Efficacy and Acceptability of Pharmacological Treatments for Depressive Disorders in Primary Care: Systematic Review and Network Meta-Analysis

Klaus Linde, MD[1]

Levente Kriston, PhD[2]

Gerta Rücker, PhD[3]

Susanne Jamil[1]

Isabelle Schumann, MD[1]

Karin Meissner, MD[1,4]

Kirsten Sigterman, MD[1]

Antonius Schneider, MD[1]

[1]Institute of General Practice, Technische Universität München, Munich, Germany

[2]Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany

[3]Institute of Medical Biometry and Statistics, University Medical Center Freiburg, Freiburg, Germany

[4]Institute of Medical Psychology, Ludwig-Maximilians-University Munich, Munich, Germany

CORRESPONDING AUTHOR

Klaus Linde, MD
Institute of General Practice
Technische Universität München
Orleansstr. 47
D-81667 München, Germany
Klaus.Linde@tum.de

ABSTRACT

PURPOSE The purpose of this study was to investigate whether antidepressants are more effective than placebo in the primary care setting, and whether there are differences between substance classes regarding efficacy and acceptability.

METHODS We conducted literature searches in MEDLINE, Embase, Cochrane Central Register of Controlled Trials (CENTRAL), and PsycINFO up to December 2013. Randomized trials in depressed adults treated by primary care physicians were included in the review. We performed both conventional pairwise meta-analysis and network meta-analysis combining direct and indirect evidence. Main outcome measures were response and study discontinuation due to adverse effects.

RESULTS A total of 66 studies with 15,161 patients met the inclusion criteria. In network meta-analysis, tricyclic and tetracyclic antidepressants (TCAs), selective serotonin reuptake inhibitors (SSRIs), a serotonin-noradrenaline reuptake inhibitor (SNRI; venlafaxine), a low-dose serotonin antagonist and reuptake inhibitor (SARI; trazodone) and hypericum extracts were found to be significantly superior to placebo, with estimated odds ratios between 1.69 and 2.03. There were no statistically significant differences between these drug classes. Reversible inhibitors of monoaminoxidase A (rMAO-As) and hypericum extracts were associated with significantly fewer dropouts because of adverse effects compared with TCAs, SSRIs, the SNRI, a noradrenaline reuptake inhibitor (NRI), and noradrenergic and specific serotonergic antidepressant agents (NaSSAs).

CONCLUSIONS Compared with other drugs, TCAs and SSRIs have the most solid evidence base for being effective in the primary care setting, but the effect size compared with placebo is relatively small. Further agents (hypericum, rMAO-As, SNRI, NRI, NaSSAs, SARI) showed some positive results, but limitations of the currently available evidence makes a clear recommendation on their place in clinical practice difficult.

Ann Fam Med 2015;13:69-79. doi: 10.1370/afm.1687.

INTRODUCTION

Epidemiological studies indicate that depressive disorders are highly prevalent in the general population worldwide.[1] Most cases are seen and managed in primary care, and only a small proportion of these are referred to specialized care.[2] Most research findings upon which treatment decisions are made, however, have involved patients cared for by mental health specialists.[3] It is not fully clear whether the findings from trials in specialty settings can be generalized to primary care. There is some evidence suggesting that primary care patients with depressive disorders are less severely depressed,[4] experience a milder course of illness,[5] have a distinct symptom profile with more complaints of fatigue and somatic symptoms,[6] and are more likely to have accompanying physical complaints[7] than are patients referred to specialty mental health care. These differences could have an impact on guideline development and management of depression in primary care.

Pharmacological interventions are a cornerstone of antidepressant treatment,[8] yet there is an ongoing debate as to whether their relatively small effects compared with placebo observed in clinical trials are clinically relevant.[9,10] Meta-analyses restricted to primary care patients have been performed for some antidepressant drugs.[11-14] Researchers conclude that these treatments are effective in primary care settings. It is not possible, however, to determine whether the available treatment options are comparable (ie, whether some treatments are superior to others in primary care). Traditional meta-analyses are restricted to the direct comparison of 2 interventions by pooling data only from trials with similar treatment arms. Network meta-analysis allows for the estimation of relative effects of interventions that have not been compared directly.[15] We systematically reviewed randomized trials of pharmacological treatments of depression in primary care settings. We used conventional and network meta-analysis to investigate whether there is evidence that in the primary care setting antidepressants are more effective than placebo and whether there are differences in efficacy and acceptability between the various substance classes.

METHODS

Details of the methods have been described in our published protocol.[16] We also reviewed trials on psychological interventions. Because trials of pharmacological and psychological interventions differ greatly regarding recruitment strategies, patients, control interventions and outcomes, these trials are analyzed separately and reported in a companion article published in this issue.[17]

Search Strategy and Study Selection

We searched MEDLINE, Embase, Cochrane Central Register of Controlled Trials (CENTRAL), and PsycINFO (main search June 2011, last update searches December 2013; see Supplemental Appendix, section 1, for the complete MEDLINE search strategy). We searched trial registries for unpublished and ongoing studies. In addition, we screened references from identified trials and published systematic reviews focusing on primary care studies of depression treatments[11-14] for additional trials.

We included randomized controlled trials that compared drugs belonging to different pharmacological classes with one another or placebo in the treatment of adult patients having prevalent or incident unipolar depressive disorder. Patients had to be recruited from a primary care setting consisting of family physicians' or general practitioners' private practices, primary care clinics or networks, internists, or other nonpsychiatrists

providing primary care in their respective countries. We excluded trials that recruited patients from community-based centers specializing in mental health care. Trials had to report results of at least 1 of the following outcomes: response to treatment, remission, mean score on a depression scale (posttreatment or change from baseline), frequency of adverse effects, or study discontinuation (for any reason or from adverse effects).

Four authors (K.L., K.S., S.J., and K.M.) reviewed all trials for screening selection and extraction. In the first screening 1 reviewer excluded clearly irrelevant records. In the second screening, 2 reviewers independently checked all remaining records against inclusion criteria. The full texts of articles were obtained for all records that were considered potentially relevant or unclear and were assessed formally for eligibility by at least 2 reviewers independently. Disagreements were resolved by discussion.

Data Extraction and Assessment of Risk of Bias

In the first extraction step, at least 2 reviewers independently extracted information on patients, methods, and results of all included studies using a pretested form. The Cochrane Collaboration's tool for assessing risk of bias was used to assess internal validity.[18] As the included studies reported results on efficacy in a highly diverse and often incomplete manner, we performed an additional extraction round using a preference approach for extracting data for response and remission (Supplemental Appendix, sTable 2). This additional extraction was done by 1 reviewer (K.L.), whereas a second reviewer (K.S. or K.M.) cross-checked all extracted data and recalculated imputations. Adequacy of dosages tested was checked against guideline recommendations.[19]

Comparisons and Outcomes

As prespecified in the study protocol, we analyzed drugs according to their substance class.[16] Efficacy end points were response (primary outcome defined as at least a 50% score reduction on a depression scale) and remission (secondary outcome defined as having a symptom score below a fixed threshold). Patients with missing data were considered nonresponders or nonremitters. In cases where responder and remission data were not reported, we imputed it from available score data.[20] Acceptability outcomes were discontinuation (dropout) because of adverse effects (primary acceptability outcome), discontinuation for any reason, and the number of patients experiencing adverse effects. For our main analysis, we used data after completion of treatment. In the rare case in which treatment duration was longer than 3 months, the measurement point closest to 3 months was used. For the analysis of long-term

effects, we used the measurement closest to 6 months after randomization.

Statistical Analyses

We used the odds ratio as the effect measure. Conventional meta-analyses of pairwise direct comparisons within studies were performed using the inverse variance weighted random effects model option in the Cochrane Informatics and Management Department RevMan 5.2 software. For network meta-analyses, a Bayesian framework (WinBUGS and R interface R2WinBUGS [http://www.r-project.org/]) following the recommendations of the UK Decision Support Unit of the National Institute of Health and Clinical Excellence (NICE) was used to combine direct and indirect evidence.[21] To examine for inconsistency, we compared the results of the consistency model with that of an inconsistency model using the deviance information criterion.[22] The potential impact of 4 prespecified (risk of bias, diagnostic subtype of depression, mean age of participants, and duration of treatment) and 2 post hoc defined covariates (dosage and sample size) was analyzed using a meta-regression model.[23] Sensitivity analyses were performed excluding outlier studies. Funnel plots were produced for all direct comparisons with data from at least 5 trials.
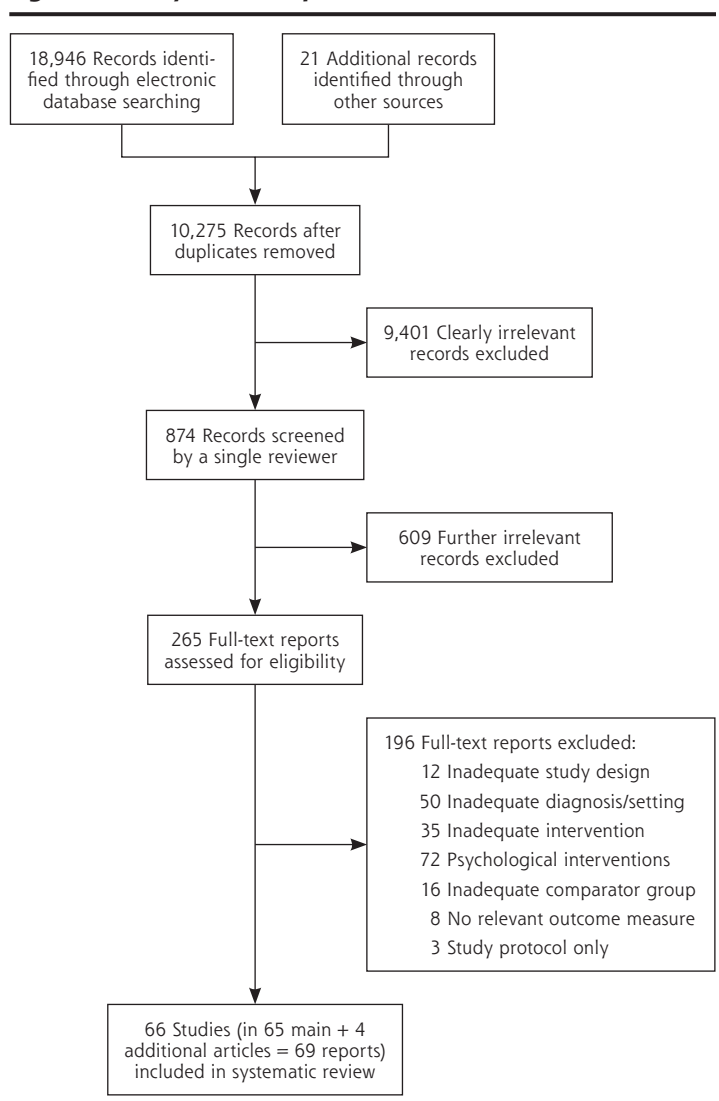
RESULTS

Study Selection and Characteristics of Included Studies

A total of 66 studies with 15,161 patients met the inclusion criteria (Figure 1; see also the Supplemental Appendix, Section 2 for references and Section 3 for characteristics of individual studies). Four trials were available as unpublished reports only (3 from a drug company's trial registry and 1 thesis). The 66 trials included 147 treatment groups. After the pooling of groups in which different dosages of the same agent or 2 agents of the same substance class had been tested, 140 treatment arms formed the basis of our analyses. In 37 treatment arms patients had received a tricyclic or tetracyclic antidepressive agent (TCA), in 37 a selective serotonin reuptake inhibitor (SSRI), in 6 a serotonin-noradrenaline reuptake inhibitor (SNRI), in 1 a noradrenaline reuptake inhibitor (NRI), in 5 a serotonin (5-HT$_2$) antagonist

and reuptake inhibitor (SARI), in 10 a noradrenergic and specific serotonergic antidepressive agent (NaSSA), in 6 a reversible inhibitor of monoamine oxidase A (rMAO-A), in 14 a hypericum extract (St. John's wort) (Table 1, which also displays single agents tested) and in 24 a placebo. Initial dosages were always within the range of recommendations[19] for starting treatment. In 11 treatment arms (from 10 trials), however, dosages remained below recommended standard dosage for all or most patients: in all 5 trazodone arms (150 mg in 3 and 100 to 200 mg in 2 trials), 2 amitriptyline arms (75 mg and 50 to 100 mg, respectively), 2 mianserin arms (40 mg and 30 to 60 mg, respectively), 1 clomipramine (40 mg), and 1 imipramine arm (mean dose 58 mg). In many trials daily dosages were in the lower range of recommendations for standard dosages. Thirty-eight (58%) studies were restricted to patients with major

Figure 1. Study selection process.

18,946 Records identified through electronic database searching

21 Additional records identified through other sources

10,275 Records after duplicates removed

9,401 Clearly irrelevant records excluded

874 Records screened by a single reviewer

609 Further irrelevant records excluded

265 Full-text reports assessed for eligibility

196 Full-text reports excluded:
    12 Inadequate study design
    50 Inadequate diagnosis/setting
    35 Inadequate intervention
    72 Psychological interventions
    16 Inadequate comparator group
    8 No relevant outcome measure
    3 Study protocol only

66 Studies (in 65 main + 4 additional articles = 69 reports) included in systematic review

depression only, whereas 28 (42%) studies also included patients who had other depressive disorders or did not provide details on the exact type of depressive disorders involved (Table 2).

Risk of Bias

Seventeen trials (26%) reported the method of generating the allocation sequence, and 14 (21%) reported an adequate method of allocation concealment (Supplemental Appendix, sTable 3 for the assessment of individual trials). All remaining trials provided no information on this issue. In 57 (86%) double-blind trials without clear indications of unblinding, we considered the risk of bias to be low; in 9 (14%) trials either no blinding or unblinding seemed likely. Bias resulting

from attrition seemed low in 21 (32%) studies and bias resulting from the reporting of outcomes was found in 49 (74%) studies. Overall risk of bias was considered to be low in 11 (17%), unclear in 23 (35%), and high in 32 (49%) studies.

Efficacy

Fifty-nine (89%) studies provided sufficient data to be included in the analysis of the main outcome measure response. For 17 of the 36 possible comparisons (46%) there was at least 1 head-to-head trial (Figure 2 for the network and Table 3 for pooled estimates; Section 4 of the Supplemental Appendix provides a forest plot with odds ratios of all individual trials). More than 3 comparative trials were available for TCAs vs SSRIs (n = 18), hypericum extracts vs placebo (n = 9), TCAs vs placebo (n = 8), SSRIs vs placebo (n = 7) and SSRIs vs hypericum extracts (n = 6). In 6 of the 13 direct comparisons with 2 or more trials, there was statistical heterogeneity (I^2 >40% and/or P <.1 in the χ^2 test). TCAs, SSRIs, SNRI, and hypericum extracts were significantly superior to placebo, but SNRI and NaSSAs were not. There were no trials comparing NRI, SARI, and rMAO-As with placebo. Funnel plots of comparisons with placebo were difficult to interpret because of the small numbers of studies per substance class (Supplemental Appendix, sFigures 3-7 for funnel plots). Visual inspection suggested asymmetry for trials comparing hypericum extract and placebo. The only significant differences between substance classes indicated superiority of TCAs compared with NaSSAs and rMAO-As, and of SARI compared with NaSSAs.

In the network meta-analysis, TCAs, SSRIs, SNRI, SARI (low-dose trazodone) and hypericum extracts were found to be significantly superior to the placebo, but effects were relatively small, with estimated odds ratios between 1.69 and 2.03 (Table 3). There were no significant differences between these drug
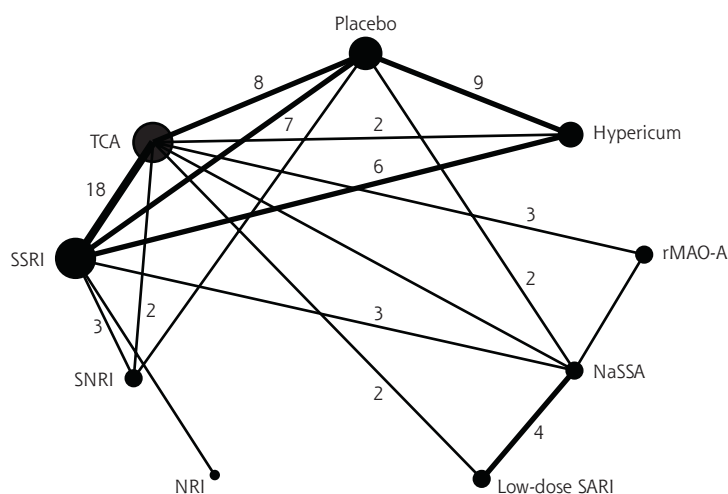
Table 1. Overview of Antidepressant Medications Tested in the Included Trials and in the Meta-Analysis

Medication Class	Trial Groups/Arms and Specific Medication Tested
TCA, tricyclic and tetracyclic antidepressants	40 Separate groups in trials, after pooling of groups in studies with more than 1 TCA, 37 arms in analyses
	14 Amitriptyline (1 pooling of 2 groups with different dosages, 1 pooled with dothiepin)
	8 Imipramine (1 pooled with desipramine)
	8 Dothiepin/dosulepin (1 pooled with amitriptyline)
	3 Clomipramine
	1 Amineptine
	1 Desipramine (pooled with imipramine)
	1 Doxepin
	1 Lofepramine
	1 Maprotiline
	1 Tianeptine
	1 Individualized TCA
SSRI, selective serotonin reuptake inhibitors	38 Arms in total, 37 after pooling of 2 arms in a study with 2 SSRI 35 arms
	12 Paroxetine
	9 Fluoxetine
	5 Sertraline
	5 Citalopram (1 pooled with escitalopram)
	2 Fluvoxamine
	3 Escitalopram (1 pooled with citalopram)
	1 Individualized SSRI
SNRI, serotonin-noradrenaline reuptake inhibitor	6 Arms in total
	6 Venlafaxine
NRI, noradrenaline reuptake inhibitor	1 Arm in total
	1 Reboxetine
SARI, serotonin (5-HT$_2$) antagonists and reuptake inhibitor	5 Arms in total
	5 Trazodone
NaSSA, noradrenergic and specific serotonergic antidepressive agents	9 Arms in total
	8 Mianserin (2 underdosing)
	1 Mirtazapine
rMAO-A, reversible inhibitors of monoaminoxidase A	8 Arms in total, 6 in analyses
	4 Moclobemide,
	4 Minaprine (2 pooling of different dosages)
Hypericum, extracts from *Hypericum perforatum L.* (St. John's wort)	15 Arms in total, 14 after pooling of 2 arms testing different extracts in 1 study; 12 different extracts

Table 2. Characteristics of Included Studies

Characteristic	Publication Year 1996 (1971, 2012)
Number of patients	
Total	15,161
Median (minimum, maximum)	162 (21, 1,385)
Diagnosis, No. (%)	
Major depression only	38 (58)
Mixed/unclear	25 (38)
Minor depression/dysthymia	3 (5)
Classification of depression, No. (%)	
Diagnostic and Statistical Manual of Mental Disorders (Fourth Edition)	15 (23)
Diagnostic and Statistical Manual of Mental Disorders (Third Edition)	24 (36)
International Classification of Diseases, Tenth Revision	6 (9)
International Classification of Diseases, Ninth Revision	4 (6)
Research Diagnostic Criteria	4 (6)
Not reported/other	13 (20)
Restricted to patients >55 y, No. (%)	8 (12)
Median length of treatment in weeks (minimum, maximum)	6 (4, 52)
Overall risk of bias, No. (%)	
High (high risk in 1 or more items)	32 (49)
Unclear (no item high risk, <3 low risk)	23 (35)
Low (at least 3 low, no high risk)	11 (17)
Data available for meta-analysis, No. (%)	
Response	59 (89)
Remission	55 (83)
Total number of patients dropping out	60 (91)
Number of patients dropping out due to adverse effects	58 (88)
Number of patients reporting adverse effects	40 (61)

Figure 2. Network for the main efficacy outcome response.

Figures indicate the number of direct comparisons (lines without figure indicate 1 comparison)

Hypericum = extract from *Hypericum perforatum L.*; NaSSA = noradrenergic and specific serotonergic antidepressive agent (mianserin, mirtazapine); NRI = noradrenaline reuptake inhibitor (reboxetine); rMAO-A = reversible inhibitor of monoaminoxidase A (moclobemide, minaprine); SARI = serotonin (5-HT$_2$) antagonist and reuptake inhibitor (trazodone); SNRI = serotonin-noradrenaline reuptake inhibitor (venlafaxine); SSRI = selective serotonin reuptake inhibitor; TCA = tricyclic and tetracyclic antidepressant.

classes, but 95% credible intervals were wide except for the comparison between TCAs and SSRIs. NRI, NaSSAs, and rMAO-As were not significantly different from placebo. TCAs, SSRIs, and hypericum extracts were significantly superior to NaSSAs and rMAO-As. Hypericum extracts were also more effective than NRI. SNRI and low-dose SARI were superior to NaSSAs. We found no evidence of inconsistency between direct and indirect comparisons.

Fixed-effects analyses yielded results very similar to random-effects analyses, thus providing no evidence for heterogeneity. Meta-regression analyses did not show a significant influence of the type of depression (major depression or not), risk of bias, restriction to elderly patients, timing of the outcome measurement, underdosing, and sample size on treatment outcome. Correspondingly, model estimates for patients with major depression, for studies with adequate drug dosages, and for large trials were very similar to the unadjusted main estimates (Supplemental Appendix, Section 6). The exclusion of outlier studies did not have any relevant impact on the findings.

In secondary efficacy analyses using remission as the outcome, all drug classes were superior to placebo without significant differences between drug classes (Supplemental Appendix, Sections 6 and 7). Only 9 trials reported long-term outcomes (>12 weeks), and of these, only 1 included a placebo control group. Point estimates were similar (between 1.55 and 1.78) for the drugs that could be included in the network (SSRI, TCA, SNRI, NaSSA, and hypericum) but confidence intervals were very wide, and differences compared with placebo were not statistically significant.

Acceptability

Fifty-eight studies (88%) reported the number of patients dropping out because of adverse effects. The data for 5 studies could not be used

Table 3. Results of Conventional and Network Meta-Analyses for the Main Efficacy Outcome Response

Medication	TCA	SSRI	SNRI	NRI	Low-Dose SARI
TCA					
OR (95% CI)[a]		0.95 (0.84-1.08)	0.83 (0.37-1.89)	na	1.43 (0.60-3.42)
No. of studies; I^2, %; *P* value[b]		18; 0; .45;	2; 55; 0.14	na	2; 79; .03
OR (95% CI)[c]		1.04 (0.88-1.21)	1.00 (0.71-1.40)	1.23 (0.73-2.14)	0.97 (0.65-1.47)
SSRI					
OR (95% CI)[a]	1.05 (0.93-1.19)		0.92 (0.67-1.25)	1.18 (0.85-1.65)	na
No. of studies; I^2, %; *P* value[b]	18; 0; .45		4; 64; .04	1	na
OR (95% CI)[c]	0.97 (0.82-1.13)		0.96 (0.71-1.29)	1.19 (0.71-2.04)	0.94 (0.63-1.43)
SNRI					
OR (95% CI)[a]	1.20 (0.53-2.72)	1.09 (0.79-1.49)		na	na
No. of studies; I^2, %; *P* value[b]	2; 55; .14	4; 64; .04		na	na
OR (95% CI)[c]	1.00 (0.72-1.41)	1.03 (0.77-1.42)		1.23 (0.67-2.28)	0.97 (0.60-1.63)
NRI					
OR (95% CI)[a]	na	0.84 (0.61-1.18)	na		na
No. of studies; I^2, %; *P* value[b]	na	1	na		na
OR (95% CI)[c]	0.60 (0.41-0.91)	0.63 (0.41-0.87)	0.61 (0.36-1.02)		0.79 (0.41-1.57)
Low-dose SARI					
OR (95% CI)[a]	0.70 (0.29-1.67)	na	na	na	
No. of studies; I^2, %; *P* value[b]	2; 79; .03	na	na	na	
OR (95% CI)[c]	1.03 (0.68-1.55)	1.06 (0.70-1.58)	1.03 (0.61-1.67)	1.26 (0.63-2.47)	
NaSSA					
OR (95% CI)[a]	0.26 (0.11-0.60)	0.95 (0.49-1.85)	na	na	0.47 (0.31-0.73)
No. of studies; I^2, %; *P* value[b]	1	3; 60; .08	na	na	4; 0; .56
OR (95% CI)[c]	0.64 (0.46-0.93)	0.67 (0.47-0.95)	0.64 (0.41-0.98)	0.79 (0.42-1.45)	0.63 (0.41-0.93)
rMAO-A					
OR (95% CI)[a]	0.61 (0.43-0.86)	na	na	na	na
No. of studies; I^2, %; *P* value[b]	3; 0; .79	na	na	na	na
OR (95% CI)[c]	0.61 (0.41-0.91)	0.63 (0.41-0.97)	0.61 (0.36-1.02)	0.74 (0.37-1.52)	0.59 (0.35-1.04)
Hypericum					
OR (95% CI)[a]	1.15 (0.71-1.85)	1.14 (0.81-1.59)	na	na	na
No. of studies; I^2, %; *P* value[b]	2; 24; .25	6; 43; .12	na	na	na
OR (95% CI)[c]	1.16 (0.81-1.50)	1.20 (0.96-1.51)	1.16 (0.80-1.68)	1.43 (0.81-2.64)	1.13 (0.72-1.73)

CI = confidence interval; CI = credible interval; hypericum = extract from *Hypericum perforatum L.*; rMAO-A = reversible inhibitor of monoaminoxidase A (moclobemide, minaprine); na = not available; NaSSA = noradrenergic and specific serotonergic antidepressive agent (mianserin, mirtazapine); NRI = noradrenaline reuptake inhibitor (reboxetine); OR = odds ratio; SARI = serotonin (5-HT$_2$) antagonist and reuptake inhibitor (trazodone); SNRI = serotonin-noradrenaline reuptake inhibitor (venlafaxine); SSRI = selective serotonin reuptake inhibitor; TCA = tricyclic and tetracyclic antidepressant.

Note: Odds ratios >1 indicate more study discontinuation in patients receiving the treatment given in the row heading.

[a] Conventional meta-analysis of within-study comparisons with pooled odds ratios.

[b] Studies with direct comparisons available, with I^2 value and *P* value from the χ^2 test for heterogeneity.

[c] From network meta-analysis.

in meta-analyses, however, because there were no dropouts attributable to side effects in any treatment group. For 16 of the 36 possible comparisons (44%) there was at least 1 head-to-head trial. In 4 of the 13 direct comparisons with 2 or more trials, there was an indication of statistical heterogeneity (Table 4). Because of the often very wide confidence intervals (caused by the small number of trials and low event rates), findings of conventional meta-analyses of head-to-head comparisons have to be interpreted with great caution (Supplemental Appendix, Section 3 for individual study findings). TCAs were associated with

significantly more dropouts resulting from adverse effects compared with rMAO-As, the NRI was associated with significantly more than SSRIs, and NaSSAs were associated with significantly more than the low-dose SARI. In network meta-analysis combining direct and indirect evidence (Table 4), TCAs, SSRIs, SNRI, NRI, and NaSSAs were associated with significantly more study discontinuations resulting from adverse effects than was placebo. Attrition was not significantly different from placebo with (mostly low-dose) SARI, rMAO-As, and hypericum extracts. rMAO-As and hypericum extracts were associated with signifi-

NaSSA	rMAO-A	Hypericum	Placebo
3.83 (1.67-8.75)	1.65 (1.16-2.34)	0.87 (0.54-1.41)	1.67 (1.24-2.25)
1	3; 0; .79;	2; 24; 0.25;	8; 25; .23;
1.55 (1.11-2.18)	1.65 (1.10-2.45)	0.86 (0.67-1.10)	1.75 (1.42-2.15)
1.05 (0.54-2.03)	na	0.88 (0.63-1.23)	1.57 (1.31-1.89)
3; 60; .08	na	6; 46; .10	7; 0; .83
1.50 (1.06-2.11)	1.60 (1.03-2.54)	0.83 (0.66-1.04)	1.69 (1.40-2.04)
na	na	na	1.94 (0.96-3.93)
na	na	na	1
1.55 (1.02-2.45)	1.65 (0.98-2,80)	0.86 (0.59-1.24)	1.75 (1.24-2.47)
na	na	na	na
na	na	na	na
0.94 (0.58-1.54)	1.34 (0.66-2.70)	0.52 (0.32-0.84)	1.42 (0.81-2.49)
2.11 (1.37-3.25)	na	na	na
4; 0; .56	na	na	na
1.60 (1.07-2.42)	1.70 (0.96-2.89)	0.89 (0.58-1.38)	1.80 (1.17-2.76)
	1.07 (0.49-2.34)	na	1.34 (0.85-2.10)
	1	na	2; 0; .98
	1.06 (0.65-1.72)	0.55 (0.37-0.82)	1.13 (0.78-1.62)
0.93 (0.43-2.03)		na	na
1		na	na
0.94 (0.58-1.54)		0.52 (0.32-0.84)	1.06 (0.67-1.66)
na	na		2.02 (1.41-2.88)
na	na		9; 43; .08
1.80 (1.21. 2.67)	1.92 (1.20-3.16)		2.03 (1.63-2.53)

cantly fewer dropouts because of adverse effects compared with TCAs, SSRIs, SNRI, NRI, and NaSSAs. (For results on dropouts for any reasons and patients with adverse effects in comparison with placebo, see the Supplemental Appendix, Section 7).

DISCUSSION

This systematic review shows that a considerable number of randomized trials have investigated the short-term (up to 12 weeks) efficacy and acceptability of pharmacological treatments for depression in pri-

mary care. Yet, although we compared groups of similar interventions instead of single specific interventions, the number of trials in some substance groups was low. TCAs, SSRIs, and hypericum extracts were investigated more often in the primary care setting than other drug classes. Our primary efficacy analyses suggest that TCAs, SSRIs, SNRI, low-dose SARI, and hypericum extracts are effective for the treatment of acute depression, but effects when compared with placebo were modest in size. With 40% of patients responding to placebo, an odds ratio of 1.69 (as found for SSRIs) would mean that 53% of patients receiving an antidepressant respond. The absolute difference of 13% more than placebo corresponds to a number needed to treat between 7 and 8 patients. In secondary analyses using remission as the outcome, NRI, NaSSAs, and rMAO-As were also found to be effective. rMAO-As, hypericum extracts, and low-dose SARI tended to be associated with more favorable results in the acceptability analyses. The quality of most of the trials included in our review was mediocre or weak. Because there was a small number of studies with observation periods of longer than 12 weeks, reliable comparative analysis of long-term effects was not possible.

Conventional meta-analyses or randomized trials restricted to primary care patients have been performed for SSRIs and TCAs compared with placebo,[11,12] SSRIs compared with TCAs,[13] and a variety of newer antidepressants compared with different comparators.[14] The reviews concluded that the reviewed antidepressants are superior to placebo in primary care patients.[13,14] The evidence on the relative efficacy of antidepressants was sparse and considered to be of variable quality but suggested no major differences. Regarding acceptability, findings favored SSRIs and most newer antidepressants when compared with TCAs.[13,14] Our review includes more primary care-based trials than the mentioned previous reviews taken together. In addition to conventional pairwise meta-analysis, we performed network meta-analysis, making efficient use of all available data. Our analyses confirm that short-term effects of SSRIs and TCAs are similar and that fewer patients receiving SSRIs report adverse

Table 4. Results of Conventional and Network Meta-Analyses for the Main Acceptability Outcome Study Discontinuation Because of Adverse Effects

Medication	TCA	SSRI	SNRI	NRI	Low-Dose SARI
TCA					
OR (95% CI)[a]		1.15 (0.85-1.55)	0.86 (0.46-1.62)	na	1.18 (0.46-3.01)
No. of studies; I^2, %; P value[b]		14; 43; .05	2; 0; .80	na	2; 54; .14
OR (95% CI)[c]		1.12 (0.84-1.47)	0.89 (0.55-1.37)	0.44 (0.18-1.06)	1.86 (0.92-3.75)
SSRI					
OR (95% CI)[a]	0.87 (0.64-1.17)		0.81 (0.58-1.13)	0.39 (0.28-0.55)	na
No. of studies; I^2, %; P value[b]	14; 43; .05		4; 38; .18	1	na
OR (95% CI)[c]	0.89 (0.68-1.19)		0.79 (0.55-1.19)	0.39 (0.16-0.90)	1.65 (0.85-3.45)
SNRI					
OR (95% CI)[a]	1.16 (0.62-2.20)	1.24 (0.89-1.74)		na	na
No. of studies; I^2, %; P value[b]	2; 0; .80	4; 38; .18		na	na
OR (95% CI)[c]	1.12 (0.73-1.81)	1.26 (0.84-1.83)		0.50 (0.19-1.23)	2.08 (0.95-4.85)
NRI					
OR (95% CI)[a]		2.53 (1.81-3.53)	na		na
No. of studies; I^2, %; P value[b]		1	na		na
OR (95% CI)[c]	2.25 (0.95-5.48)	2.54 (1.11-6.37)	2.01 (0.82-5.20)		4.19 (1.59-12.48)
SARI					
OR (95% CI)[a]	0.84 (0.33-2.15)	na	na	na	
No. of studies; I^2, %; P value[b]	2; 54; .14	na	na	na	
OR (95% CI)[c]	0.54 (0.26-1.08)	0.61 (0.29-1.17)	0.48 (0.21-1.05)	0.24 (0.08-0.63)	
NaSSA					
OR (95% CI)[a]	1.97 (0.71-5.51)	1.28 (0.69-2.36)	na	na	3.55 (1.27-9.90)
No. of studies; I^2, %; P value[b]	1	3; 31; .23	na	na	3; 58; .09
OR (95% CI)[c]	1.56 (0.88-2.76)	1.75 (1.04-3.02)	1.39 (0.73, 2.76)	0.69 (0.27-1.71)	2.89 (1.56-6.18)
rMAO-A					
OR (95% CI)[a]	0.39 (0.18-0.85)	na	na	na	na
No. of studies; I^2, %; P value[b]	5; 47; .11	na	na	na	na
OR (95% CI)[c]	0.38 (0.20-0.67)	0.43 (0.21-0.81)	0.34 (0.15-0.71)	0.17 (0.06-0.48)	0.71 (0.28-1.73)
Hypericum					
OR (95% CI)[a]	0.43 (0.17-1.09)	0.68 (0.35-1.30)	na	na	na
No. of studies; I^2, %; P value[b]	2; 0; 0.89	6; 0; .92	na	na	na
OR (95% CI)[c]	0.42 (0.22-0.78)	0.47 (0.24-0.85)	0.38 (0.18-0.75)	0.19 (0.07-0.50)	0.78 (0.31-1.83)

CI = confidence interval; CI = credible interval; hypericum = extract from *Hypericum perforatum L.*; rMAO-A = reversible inhibitor of monoaminoxidase A (moclobemide, minaprine); na = not available; NaSSA = noradrenergic and specific serotonergic antidepressive agent (mianserin, mirtazapine); NRI = noradrenaline reuptake inhibitor (reboxetine); OR = odds ratio; SARI = serotonin (5-HT$_2$) antagonist and reuptake inhibitor (trazodone); SNRI = serotonin-noradrenaline reuptake inhibitor (venlafaxine); SSRI = selective serotonin reuptake inhibitor; TCA = tricyclic and tetracyclic antidepressant.

Note: Odds ratios >1 indicate more study discontinuation in patients receiving the treatment given in the row heading.

[a] Conventional meta-analysis of within-study comparisons with pooled odds ratios.

[b] Studies with direct comparisons available, with I^2 value and P value from the χ^2 test for heterogeneity.

[c] From network meta-analysis.

effects, but we did not find any significant differences regarding study discontinuation resulting from adverse effects or for any reasons among the investigated substance classes. Our findings are in accordance with a conventional meta-analysis of trials mostly performed in specialist mental health care showing that the NRI reboxetine has relevant adverse effects and little efficacy.[24] In the network meta-analysis hypericum extracts showed similar efficacy but better acceptability than SSRIs and TCAs. This result is in line with a conventional meta-analysis of hypericum trials from all settings.[25] Hypericum extracts have a high risk of adverse interactions with other drugs,[26] however, and patients with multiple morbid conditions are typically excluded from randomized trials. Furthermore, the hypericum meta-analysis[25] also found that evidence of efficacy was considerably more solid in German-speaking regions than in other countries, including the United States, which makes interpretation difficult.

Because of the limited number of trials per single agents, our review cannot provide estimates of relative efficacy on this level. Network meta-analyses of single

NaSSA	rMAO-A	Hypericum	Placebo
0.51 (0.18-1.42)	2.57 (1.18-5.60)	2.33 (.92-5.93)	2.30 (1.10-4.81)
1	5; 47; .11	2; 0; .89	7; 24; .25
0.64 (0.36-1.14)	2.63 (1.49-5.09)	2.37 (1.28-4.49)	2.47 (1.59-3.84)
0.78 (0.42-1.44)	na	1.48 (0.77-2.85)	1.86 (1.16-2.98)
3; 31; .23	na	6; 0; .92	8; 0; .64
0.57 (0.33-0.96)	2.34 (1.24-4.84)	2.11 (1.18-4.14)	2.20 (1.44-3.36)
na	na	na	2.96 (0.90-9.73)
na	na	na	1
0.72 (0.36-1.37)	2.95 (1.41-6.69)	2.66 (1.33-5.57)	2.77 (1.58-4.87)
na	na	na	na
na	na	na	na
1.45 (0.58-3.74)	5.94 (2.09-17.13)	5.35 (1.99-14.94)	5.58 (2.28-13.66)
0.28 (0.10-0.79)	na	na	na
3; 58; 0.09	na	na	na
0.35 (0.16-0.64)	1.42 (0.58-3.55)	0.78 (0.31-1.83)	1.33 (0.61-2.89)
	na	na	3.52 (1.40-8.84)
	na	na	2; 0; .90
	4.10 (1.81-10.38)	3.69 (1.71-8.56)	3.85 (2.05-7.25)
na		na	na
na		na	na
0.24 (0.10-0.55)		0.90 (0.35-2.04)	0.94 (0.43-2.05)
na	na		0.79 (0.31-1.97)
na	na		5; 0; .48
0.27 (0.11-0.58)	1.11 (0.49-2.83)		1.04 (0.53-2.06)

guidelines make recommendations that are specific to the primary care setting. Generally, for patients with whom pharmacotherapy is considered, the UK guideline recommends generic SSRIs as the first choice; the Canadian guideline recommends all second-generation antidepressants, and the US and the German guidelines are somewhat less explicit by listing a number of situations in which other choices might be preferable. The UK guideline explicitly states that antidepressants should not be used routinely for persistent subthreshold depressive symptoms or mild depression. All guidelines agree that there is evidence for the effectiveness of hypericum extracts for mild to moderate depression. But while the German and the Canadian guidelines are supportive of them when the risk of interactions is taken into account adequately, the UK guideline explicitly discourages their use. The US guideline does not make a recommendation.

Despite guideline recommendations, evidence from conventional and network meta-analyses, and widespread use, there is an ongoing discussion of the extent to which antidepressants have clinically relevant effects compared with placebo. In the meta-analyses the effects size compared with placebo is frequently considered rather small. Yet, findings from published trials tend to overestimate these effects because of publication and reporting bias. A meta-analysis of 74 trials of 12 second-generation antidepressants registered with the US Food and Drug Administration (FDA) indicated statistically significant effects when compared with placebo for all agents, but meta-analysis of published trials only would have inflated the effects size by 32% on average.[34] For our review, we searched several trial registries of manufacturers, but because of the large number of agents covered in our review, we were unable to comprehensively search for unpublished trials. Owing to our limited resources and that many primary care trials are performed outside the approval process, we did not search the FDA database. We cannot rule out that our findings are distorted by publication or reporting bias to some extent.

second-generation antidepressants have been performed across settings (mainly including trials from specialist mental health care). One review concluded that sertraline might be the best choice when starting treatment for moderate and severe depression,[27] one favored escitalopram,[28] and another considered the evidence insufficient to recommend a particular agent.[29,30]

We compared our findings with recommendations of current high-quality clinical practice guidelines from the United States,[8] the United Kingdom,[31] Canada,[32,33] and Germany.[19] Currently, none of these

Another widely cited meta-analysis of FDA-registered trials found that compared with placebo,

medication effects became clinically relevant only in patients with very severe depression.[9] This finding would be particularly important for primary care, as family physicians see many patients with less than severe depression. The meta-analysis included very few trials in patients with mild to moderate depression, however, and used aggregate data to investigate the association between baseline severity and outcome. A recent reanalysis of the same data set using different meta-analytic methods did not confirm the influence of initial severity on efficacy.[35] Also, the findings of 2 newer meta-analyses of individual patient data—a more appropriate method—are contradictory.[36,37] In our analyses, there were no major differences in trials limited to patients with major depression and trials that included other depressive patients. We could not investigate the association between baseline severity and outcome in detail because of the multiple different depression scales used in the primary studies.

Overall, despite publication bias and problems with methodological quality, findings seem to agree that antidepressants are significantly more effective than placebo. Most critical discussions focus on the clinical relevance of effects on the basis of meta-analytical effect sizes. Still, it should be noted that when comparing antidepressive medications with placebo alone, effect sizes cannot provide sufficient information on the clinical relevance of treatment effects in individual clinical contexts, so this issue is likely to remain controversial.[9,38] The results of our analyses indicate that antidepressants have higher short-term effects when compared with placebo also in primary care.

For decision making in individual patients, family physicians should be aware that SSRIs and TCAs have a somewhat more solid evidence base than other substance classes (with SSRIs having a slightly better acceptability profile). Further agents (hypericum, rMAO-A, SNRI, NRI, NaSSA, SARI) showed some positive results, but limitations of the currently available evidence make difficult a clear recommendation on their place in clinical practice. Differences between substance classes and between single second-generation antidepressants seem to be relatively minor. The latter findings come from network meta-analyses of trials mostly performed in specialized mental health care, but our results suggest that results from primary care trials and other trials are broadly similar.

It must be emphasized that there are very few data from primary care trials regarding long-term effectiveness and acceptability. Future research should prioritize large, long-term, pragmatic trials and observational studies addressing clinically relevant questions, such as the best management of mild-to-moderate depression and comparison of pharmacological and psychological treatments under conditions of routine care and stepped-care strategies.

REFERENCES

1. Demyttenaere K, Bruffaerts R, Posada-Villa J, et al; WHO World Mental Health Survey Consortium. Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. *JAMA.* 2004;291(21):2581-2590.

2. Bijl RV, Ravelli A. Psychiatric morbidity, service use, and need for care in the general population: results of The Netherlands Mental Health Survey and Incidence Study. *Am J Public Health.* 2000;90(4):602-607.

3. Peveler R, Kendrick T. Treatment delivery and guidelines in primary care. *Br Med Bull.* 2001;57:193-206.

4. Schwenk TL, Coyne JC, Fechner-Bates S. Differences between detected and undetected patients in primary care and depressed psychiatric patients. *Gen Hosp Psychiatry.* 1996;18(6):407-415.

5. Simon GE, VonKorff M. Recognition, management, and outcomes of depression in primary care. *Arch Fam Med.* 1995;4(2):99-105.

6. Williamson PS, Yates WR. The initial presentation of depression in family practice and psychiatric outpatients. *Gen Hosp Psychiatry.* 1989;11(3):188-193, discussion 216-221.

7. Kroenke K, Price RK. Symptoms in the community. Prevalence, classification, and psychiatric comorbidity. *Arch Intern Med.* 1993; 153(21):2474-2480.

8. Trangle M, Dieperink B, Gabert T, et al. Major depression in adults in primary care. Bloomington, MN: Institute for Clinical Systems Improvement; 2012.

9. Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med.* 2008;5(2):e45.

10. Adli M, Hegerl U. Do we underestimate the benefits of antidepressants? *Lancet.* 2014;383(9926):1361-1362.

11. Arroll B, Elley CR, Fishman T, et al. Antidepressants versus placebo for depression in primary care. *Cochrane Database Syst Rev.* 2009; (3):CD007954.

12. Arroll B, Macgillivray S, Ogston S, et al. Efficacy and tolerability of tricyclic antidepressants and SSRIs compared with placebo for treatment of depression in primary care: a meta-analysis. *Ann Fam Med.* 2005;3(5):449-456.

13. MacGillivray S, Arroll B, Hatcher S, et al. Efficacy and tolerability of selective serotonin reuptake inhibitors compared with tricyclic antidepressants in depression treated in primary care: systematic review and meta-analysis. *BMJ.* 2003;326(7397):1014.

14. Mulrow CD, Williams JW Jr, Chiquette E, et al. Efficacy of newer medications for treating depression in primary care patients. *Am J Med.* 2000;108(1):54-64.

15. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ.* 2005;331(7521):897-900.

16. Linde K, Schumann I, Meissner K, et al. Treatment of depressive disorders in primary care—protocol of a multiple treatment systematic review of randomized controlled trials. *BMC Fam Pract.* 2011;12:127.

17. Linde K, Sigterman K, Kriston L, et al. Effectiveness of psychological treatment for depressive disorders in primary care: systematic review and meta-analysis. *Ann Fam Med.* 2014;12(6):56-68.

18. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0. Updated March 2011. London, England: The Cochrane Collaboration; 2008. http://handbook.cochrane.org.

19. DGPPN, BÄK, KBV, AWMF, AkdÄ, BPtK, BApK, DAGSHG, DEGAM, DGPM, DGPs, DGRW: S3-Leitlinie/Nationale Versorgungsleitlinie Unipolare Depression - Langfassung. Berlin, Düsseldorf: DGPPN, ÄZQ, AWMF; 2009. http://www.leitlinien.de/nvl/depression/.

20. Furukawa TA, Cipriani A, Barbui C, Brambilla P, Watanabe N. Imputing response rates from means and standard deviations in meta-analyses. *Int Clin Psychopharmacol.* 2005;20(1):49-52.

21. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making.* 2013;33(5):607-617.

22. Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. Evidence synthesis for decision making 4: inconsistency in networks of evidence based on randomized controlled trials. *Med Decis Making.* 2013;33(5):641-656.

23. Dias S, Sutton AJ, Welton NJ, Ades AE. Evidence synthesis for decision making 3: heterogeneity—subgroups, meta-regression, bias, and bias-adjustment. *Med Decis Making.* 2013;33(5):618-640.

24. Eyding D, Lelgemann M, Grouven U, et al. Reboxetine for acute treatment of major depression: systematic review and meta-analysis of published and unpublished placebo and selective serotonin reuptake inhibitor controlled trials. *BMJ.* 2010;341:c4737.

25. Linde K, Berner MM, Kriston L. St John's wort for major depression. *Cochrane Database Syst Rev.* 2008;(4):CD000448.

26. Mannel M. Drug interactions with St John's wort: mechanisms and clinical implications. *Drug Saf.* 2004;27(11):773-797.

27. Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet.* 2009;373(9665):746-758.

28. Ramsberg J, Asseburg C, Henriksson M. Effectiveness and cost-effectiveness of antidepressants in primary care: a multiple treatment comparison meta-analysis and cost-effectiveness model. *PLoS One.* 2012;7(8):e42003.

29. Gartlehner G, Gaynes BN, Hansen RA, et al. Comparative benefits and harms of second-generation antidepressants: background paper for the American College of Physicians. *Ann Intern Med.* 2008;149(10):734-750.

30. Gartlehner G, Hansen RA, Morgan LC, et al. Comparative benefits and harms of second-generation antidepressants for treating major depressive disorder: an updated meta-analysis. *Ann Intern Med.* 2011;155(11):772-785.

31. National Collaborating Centre for Mental Health. Depression: the treatment and management of depression in adults. http://www.nice.org.uk/nicemedia/live/12329/45896/45896.pdf. Updated 2010.

32. Lam RW, Kennedy SH, Grigoriadis S, et al; Canadian Network for Mood and Anxiety Treatments. Canadian Network for Mood and Anxiety Treatments (CANMAT) clinical guidelines for the management of major depressive disorder in adults. III. Pharmacotherapy. *J Affect Disord.* 2009;117(Suppl 1):S26-S43.

33. Ravindran AV, Lam RW, Filteau MJ, et al; Canadian Network for Mood and Anxiety Treatments. Canadian Network for Mood and Anxiety Treatments (CANMAT) Clinical guidelines for the management of major depressive disorder in adults. V. Complementary and alternative medicine treatments. *J Affect Disord.* 2009;117(Suppl 1):S54-S64.

34. Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med.* 2008;358(3):252-260.

35. Fountoulakis KN, Veroniki AA, Siamouli M, Möller HJ. No role for initial severity on the efficacy of antidepressants: results of a multi-meta-analysis. *Ann Gen Psychiatry.* 2013;12(1):26.

36. Fournier JC, DeRubeis RJ, Hollon SD, et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA.* 2010;303(1):47-53.

37. Gibbons RD, Hur K, Brown CH, Davis JM, Mann JJ. Benefits from antidepressants: synthesis of 6-week patient-level outcomes from double-blind placebo-controlled randomized trials of fluoxetine and venlafaxine. *Arch Gen Psychiatry.* 2012;69(6):572-579.

38. Cuijpers P, Turner EH, Koole SL, van Dijke A, Smit F. What is the threshold for a clinically relevant effect? The case of major depressive disorders. *Depress Anxiety.* 2014;31(5):374-378.