwarranted, which in turn can help to explain the TPP.

Half or more of the population of first-onset cases of depression will never experience a recurrence, with the majority of those who do probably never experiencing another (Eaton et al., 2008; S. M. Monroe, Anderson, & Harkness, 2019; S. M. Monroe & Harkness, 2011; J. Ormel et al., 1993; Rottenberg, Devendorf, Kashdan, & Disabato, 2018). Even the lifetime recurrence risk for first episodes of depression presenting to psychiatrists in secondary care may be lower than is widely believed, being of the order of 50% (Lee, 2003). Recurrence rates in primary care settings vary between 30 and 40% during >5-year follow-ups (van Weel-Baumgarten, Schers, van den Bosch, van den Hoogen, & Zitman, 2000). Succinctly stated, the preponderance of people who ever experience depression do not lead lives riven by repeated recurrences. In stark clinical contrast, a smaller, very important subset of the first-depressed will develop chronic-recurrent depression, with many recurrences or long illness periods throughout their lives (Keller et al., 1984; Keller & Boland, 1998; Lee, 2003). Approximately a quarter to a fifth of the lifetime prevalence suffers from chronic episodes (2+ years) (Angst, Gamma, Roessler, Ajdacic, & Klein, 2009; J. A. Murphy & Byrne, 2012).

There are at least four implications of this chronic-recurrent versus nonrecurrent subgroup distinction for the TPP. First and most generally, even with better treatments being more widely available, the impact on prevalence for nonrecurrent cases relative to chronic-recurrent cases may be more difficult to evaluate. This is because the only opportunity to do so for the large nonrecurrent group is by abbreviating the index, and only, lifetime episode (after which these people disappear from prevalence estimates). In contrast, the smaller chronic-recurrent group will provide many more opportunities for treatment to impact prevalence, from abbreviating multiple episodes through preventing relapses and recurrences. As next described, the net effect of these differing clinical circumstances would be for the larger group of nonrecurrent cases to attenuate or inflate any treatment-driven impact on prevalence attributable to the smaller group of chronic-recurrent cases.

Second, it follows that treatment can only impact prevalence if treatment is received, and if treatment abbreviates the natural course of a depressive episode. Yet people who have never been depressed previously may be less likely to seek treatment for their episode, and be more likely to recover spontaneously within a few months. Thus for a significant proportion of the 50% or more of non-recurrent first onset cases treatment effects may be irrelevant, and for those who seek treatment there may be only a time-limited 'in-episode' opportunity for detecting effects (i.e., they may only experience a slight reduction in episode duration). Add to this the likelihood that first onset cases may be more delayed in seeking treatment, establishing treatment effects would become even more challenging for this subject (S. M. Monroe & Harkness, 2011; Sareen et al., 2013; Spijker et al., 2002; Whiteford et al., 2013). In general, the availability of more and better treatments could be mostly irrelevant for the largest subset of the population of depressed persons, which would compromise investigative attempts to demonstrate treatment impact on depression's prevalence.

Third and optimistically, wider access to more effective treatments could robustly impact prevalence for the chronic-recurrent subset. By abbreviating multiple episodes, forestalling relapses, and reducing chronicity and recurrences, the opportunities for a treatment-driven drop in prevalence among these cases are many. In theory, the time spent depressed per person would decrease substantially, yielding in a clear prevalence drop *for this subgroup*.

Yet fourth and more pessimistically, people suffering from chronic-recurrent depression often have more severe episodes (i.e., greater number of symptoms, impaired functioning, higher suicide risk, longer duration and chronicity). Historically, too, they are at heightened risk of becoming treatment resistant. By all clinical accounts, the chronic-recurrent subgroup is a very challenging one to treat (Burcusa & Iacono, 2007; Lee, 2003). Yet RCTs rarely if ever distinguish between incident (i.e., first onset), chronic, and recurrent cases. The question remains open as to whether or not the advances in treatment efficacy and availability are as effective for, and available to, the chronic-recurrent individuals. If the modest efficacy of acute-phase medication and psychotherapy has largely been based on shortening episodes only for the many nonrecurrent cases and not on impacting the time depressed for the chronic-recurrent cases, the benefit of more and better acute-phase treatments would be minimal at the population level.

The previously mentioned successes of treatments developed specifically to prevent relapse and recurrence presuppose that these maintenance and preventive treatments are actually used in routine care and in guideline-consistent ways. Regarding maintenance medication, there is evidence that it is practiced, although it is less clear how adequately, especially in primary care. In addition, patient attrition is a serious problem in medication treatment. Regarding preventive psychotherapies, we could not locate robust evidence that indicates to what extent and how adequately they are provided in routine care. Our anecdotical experience suggests that they are still rather rare in routine clinical practice.

7.1. Conclusions on treatment impact on nonrecurrent versus chronic-recurrent depression

Overall, the impact of more and better treatments on depression's prevalence is likely to be limited for most incident cases – those who never become depressed again. This is not because treatments in theory do not necessarily help them, it is simply that treatments – if received – have just one time-limited opportunity to alleviate their episode. In clear contrast, the impact of more and better treatments should be more readily apparent for the smaller subset of cases who are at risk for having longer episodes and many recurrences or chronicity. Yet it remains unclear to what degree treatment advances affect the lifetime course for these chronic-recurrent subset, if these treatments are provided to them in routine clinical care, and whether or not their impact can eventually reveal a more robust treatment-induced prevalence drop.

All of the foregoing converges on the conclusion that any expected treatment-driven reduction in population prevalence of depression depends both on the opportunities for treatment to impact time depressed people, as well as on the efficacy of these treatments specifically for those who spend the most time in their lives depressed. Explanation 6 thereby remains a strong candidate for helping to understand the TPP.

8. Explanation 7: Can treatment be counterproductive?

Our analysis of the TPP has presumed that medication and psychotherapy only have positive effects, but it is possible that 'medication monotherapy' and 'low-fidelity psychotherapy' can also have counterproductive consequences. Two adverse consequences have been proposed: (1) Reduction of self-help activities and loss of agency (Meadows et al., 2019)and (2) oppositional perturbation (Andrews et al., 2011). Both possibilities merit consideration as possible contributors to the TPP, albeit each is considerably more speculative than the other explanations.

RCTs have not actively searched for these two adverse consequences, probably because there was no readily apparent reason to do so.

Medication-monotherapy and low-fidelity psychotherapy are uncommon in RCTs where treatment protocols are specified and carefully monitored, unlike treatment in real-world settings. In addition, medication-withdrawal studies may have missed the bigger picture of improved ultimate outcomes, due to misinterpreted withdrawal symptoms (Fava et al., 2015; Fava & Offidani, 2011) and too short follow-ups, as has been the case with antipsychotic medication withdrawal studies (Wunderink, Nieboer, Wiersma, Sytema, & Nienhuis, 2013).

8.1. Reduction of self-help activities and loss of agency

Meadows and colleagues hypothesized that medication treatment without behavioral management (mono-medication) is counterproductive if it reduces self-help activity and active coping (Meadows et al., 2019). The same might apply to low-fidelity-to-guideline psychotherapy. The argument is that depressed people often engage in strategies subsumed in self-help programs and psychological treatments, such as exercising, increasing pleasant activities, reducing stressful situations, and meditating. These self-help strategies can have multiple benefits, including direct effects on depressive symptoms and more indirect effects on their 'agency' and 'self-efficacy' for coping with depression and underlying problems. Successful experiences provide individuals with a greater sense of their own abilities, rather than feeling broken and dependent on others or medication to fix them (Haslam, 2016). If people on monotherapy avoid or reduce self-help activities, benefits of the particular treatment may be more than offset by the loss of agency. Another possible mechanism that could mediate the counterproductive effect of monotherapy is the following. Medications are thought to enhance neuronal plasticity, allowing environmental inputs to modify neuronal networks to better fine tune the individual to the outside world (Vetencourt et al., 2008). Recent observations in the visual cortex directly support this premise (Castren & Rantamaki, 2010). This suggests that medication should not be administered alone but should be combined with interventions to guide plasticity within the brain, by providing appropriate environmental input (e.g., behavioral activation, meditation). Although exact data on the frequency of mono-medication and low-fidelity psychotherapy are lacking, there is ample evidence of major treatment quality gaps.

8.2. Oppositional perturbation and symptom return

The "Oppositional Perturbation" hypothesis proposes to account for unintended and unwanted effects of medication on illness course, including symptom return after discontinuation, and a progressive loss of effectiveness (tachyphylaxis) across repeated ADM trials (Amsterdam, Lorenzo-Luaces, & DeRubeis, 2016; P. W. Andrews, Kornstein, Halberstadt, Gardner, & Neale, 2011; Fava & Offidani, 2011; J. Ormel, Bosker, Hollon, and Ruhe, 2020b). Both symptom return and loss of effectiveness represent worrying within-person findings that have been repeatedly reported. The explanatory hypothesis runs approximately as follows. Medications initially increase neurotransmitter levels in the synapse (up to four times anything ever seen in nature), causing the homeostatic regulatory mechanisms to respond by shutting down neurotransmitter synthesis pre-synaptically and reducing sensitivity post-synaptically. This reestablishes homeostatic regulation and these changes are maintained for the entire time that the patient remains on ADM. However, this ADM-driven perturbation may "bounce back" when medication is discontinued and overshoot the normal balance of monoamine storage and release, *increasing the risk for symptom return relative to patients who get better via spontaneous remission*. In effect, medication may "hijack" the homeostatic monoamine regulatory mechanisms, creating a persistent state of neuroregulatory perturbation. Importantly, *direct* evidence for oppositional perturbation is lacking, but intriguing *indirect* evidence is available. Specifically, the overshoot appears proportional to the extent that a given class of medications perturbs the underlying neurotransmitter systems and this corresponds directly to the likelihood of symptom return once medication is discontinued (P. W. Andrews et al., 2011; Fava & Offidani, 2011).

Although it is well established that medications impact underlying homeostatic regulatory processes, the consequences of this are not fully understood. They may shift from producing desired effects for so long as the medications are active (e.g., symptom suppression) to iatrogenic effects when they are discontinued (e.g., symptom return). It is noteworthy that CBT has an enduring effect, possibly shared by BA, that cuts risk for symptom return by more than half relative to medication following withdrawal (P. Cuijpers et al., 2013). The excess risk in the ADM group usually has been attributed to medication's beneficial effects ending with discontinuation, based on the plausible assumption that exposure to medication is benign and has no lingering negative effects. Oppositional perturbation provides a more provocative explanation: the apparent prophylactic effects of CBT and continued medication are deceptive, owing to enhanced risk of relapse/recurrence *due* to withdrawal of medication and the oppositional perturbation it unleashes.

Counterproductive effects of medication and psychotherapy provided without forms of behavioral management probably depend on provider characteristics, patient's personality, and contextual factors. Nowadays, more than 90% of SSRI prescriptions are written by GPs, who have fewer empowering strategies in their armamentarium or time to implement the ones they have. People in disadvantaged communities might thus be doubly disadvantaged because they tend to receive more mono-medication and less rigorous psychotherapy compared to the more comprehensive delivery of combined and empowering treatment in affluent circumstances (Meadows et al., 2019).

8.3. Conclusion on counterproductive effects

Unlike the earlier explanations for the TPP, only very limited indirect data are available regarding possible negative effects of present-day treatments for MDD. But in theory, counterproductive effects (if they exist) significantly expand opportunities for understanding the TPP, and our goal is to put forth as many credible explanations as possible. If operative, any beneficial impact of treatment would be diluted at the population level, and thus could help explain the TPP.

9. Concluding remarks on the treatment-prevalence paradox

We evaluated seven possible explanations for understanding the TPP (Overview in Table 3). Although no singular conclusion fully explains the TPP, there are clear differences in the viability and evidence of the different explanations. The explanations upon which all others depend concerns an increase in 1) the rate of "false positives" due to misdiagnosing distress as depression and 2) the actual incidence of depression, which in turn could offset any treatment-driven decrease in prevalence. Although neither can be entirely ruled out, there is little reason to conclude that prevalence has increased either in error or in fact. There are no strong signals or even hints of a pattern that supports such a premise. This opens the door for the five remaining explanations.

Compared to its absence, as indexed by spontaneous remission and natural history (Sareen et al., 2013; Wang et al., 2017; Whiteford et al., 2013), the absolute long-term effectiveness of treatment in real-world settings is disappointingly modest. The overestimated efficacy in controlled trials is largely attributable to a variety of biases (Explanations 3 and 4) and the modest RCT-based efficacy is strongly amplified by substantial gaps in the quality of implementation in real-world settings (Explanation 5). *Collectively*, these three explanations likely go a long way to help explain the TPP. It also is likely that a significant portion of the increment in treatment, to a large extent medication, has gone to patients not likely to experience chronic depression or multiple recurrences, although this is admittedly speculative (Explanation 6). (Note that this does not exclude that long-term users account for most units of medication).The final, more speculative, explanation suggests that existing treatments may have unrecognized counterproductive

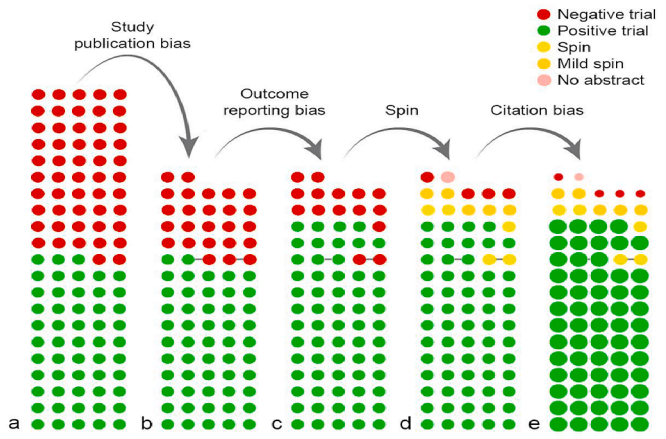

Fig. 2. The cumulative impact of reporting and citation biases on the evidence base for ADM.
The cumulative impact of reporting and citation biases on the evidence base for antidepressant medications. (a) displays the initial, complete cohort of trials, while (b) through (e) show the cumulative effect of biases. Each circle indicates a trial, while the color indicates the results or the presence of spin. Circles connected by a grey line indicate trials that were published together in a pooled publication. In (e), the size of the circle indicates the (relative) number of citations received by that category of studies (De Vries et al., 2018. With Permission).

effects. It is very important to investigate undertreatment of chronic-recurrent cases and iatrogenic effects of current treatments, especially medication as the increased treatment rate in recent decades largely consist of medications.

The image of overrated efficacy of treatments is due to not only publication and outcome reporting bias but to spin and citing bias as well (Fig. 2). Two examples: Regarding spin, out of 49 negative trials only 12 abstracts concluded that psychotherapy was not more effective than a control condition (De Vries et al., 2018). The remaining abstracts were either positive (73%) or mixed (19%) (e.g., concluding that the treatment was effective for one outcome but not another). Regarding citation bias: positive medication trials were cited three times as frequently as negative trials. In addition, negative trials with a positive or mixed abstract were cited more often than those with a negative abstract (59 and 87 citations, respectively v. 26). Positive psychotherapy trials were cited nearly twice as frequently as negative trials.

We envision major implications of our investigation of the TPP for research and clinical practice. Urgently needed research should address *long-term* outcomes, both in terms of (sustained) remission and recovery, as well as functioning, of treatment and non-treatment seekers. Another important question that emerges involves the nature of nonspecific treatment effects over time, substantial in the acute treatment phase, but unclear how enduring. Finally, it is essential to evaluate the Loss of Agency and Oppositional Perturbation hypotheses possible contribution to explaining the TPP. Notwithstanding these uncertainties, it seems likely that closing the quality gaps, especially for recurrent and chronic cases, will reduce prevalence.

From a public health perspective, investing in treatment and prevention for the high-risk subtype of depression is a realistic goal that has potential for decreasing depression's prevalence. Providing those at high risk for recurrence and chronicity with acute-phase psychotherapy with enduring effects or relaps/recurrence preventive psychotherapy or continuation and maintenance medication including behavioral management could have a positive impact on prevalence, especially when psychotherapy is made more readily available and disseminated in an adequate manner for suitable individuals. Resources could be freed up, targeting those for whom prevalence reduction can be most readily achieved. Reducing recurrence and chronicity risk will require addressing their determinants(S. M. Monroe, Anderson, & Harkness, 2019), a matter about which there is currently little theory or evidence (Buckman et al., 2018; Burcusa & Iacono, 2007). An approach worthy of investigation (J. Ormel, Cuijpers, Jorm, and Schoevers, 2020a) could be institutionally embedded structural prevention targeting parent's parenting and children's psychosocial skills.

Author contributions

JO: conceptualization, method (literature searches), writing: original draft.
SH and SM: conceptualization, writing: reviewing and editing.
RK and PC: Writing: reviewing + editing.
All authors approved the final manuscript.

Role of funding sources

No involvement of any funding source. All work was done as regular university employee.

Declaration of Competing Interest

All authors declare to have no conflict of interest regarding and no disclosure to report.

References

Amsterdam, J. D., Lorenzo-Luaces, L., & DeRubeis, R. J. (2016). Step-wise loss of antidepressant effectiveness with repeated antidepressant trials in bipolar II depression. *Bipolar Disorders, 18*(7), 563–570. https://doi.org/10.1111/bdi.12442
Andrews, G., & Peters, L. (1998). The psychometric properties of the composite international diagnostic interview. *Social Psychiatry and Psychiatric Epidemiology, 33*(2), 80–88.
Andrews, P. W., Kornstein, S. G., Halberstadt, L. J., Gardner, C. O., & Neale, M. C. (2011). Blue again: Perturbational effects of antidepressants suggest monoaminergic homeostasis in major depression. *Frontiers in Psychology, 2*, 159. https://doi.org/10.3389/fpsyg.2011.00159
Angst, J., Gamma, A., Roessler, W., Ajdacic, V., & Klein, D. N. (2009). Long-term depression versus episodic major depression: Results from the prospective Zurich study of a community sample. *Journal of Affective Disorders, 115*(1–2), 112–121. https://doi.org/10.1016/j.jad.2008.09.023
Barth, J., Munder, T., Gerger, H., Nueesch, E., Trelle, S., Znoj, H., & Cuijpers, P. (2013). Comparative efficacy of seven psychotherapeutic interventions for patients with depression: A network meta-analysis. *PLoS Medicine, 10*(5), Article e1001454. https://doi.org/10.1371/journal.pmed.1001454
Baxter, A. J., Scott, K. M., Ferrari, A. J., Norman, R. E., Vos, T., & Whiteford, H. A. (2014). Challenging the myth of an "epidemic" of common mental disorders: Trends in the global prevalence of anxiety and depression between 1990 and 2010. *Depression and Anxiety, 31*(6), 506–516. https://doi.org/10.1002/da.22230
Biesheuvel-Leliefeld, K. E. M., Kok, G. D., Bockting, C. L. H., Cuijpers, P., Hollon, S. D., van Marwijk, H. W. J., & Smit, F. (2015). Effectiveness of psychological interventions in preventing recurrence of depressive disorder: Meta-analysis and meta-regression. *Journal of Affective Disorders, 174*, 400–410. https://doi.org/10.1016/j.jad.2014.12.016
Bijl, R. V., de Graaf, R., Ravelli, A., Smit, F., & Vollebergh, W. A. (2002). Gender and age-specific first incidence of DSM-III-R psychiatric disorders in the general population. Results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology, 37*(8), 372–379.
Bockting Claudi, L. H. C. L., ten Doesschate, M. C., Spijker, J., Spinhoven, P., & Koeter, M. W. J. (2008). Continuation and maintenance use of antidepressants in recurrent depression. *Psychotherapy and Psychosomatics, 77*(1), 17–26.
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. https://doi.org/10.1016/j.cpr.2015.02.003
Bockting, C. L. H., Klein, N. S., Elgersma, H. J., van Rijsbergen, G. D., Slofstra, C., Ormel, J., & Burger, H. (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. *Lancet Psychiatry, 5*(5), 401–410. https://doi.org/10.1016/S2215-0366(18)30100-7
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. https://doi.org/10.1186/s12888-017-1531-8
Bongers, F. (2009). General practice 1987 and 2001: changes in morbidity and interventions: Findings from the two Dutch national surveys of general practice. Utrecht: NIVEL.

Boutron, I., Altman, D. G., Hopewell, S., Vera-Badillo, F., Tannock, I., & Ravaud, P. (2014). Impact of spin in the abstracts of articles reporting results of randomized controlled trials in the field of cancer: The SPIIN randomized controlled trial. *Journal of Clinical Oncology, 32*(36), 4120–U346. https://doi.org/10.1200/JCO.2014.56.7503

Bromet, E., Andrade, L. H., Hwang, I., Sampson, N. A., Alonso, J., de Girolamo, G., & Kessler, R. C. (2011). Cross-national epidemiology of DSM-IV major depressive episode. *BMC Medicine, 9*. https://doi.org/10.1186/1741-7015-9-90; 10.1186/1741-7015-9-90, 90–7015-9-90.

Brugha, T. S., Bebbington, P. E., Singleton, N., Melzer, D., Jenkins, R., Lewis, G., & Meltzer, H. (2004). Trends in service use and treatment for mental disorders in adults throughout Great Britain. *British Journal of Psychiatry, 185*, 378–384.

Brugha, T. S., Jenkins, R., Taub, N., Meltzer, H., & Bebbington, P. E. (2001). A general population comparison of the composite international diagnostic interview (CIDI) and the schedules for clinical assessment in neuropsychiatry (SCAN). *Psychological Medicine, 31*(6), 1001–1013.

Bruijniks, S. J. E., Lemmens, L. H. J. M., Hollon, S. D., Peeters, F. P. M. L., Cuijpers, P., Arntz, A., & Huibers, M. J. H. (2020). The effects of once- versus twice-weekly sessions on psychotherapy outcomes in depressed patients. *British Journal of Psychiatry, 216*(4), 222–230. https://doi.org/10.1192/bjp.2019.265

Buckman, J. E. J., Underwood, A., Clarke, K., Saunders, R., Hollon, S. D., Fearon, P., & Pilling, S. (2018). Risk factors for relapse and recurrence of depression in adults and how they operate: A four-phase systematic review and meta-synthesis. *Clinical Psychology Review, 64*, 13–38. https://doi.org/10.1016/j.cpr.2018.07.005

Burcusa, S. L., & Iacono, W. G. (2007). Risk for recurrence in depression. *Clinical Psychology Review, 27*(8), 959–985. https://doi.org/10.1016/j.cpr.2007.02.005

Castren, E., & Rantamaki, T. (2010). The role of BDNF and its receptors in depression and antidepressant drug action: Reactivation of developmental plasticity. *Developmental Neurobiology, 70*(5), 289–297. https://doi.org/10.1002/dneu.20758

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. https://doi.org/10.1016/S0140-6736(17)32802-7

Collaborators, G. (2018). *Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: A systematic analysis for the Global Burden of Disease Study 2017*.

Compton, W. M., Conway, K. P., Stinson, F. S., & Grant, B. F. (2006). Changes in the prevalence of major depression and comorbid substance use disorders in the United States between 1991-1992 and 2001-2002. *American Journal of Psychiatry, 163*(12), 2141–2147. https://doi.org/10.1176/appi.ajp.163.12.2141

Cooney, G. M., Dwan, K., Greig, C. A., Lawlor, D. A., Rimer, J., Waugh, F. R., & Mead, G. E. (2013). Exercise for depression. *Cochrane Database of Systematic Reviews, 9*, CD004366. https://doi.org/10.1002/14651858.CD004366.pub6

Cuijpers, P., Hollon, S. D., van Straten, A., Bockting, C., Berking, M., & Andersson, G. (2013). Does cognitive behaviour therapy have an enduring effect that is superior to keeping patients on continuation pharmacotherapy? A meta-analysis. *Bmj Open, 3*(4), Article e002542. https://doi.org/10.1136/bmjopen-2012-002542

Cuijpers, P., Karyotaki, E., Ciharova, M., Miguel, C., Noma, H., & Furukawa, T. A. (2021). The effects of psychotherapies for depression on response, remission, reliable change, and deterioration: A meta-analysis. *Acta Psychiatrica Scandinavica, 144*(3), 288–299.

Cuijpers, P., Noma, H., Karyotaki, E., Vinkers, C. H., Cipriani, A., & Furukawa, T. A. (2020). A network meta-analysis of the effects of psychotherapies, pharmacotherapies and their combination in the treatment of adult depression. *World Psychiatry, 19*(1), 92–107.

Cuijpers, P., Karyotaki, E., Reijnders, M., & Ebert, D. D. (2019). Was eysenck right after all? A reassessment of the effects of psychotherapy for adult depression. *Epidemiology and Psychiatric Sciences, 28*(1), 21–30. https://doi.org/10.1017/S2045796018000057

Cuijpers, P. (2017). Four decades of outcome research on psychotherapies for adult depression: An overview of a series of meta-analyses. *Canadian Psychology-Psychologie Canadienne, 58*(1), 7–19. https://doi.org/10.1037/cap0000096

Cuijpers, P., Hollon, S. D., van Straten, A., Bockting, C., Berking, M., & Andersson, G. (2013). Does cognitive behaviour therapy have an enduring effect that is superior to keeping patients on continuation pharmacotherapy?A meta-analysis. *Bmj Open, 3*(4), Article e002542. https://doi.org/10.1136/bmjopen-2012-002542

Cuijpers, P., Sijbrandij, M., Koole, S. L., Andersson, G., Beekman, A. T., Reynolds, C. F., & III.. (2014). Adding psychotherapy to antidepressant medication in depression and anxiety disorders: A meta-analysis. *World Psychiatry, 13*(1), 56–67. https://doi.org/10.1002/wps.20089

Daly, M. (2021). Prevalence of depression among adolescents in the US from 2009 to 2019: Analysis of trends by sex, Race/Ethnicity, and income. *Journal of Adolescent Health*. In press.

de Graaf, R., ten Have, M., Tuithof, M., & van Dorsselaer, S. (2013). First-incidence of DSM-IV mood, anxiety and substance use disorders and its determinants: Results from the Netherlands Mental Health Survey and Incidence Study-2. *Journal of Affective Disorders, 149*(1–3), 100–107. https://doi.org/10.1016/j.jad.2013.01.009

de Graaf, R., Ten Have, M., van Gool, C., & van Dorsselaer, S. (2012). Prevalence of mental disorders, and trends from 1996 to 2009. Results from NEMESIS-2. [Prevalentie van psychische aandoeningen en trends van 1996 tot 2009; resultaten van NEMESIS-2]. *Tijdschrift Voor Psychiatrie, 54*(1), 27–38.

De Vries, Y., Roest, A., de Jonge, P., Cuijpers, P., Munafò, M., & Bastiaansen, J. (2018). The cumulative effect of reporting and citation biases on the apparent efficacy of treatments: The case of depression. *Psychological Medicine, 48*(15), 2453–2455.

Demyttenaere, K., Bruffaerts, R., Posada-Villa, J., Gasquet, I., Kovess, V., Lepine, J. P., & Chatterji, S. (2004). Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization world mental health surveys. *Jama-Journal of the American Medical Association, 291*(21), 2581–2590.

DeRubeis, R. J., Hollon, S. D., Amsterdam, J. D., Shelton, R. C., Young, P. R., Salomon, R. M., & Gallop, R. (2005). Cognitive therapy vs medications in the treatment of moderate to severe depression. *Archives of General Psychiatry, 62*(4), 409–416.

Dimidjian, S., Hollon, S. D., Dobson, K. S., Schmaling, K. B., Kohlenberg, R. J., Addis, M. E., & Jacobson, N. S. (2006). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of Consulting and Clinical Psychology, 74*(4), 658–670. https://doi.org/10.1037/0022-006X.74.4.658

Donohue, J. M., Cevasco, M., & Rosenthal, M. B. (2007). A decade of direct-to-consumer advertising of prescription drugs. *New England Journal of Medicine, 357*(7), 673–681.

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), Article e0137864. https://doi.org/10.1371/journal.pone.0137864

Eaton, W. W., Kramer, M., Anthony, J. C., Dryman, A., Shapiro, S., & Locke, B. Z. (1989). The incidence of specific dis Dsm-Iii mental-disorders - Data from the nimh epidemiologic catchment-area program. *Acta Psychiatrica Scandinavica, 79*(2), 163–178. https://doi.org/10.1111/j.1600-0447.1989.tb08584.x

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. https://doi.org/10.1001/archpsyc.65.5.513

Elkin, I., Shea, M. T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., & Parloff, M. B. (1989). National Institute of Mental Health treatment of depression collaborative research program: General effectiveness of treatments. *Archives of General Psychiatry, 46*, 971–982.

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. https://doi.org/10.1159/000370338

Fava, G. A., & Offidani, E. (2011). The mechanisms of tolerance in antidepressant action. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 35*(7), 1593–1602. https://doi.org/10.1016/j.pnpbp.2010.07.026

Ferrari, A. J., Somerville, A. J., Baxter, A. J., Norman, R., Patten, S. B., Vos, T., & Whiteford, H. A. (2013). Global variation in the prevalence and incidence of major depressive disorder: A systematic review of the epidemiological literature. *Psychological Medicine, 43*(3), 471–481. https://doi.org/10.1017/S0033291712001511

Frank, E., Prien, R. F., Jarrett, R. B., Keller, M. B., Kupfer, D. J., Lavori, P. W., & Weissman, M. M. (1991). Conceptualization and rationale for consensus definitions of terms in major depressive disorder. Remission, recovery, relapse, and recurrence. *Archives of General Psychiatry, 48*(9), 851–855.

Fullerton, C. A., Busch, A. B., Normand, S. T., McGuire, T. G., & Epstein, A. M. (2011). Ten-year trends in quality of care and spending for depression 1996 through 2005. *Archives of General Psychiatry, 68*(12), 1218–1226. https://doi.org/10.1001/archgenpsychiatry.2011.146

Furer, J. W. (2001). *Protoprofessionalisering: een empirisch onderzoek naar de validiteit van protoprofessionalisering en naar haar verbanden met gezond(heids)gedrag, ziektegedrag en gezondheidstoestand*. Retrieved from http://hdl.handle.net/2066/146792.

Furukawa, T. A., Noma, H., Caldwell, D. M., Honyashiki, M., Shinohara, K., Imai, H., & Churchill, R. (2014). Waiting list may be a nocebo condition in psychotherapy trials: A contribution from network meta-analysis. *Acta Psychiatrica Scandinavica, 130*(3), 181–192. https://doi.org/10.1111/acps.12275

Furukawa, T. A., Shinohara, K., Sahker, E., Karyotaki, E., Miguel, C., Ciharova, M., & Cuijpers, P. (2021). Initial treatment choices to achieve sustained response in major depression: A systematic review and network meta-analysis. *World Psychiatry, 20*(3), 387–396. https://doi.org/10.1002/wps.20906

Gaynes, B. N., Rush, A. J., Trivedi, M. H., Wisniewski, S. R., Balasubramani, G., McGrath, P. J., & Yates, W. R. (2008). Primary versus specialty care outcomes for depressed outpatients managed with measurement-based care: Results from STAR* D. *Journal of General Internal Medicine, 23*(5), 551–560.

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.

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. https://doi.org/10.3109/00048671003705441

Grant, B. F., Goldstein, R. B., Chou, S. P., Huang, B., Stinson, F. S., Dawson, D. A., & Compton, W. M. (2009). Sociodemographic and psychopathologic predictors of first incidence of DSM-IV substance use, mood and anxiety disorders: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. *Molecular Psychiatry, 14*(11), 1051–1066. https://doi.org/10.1038/mp.2008.41

Hagnell, O., Lanke, J., Rorsman, B., & Öjesjö, L. (1982). Are we entering an age of melancholy? : Depressive illnessess in a prospective epidemiological study over 25 years: The lundby study, Sweden. *Psychological Medicine, 12*, 279–289.

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. https://doi.org/10.1176/appi.ps.59.10.1121

Haro, J. M., Rbabzadeh-Bouchez, S., Brugha, T. S., de Girolamo, G., Guyer, M. E., Jin, R., & Kessler, R. C. (2006). Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health surveys. *International Journal of Methods in Psychiatric Research, 15*(1049-8931; 4), 167–180.

Harris, M. G., Diminic, S., Burgess, P. M., Carstensen, G., Stewart, G., Pirkis, J., & Whiteford, H. A. (2014). Understanding service demand for mental health among australians aged 16 to 64 years according to their possible need for treatment. *Australian and New Zealand Journal of Psychiatry, 48*(9), 838–851. https://doi.org/10.1177/0004867414531459

Harris, M. G., Hobbs, M. J., Burgess, P. M., Pirkis, J. E., Diminic, S., Siskind, D. J., & Whiteford, H. A. (2015). Frequency and quality of mental health treatment for affective and anxiety disorders among australian adults. *Medical Journal of Australia, 202*(4), 184–189. https://doi.org/10.5694/mja14.00297

Haslam, N. (2016). Concept creep: Psychology's expanding concepts of harm and pathology. *Psychological Inquiry, 27*(1), 1–17. https://doi.org/10.1080/1047840X.2016.1082418

Hollon, S. D., Areán, P. A., Craske, M. G., Crawford, K. A., Kivlahan, D. R., Magnavita, J. J., & Bufka, L. F. (2014). Development of clinical practice guidelines. *Annual Review of Clinical Psychology, 10*, 213–241.

Horwitz, A. V. (2007). Distinguishing distress from disorder as psychological outcomes of stressful social arrangements. *Health, 11*(3), 273–289.

Jeronimus, B. F., Kotov, R., Riese, H., & Ormel, J. (2016). Neuroticism's prospective association with mental disorders halves after adjustment for baseline symptoms and psychiatric history, but the adjusted association hardly decays with time: A meta-analysis on 59 longitudinal/prospective studies with 443 313 participants. *Psychological Medicine, 46*(14), 2883–2906. S0033291716001653 [pii].

Jorm, A. F. (2014). Why hasn't the mental health of Australians improved? The need for a national prevention strategy. *Australian and New Zealand Journal of Psychiatry, 48*(9), 795–801. https://doi.org/10.1177/0004867414546387

Jorm, A. F., Patten, S. B., Brugha, T. S., & Mojtabai, R. (2017). Has increased provision of treatment reduced the prevalence of common mental disorders? Review of the evidence from four countries. *World Psychiatry, 16*(1), 90–99. https://doi.org/10.1002/wps.20388

Karyotaki, E., Smit, Y., Henningsen, K. H., Huibers, M., Robays, J., de Beurs, D., & Cuijpers, P. (2016). Combining pharmacotherapy and psychotherapy or monotherapy for major depression? A meta-analysis on the long-term effects. *Journal of Affective Disorders, 194*, 144–152.

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-+.

Keller, M. B., & Boland, R. J. (1998). Implications of failing to achieve successful long-term maintenance treatment of recurrent unipolar major depression. *Biological Psychiatry, 44*(5), 348–360. https://doi.org/10.1016/S0006-3223(98)00110-3

Keller, M. B., Klerman, G. L., Lavori, P. W., Coryell, W., Endicott, J., & Taylor, J. (1984). Long-term outcome of episodes of major depression: Clinical and public health significance. *Jama-Journal of the American Medical Association, 2252*, 788–792.

Kessler, R. C., Avenevoli, S., Green, J., Gruber, M. J., Guyer, M., He, Y., & Zaslavsky, A. M. (2009). National comorbidity survey replication adolescent supplement (NCS-A): III. Concordance of DSM-IV/CIDI diagnoses with clinical reassessments. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*(4), 386–399. https://doi.org/10.1097/CHI.0b013e31819a1cbc; 10.1097/CHI.0b013e31819a1cbc

Kessler, R. C., Demler, O., Frank, R. G., Olfson, M., Pincus, H. A., Walters, E. E., & Zaslavsky, A. M. (2005). Prevalence and treatment of mental disorders, 1990 to 2003. *The New England Journal of Medicine, 352*(24), 2515–2523. https://doi.org/10.1056/NEJMsa043266

Kirsch, I. (2014). Antidepressants and the Placebo Effect. *Zeitschrift Fur Psychologie-Journal of Psychology, 222*(3), 128–134. https://doi.org/10.1027/2151-2604/a000176

Klerman, G. L. (1988). The current age of youthful melancholia. Evidence for increase in depression among adolescents and young adults. *British Journal of Psychiatry, 152*, 4–14.

Kotov, R., Gamez, W., Schmidt, F., & Watson, D. (2010). Linking "big" personality traits to anxiety, depressive, and substance use disorders: A meta-analysis. *Psychological Bulletin, 136*(5), 768–821.

Kuyken, W., Warren, F. C., Taylor, R. S., Whalley, B., Crane, C., Bondolfi, G., & Dalgleish, T. (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. https://doi.org/10.1001/jamapsychiatry.2016.0076

Lee, A. S. (2003). Better outcomes for depressive disorders? *Psychological Medicine, 33*(5), 769–774.

Lehtinen, V., Sohlman, B., Nummelin, T., Salomaa, M., Ayuso-Mateos, J. L., & Dowrick, C. (2005). The estimated incidence of depressive disorder and its determinants in the Finnish ODIN sample. *Social Psychiatry and Psychiatric Epidemiology, 40*(10), 778–784. https://doi.org/10.1007/s00127-005-0956-4

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. https://doi.org/10.1192/bjp.bp.111.096594

Malouff, J. M., Thorsteinsson, E. B., & Schutte, N. S. (2005). The relationship between the five-factor model of personality and symptoms of clinical disorders: A meta-analysis. *Journal of Psychopathology and Behavioral Assessment, 27*(2), 101–114.

Marcus, S. C., & Olfson, M. (2010a). National Trends in the treatment for depression from 1998 to 2007. *Archives of General Psychiatry, 67*(12), 1265–1273. https://doi.org/10.1001/archgenpsychiatry.2010.151

Marcus, S. C., & Olfson, M. (2010b). National Trends in the treatment for depression from 1998 to 2007. *Archives of General Psychiatry, 67*(12), 1265–1273. https://doi.org/10.1001/archgenpsychiatry.2010.151

Mark, T. L., Levit, K. R., Vandivort-Warren, R., Buck, J. A., & Coffey, R. M. (2011). Changes in US spending on mental health and substance abuse treatment, 1986–2005. *And Implications For Policy. Health Affairs, 30*(2), 284–292. https://doi.org/10.1377/hlthaff.2010.0765

McHugh, R. K., Whitton, S. W., Peckham, A. D., Welge, J. A., & Otto, M. W. (2013). Patient preference for psychological vs pharmacologic treatment of psychiatric disorders: A meta-analytic review. *The Journal of Clinical Psychiatry, 74*(6), 595–602. https://doi.org/10.4088/JCP.12r07757 [doi]

Meadows, G. N., Prodan, A., Patten, S., Shawyer, F., Francis, S., Enticott, J., & Kakuma, R. (2019). Resolving the paradox of increased mental health expenditure and stable prevalence. *Australian & New Zealand Journal of Psychiatry, 0004867419857821*.

Moffitt, T. E., Caspi, A., Taylor, A., Kokaua, J., Milne, B. J., Polanczyk, G., & Poulton, R. (2010). How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. *Psychological Medicine, 40*(6), 899–909. https://doi.org/10.1017/S0033291709991036

Meertens, V., Scheepers, P., & Tax, B. (2003). Depressive symptoms in the Netherlands 1975-1996: a theoretical framework and an empirical analysis of socio-demographic characteristics, gender differences and changes over time. *Sociology of Health & Illness, 25*(2), 208–231. https://doi.org/10.1111/1467-9566.00332

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., & Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences of the United States of America, 108*(7), 2693–2698. https://doi.org/10.1073/pnas.1010076108

Mojtabai, R. (2007). Americans' attitudes toward mental health treatment seeking: 1990–2003. *Psychiatric Services, 58*(5), 642–651. https://doi.org/10.1176/appi.ps.58.5.642

Mojtabai, R., & Jorm, A. F. (2015). Trends in psychological distress, depressive episodes and mental health treatment-seeking in the United States: 2001-2012. *Journal of Affective Disorders, 174*, 556–561. https://doi.org/10.1016/j.jad.2014.12.039

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. https://doi.org/10.4088/JCP.13m08443

Mojtabai, R., Olfson, M., & Han, B. (2016). National trends in the prevalence and treatment of depression in adolescents and young adults. *Pediatrics, 138*(6), Article e20161878. https://doi.org/10.1542/peds.2016-1878

Moncrieff, J. (2003). A comparison of antidepressant trials using active and inert placebos. *International Journal of Methods in Psychiatric Research, 12*(3), 117–127. https://doi.org/10.1002/mpr.148

Moncrieff, J. (2018). Against the stream: Antidepressants are not antidepressants - an alternative approach to drug action and implications for the use of antidepressants. *Bjpsych Bulletin, 42*(1), 42–44. https://doi.org/10.1192/bjb.2017.11

Moncrieff, J., & Kirsch, I. (2015). Empirically derived criteria cast doubt on the clinical significance of antidepressant-placebo differences. *Contemporary Clinical Trials, 43*, 60–62. https://doi.org/10.1016/j.cct.2015.05.005

Moncrieff, J., Wessely, S., & Hardy, R. (1998). Meta-analysis of trials comparing antidepressants with active placebos. *British Journal of Psychiatry, 172*, 227–231.

Monroe, S. M., Anderson, S. F., & Harkness, K. L. (2019). Life stress and major depression: The mysteries of recurrences. *Psychological Review, 126*(6), 791–816.

Monroe, S. M., & Harkness, K. L. (2011). Recurrence in major depression: A conceptual analysis. *Psychological Review, 118*(4), 655–674. https://doi.org/10.1037/a0025190

Mroczek, D., & Spiro, A. (2003). Modeling intraindividual change in personality traits: Findings from the normative aging study. *Journals of Gerontology Series B-Psychological Sciences and Social Sciences, 58*(3), P153–P165.

Mulder, R. T. (2008). An epidemic of depression or the medicalization of distress? *Perspectives in Biology and Medicine, 51*(2), 238–250.

Murphy, J. M., Laird, N. M., Monson, R. R., Sobol, A. M., & Leighton, A. H. (2000). A 40-year perspective on the prevalence of depression. The Stirling County study. *Archives of General Psychiatry, 57*, 209–215.

Murphy, J. A., & Byrne, G. J. (2012). Prevalence and correlates of the proposed DSM-5 diagnosis of chronic depressive disorder. *Journal of Affective Disorders, 139*(2), 172–180. https://doi.org/10.1016/j.jad.2012.01.033

Murray, C. J. L., Barber, R. M., Foreman, K. J., Ozgoren, A. A., Abd-Allah, F., Abera, S. F., & Collaborator, H. A. L. E. (2015). Global, regional, and national disability-adjusted life years (DALYs) for 306 diseases and injuries and healthy life expectancy (HALE) for 188 countries, 1990–2013: Quantifying the epidemiological transition. *Lancet, 386*(10009), 2145–2191. https://doi.org/10.1016/S0140-6736(15)61340-X

Mynors-Wallis, L. M., Gath, D. H., Lloyd-Thomas, A. R., & Tomlinson, D. (1995). Randomised controlled trial comparing problem solving treatment with amitriptyline and placebo for major depression in primary care. *British Medical Journal, 310*, 441–445.

National Heart, Lung, and Blood Institute. (2006). *Incidence and prevalence: 2006 chart book on cardiovascular and lung diseases*. Bethesda, MD: National Institutes of Health.

Newman, S. C., & Bland, R. C. (1998). Incidence of mental disorders in Edmonton: estimates of rates and methodological issues. *Journal of Psychiatric Research, 32*(5), 273–282. https://doi.org/10.1016/S0022-3956(98)00011-9

Niaounakis, T. K. (2013). *Productiviteitstrends in de geestelijke gezondheidszorg: Een empirisch onderzoek naar het effect van regulering op de productiviteitsontwikkeling.* Delft: IPSE Studies. tussen 1982 en 2010.

Nuijen, J., Bon-Marten, M. V., Graaf, R. D., Have, M. T., Poel, A., Beurs, D. D., & Voorrips, L. (2018). Zicht op depressie: de aandoening, preventie en zorg. In *Themarapportage van de Staat van Volksgezondheid en Zorg*.

Olfson, M., & Marcus, S. C. (2010). National trends in outpatient psychotherapy. *American Journal of Psychiatry, 167*(12), 1456–1463. https://doi.org/10.1176/appi.ajp.2010.10040570

Olfson, M., Marcus, S., Druss, B., & Pincus, H. (2002). National trends in the use of outpatient psychotherapy. *American Journal of Psychiatry, 159*(11), 1914–1920. https://doi.org/10.1176/appi.ajp.159.11.1914

Olfson, M., Marcus, S., Wan, G., & Geissler, E. (2004). National trends in the outpatient treatment of anxiety disorders. *Journal of Clinical Psychiatry, 65*(9), 1166–1173. https://doi.org/10.4088/JCP.v65n0903

Olfson, M., Blanco, C., Wang, S., Laje, G., & Correll, C. U. (2014). National Trends in the mental health Care of Children, adolescents, and adults by office-based physicians. *JAMA Psychiatry, 71*(1), 81–90. https://doi.org/10.1001/jamapsychiatry.2013.3074

Olfson, M., Kroenke, K., Wang, S., & Blanco, C. (2014). Trends in office-based mental health care provided by psychiatrists and primary care physicians. *Journal of Clinical Psychiatry, 75*(3), 247–253. https://doi.org/10.4088/JCP.13m08834

Ormel, J., Cuijpers, P., Jorm, A., & Schoevers, R. (2020). What is needed to eradicate the depression epidemic, and why. *Mental Health and Prevention, 17*. https://doi.org/10.1016/j.mhp.2019.200177. in press.

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 pschopathology and disability. *Archives of General Psychiatry, 50*(10), 759–766.

Ormel, J., Von Korff, M., Ustun, T. B., Pini, S., Korten, A., & Oldehinkel, A. J. (1994). Common mental disorders and disability across cultures. Results from the WHO collaborative primary care study. *Jama-Journal of the American Medical Association, 272*, 1741–1748.

Ormel, J., VonKorff, M. J. B. F., & Riese, H. (2017). Set point theory and personality development: resolution of a paradox. In J. Specht (Ed.), *Personality development across the life span* (pp. 117–137). New York: Elsevier.

Ormel, J., Bosker, F. J., Hollon, S. D., & Ruhe, H. G. (2020). Can loss of agency and oppositional perturbation associated with antidepressant monotherapy and low-fidelity psychological treatment dilute the benefits of guideline-consistent depression treatment at the population level? *European Psychiatry, 63*(1), Article e89. https://doi.org/10.1192/j.eurpsy.2020.86

Patten, S. B. (2009). Accumulation of major depressive episodes over time in a prospective study indicates that retrospectively assessed lifetime prevalence estimates are too low. *BMC Psychiatry, 9*, 19. https://doi.org/10.1186/1471-244X-9-19

Patten, S. B., Williams, J. V. A., Lavorato, D. H., Bulloch, A. G. M., Wiens, K., & Wang, J. (2016). Why is major depression prevalence not changing? *Journal of Affective Disorders, 190*, 93–97. https://doi.org/10.1016/j.jad.2015.09.002

Patten, S. B., Williams, J. V. A., Lavorato, D. H., Wang, J. L., McDonald, K., & Bulloch, A. G. M. (2016). Major depression in Canada: What has changed over the past 10 years? *Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie, 61*(2), 80–85. https://doi.org/10.1177/0706743715625940

Phelan, J. C., Link, B. G., Stueve, A., & Pescosolido, B. A. (2000). Public conceptions of mental illness in 1950 and 1996: What is mental illness and is it to be feared? *Journal of Health and Social Behavior*, 188–207.

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. https://doi.org/10.1159/000318293

Reavley, N. J., & Jorm, A. F. (2014). Willingness to disclose a mental disorder and knowledge of disorders in others: Changes in Australia over 16 years. *Australian and New Zealand Journal of Psychiatry, 48*(2), 162–168. https://doi.org/10.1177/0004867413495317

Richter, D., Wall, A., Bruen, A., & Whittington, R. (2019). Is the global prevalence rate of adult mental illness increasing? Systematic review and meta-analysis. *Acta Psychiatrica Scandinavica, 140*(5), 393–407.

Ries, L. A., Harkins, D., Krapcho, M., Mariotto, A., Miller, B., Feuer, E. J., & Howlader, N. (2006). *SEER cancer statistics review, 1975-2003*.

Roberts, B. W., Hill, P. L., & Davis, J. P. (2017). How to change conscientiousness: The sociogenomic trait intervention model. *Personality Disorders-Theory Research and Treatment, 8*(3), 199–205. https://doi.org/10.1037/per0000242

Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science, 17*(1), 31–35. https://doi.org/10.1111/j.1467-8721.2008.00543.x

Rottenberg, J., Devendorf, A. R., Kashdan, T. B., & Disabato, D. J. (2018). The curious neglect of high functioning after psychopathology: The case of depression. *Perspectives on Psychological Science, 13*(5), 549–566. https://doi.org/10.1177/1745691618769868

Rush, A. J., Trivedi, M. H., Wisniewski, S. R., Nierenberg, A. A., Stewart, J. W., Warden, D., & Fava, M. (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.

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. https://doi.org/10.1017/S003329171200284X

Schoemaker, C., Ten Have, M., Sytema, S., & Verhaak, P. (2007). Trends in de geestelijke volksgezondheid in Nederland. *Maandblad Geestelijek Volksgezondheid, 62*(10), 824–835.

Schomerus, G., Schwahn, C., Holzinger, A., Corrigan, P. W., Grabe, H. J., Carta, M. G., & Angermeyer, M. C. (2012). Evolution of public attitudes about mental illness: A systematic review and meta-analysis. *Acta Psychiatrica Scandinavica, 125*(6), 440–452.

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). https://doi.org/10.1093/ijnp/pyv076

Simon, G. E., & Von Korff, M. (1995). Recall of psychiatric history in cross-sectional surveys: Implications for epidemiologic research. *Epidemiologic Reviews, 17*, 221–227.

Smits, I. A. M., Dolan, C. V., Vorst, H. C. M., Wicherts, J. M., & Timmerman, M. E. (2011). Cohort differences in big five personality factors over a period of 25 years. *Journal of Personality and Social Psychology, 100*(6), 1124–1138. https://doi.org/10.1037/a0022874

Spiers, N., Qassem, T., Bebbington, P., McManus, S., King, M., Jenkins, R., & Brugha, T. S. (2016). Prevalence and treatment of common mental disorders in the English national population, 1993–2007. *British Journal of Psychiatry, 209*(2), 150–156. https://doi.org/10.1192/bjp.bp.115.174979

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.

Stirman, S. W., DeRubeis, R. J., Crits-Christoph, P., & Rothman, A. (2005). Can the randomized controlled trial literature generalize to nonrandomized patients? *Journal of Consulting and Clinical Psychology, 73*(1), 127–135. https://doi.org/10.1037/0022-006X.73.1.127

Streiner, D. L. (2002). The 2 "Es" of research: Efficacy and effectiveness trials. *Canadian Journal of Psychiatry-Revue Canadienne De Psychiatrie, 47*(6), 552–556. https://doi.org/10.1177/070674370204700607

Terracciano, A., McCrae, R. R., Brant, L. J., & Costa, P. T., Jr. (2005). Hierarchical linear modeling analyses of the NEO-PI-R scales in the Baltimore longitudinal study of aging. *Psychology and Aging, 20*(3), 493–506. https://doi.org/10.1037/0882-7974.20.3.493

Terracciano, A. (2010). Secular trends and personality: Perspectives from longitudinal and cross-cultural studies-commentary on Trzesniewski & Donnellan (2010). *Perspectives on Psychological Science, 5*(1), 93–96. https://doi.org/10.1177/1745691609357017

Terracciano, A., McCrae, R. R., & Costa, P. T., Jr. (2010). Intra-individual change in personality stability and age. *Journal of Research in Personality, 44*(1), 31–37. https://doi.org/10.1016/j.jrp.2009.09.006

Thornicroft, G., Chatterji, S., Evans-Lacko, S., Gruber, M., Sampson, N., Aguilar-Gaxiola, S., & Kessler, R. C. (2017). Undertreatment of people with major depressive disorder in 21 countries. *British Journal of Psychiatry, 210*(2), 119–124. https://doi.org/10.1192/bjp.bp.116.188078

Thornicroft, G., Mehta, N., Clement, S., Evans-Lacko, S., Doherty, M., Rose, D., & Henderson, C. (2016). Evidence for effective interventions to reduce mental-health-related stigma and discrimination. *Lancet, 387*(10023), 1123–1132. https://doi.org/10.1016/S0140-6736(15)00298-6

Trivedi, M. H., Rush, A. J., Crismon, M. L., Kashner, T. M., Toprac, M. G., Carmody, T. J., & Shon, S. P. (2004). Clinical results for patients with major depressive disorder in the Texas medication algorithm project. *Archives of General Psychiatry, 61*(7), 669–680.

Trzesniewski, K. H., & Donnellan, M. B. (2010). Rethinking "generation me": Astudy of cohort effects from 1976–2006. *Perspectives on Psychological Science, 5*(1), 58–75. https://doi.org/10.1177/1745691609356789

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., de Wever, W. W. H., van der Wee, N. J. A., van Veen, T., Cuijpers, P., & Zitman, F. G. (2012). The generalizability of psychotherapy efficacy trials in major depressive disorder: an analysis of the influence of patient selection in efficacy trials on symptom outcome in daily practice. *BMC Psychiatry, 12*, 192. https://doi.org/10.1186/1471-244X-12-192

van der Lem, R., van der Wee, N. J. A., van Veen, T., & Zitman, F. G. (2012). Efficacy 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*(4), 226–234. https://doi.org/10.1159/000330890

Van der Linden, M., Westert, G., Bakker, D. D., & Schellevis, F. (2004). *Tweede Nationale Studie naar ziekten en verrichtingen in de huisartspraktijk: klachten en aandoeningen in de bevolking en in de huisartspraktijk*.

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.

Velden, J. V., De Bakker, D., Claessens, A., & Schellevis, F. (1991). *Een Nationale Studie van ziekten en verrichtingen in de huisartspraktijk: basisrapport, morbiditeit in de huisartspraktijk*.

Vetencourt, J. F. M., Sale, A., Viegi, A., Baroncelli, L., De Pasquale, R., O'Leary, O. F., & Maffei, L. (2008). The antidepressant fluoxetine restores plasticity in the adult visual cortex. *Science, 320*(5874), 385–388.

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. https://doi.org/10.1037/0022-006X.75.3.475

Vos, T., Abajobir, A. A., Abbafati, C., Abbas, K. M., Abate, K. H., Abd-Allah, F., & GBD 2016 Dis Injury Incidence Prev. (2017). Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet, 390*(10100), 1211–1259. https://doi.org/10.1016/S0140-6736(17)32154-2

Vos, T., Barber, R. M., Bell, B., Bertozzi-Villa, A., Biryukov, S., Bolliger, I., & Global Burden Dis Study. (2015). Global, regional, and national incidence, prevalence, and years lived with disability for 301 acute and chronic diseases and injuries in 188 countries, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet, 386*(9995), 743–800. https://doi.org/10.1016/S0140-6736(15)60692-4

Wakefield, J. C. (2002). Diagnosing DSM-IV-Part I: DSM-IV and the concept of disorder. *Behaviour Research and Therapy, 35*(7), 633–649.

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. https://doi.org/10.1007/s10488-016-0745-2

Weinberger, A. H., Gbedemah, M., Martinez, A. M., Nash, D., Galea, S., & Goodwin, R. D. (2018). Trends in depression prevalence in the USA from 2005 to 2015: widening disparities in vulnerable groups. *Psychological Medicine, 48*(8), 1308–1315. https://doi.org/10.1017/S0033291717002781

Wells, K. B. (1999). Treatment research at the crossroads: The scientific interface of clinical trials and effectiveness research. *American Journal of Psychiatry, 156*(1), 5–10. https://doi.org/10.1176/ajp.156.1.5

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. https://doi.org/10.1017/S0033291712001717

Wittchen, H., Kessler, R. C., Zhao, S., & Abelson, J. L. (1995). Reliability and clinical validity of UM-CIDI DSM-III-R generalized anxiety disorder. *Journal of Psychiatric Research, 29*(2), 95–110.

Wunderink, L., Nieboer, R. M., Wiersma, D., Sytema, S., & Nienhuis, F. J. (2013). Recovery in remitted first-episode psychosis at 7 years of follow-up of an early dose reduction/discontinuation or maintenance treatment strategy long-term follow-up of a 2-year randomized clinical trial. *JAMA Psychiatry, 70*(9), 913–920. https://doi.org/10.1001/jamapsychiatry.2013.19

Young, A. S., Klap, R., Shoai, R., & Wells, K. B. (2008). Persistent depression and anxiety in the United States: Prevalence and. *Psychiatric Services, 59*(12), 1391–1398. https://doi.org/10.1176/appi.ps.59.12.1391

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/10.1097/JCP.0b013e31803308e1