Comparative Benefits and Harms of Second-Generation Antidepressants for Treating Major Depressive Disorder

An Updated Meta-analysis

Gerald Gartlehner, MD, MPH; Richard A. Hansen, PhD, RPh; Laura C. Morgan, MA; Kylie Thaler, MD, MPH; Linda Lux, MPA; Megan Van Noord, MSIS; Ursula Mager, PhD, MPH; Patricia Thieda, MA; Bradley N. Gaynes, MD, MPH; Tania Wilkins, MSc; Michaela Strobelberger, MA; Stacey Lloyd, MPH; Ursula Reichenpfader, MD, MPH; and Kathleen N. Lohr, PhD

Background: Second-generation antidepressants dominate the management of major depressive disorder (MDD), but evidence on the comparative benefits and harms of these agents is contradictory.

Purpose: To compare the benefits and harms of second-generation antidepressants for treating MDD in adults.

Data Sources: English-language studies from PubMed, Embase, the Cochrane Library, PsycINFO, and International Pharmaceutical Abstracts from 1980 to August 2011 and reference lists of pertinent review articles and gray literature.

Study Selection: 2 independent reviewers identified randomized trials of at least 6 weeks’ duration to evaluate efficacy and observational studies with at least 1000 participants to assess harm.

Data Extraction: Reviewers abstracted data about study design and conduct, participants, and interventions and outcomes and rated study quality. A senior reviewer checked and confirmed extracted data and quality ratings.

Data Synthesis: Meta-analyses and mixed-treatment comparisons of response to treatment and weighted mean differences were conducted on specific scales to rate depression. On the basis of 234 studies, no clinically relevant differences in efficacy or effectiveness were detected for the treatment of acute, continuation, and maintenance phases of MDD. No differences in efficacy were seen in patients with accompanying symptoms or in subgroups based on age, sex, ethnicity, or comorbid conditions. Individual drugs differed in onset of action, adverse events, and some measures of health-related quality of life.

Limitations: Most trials were conducted in highly selected populations. Publication bias might affect the estimates of some comparisons. Mixed-treatment comparisons cannot conclusively exclude differences in efficacy. Evidence within subgroups was limited.

Conclusion: Current evidence does not warrant recommending a particular second-generation antidepressant on the basis of differences in efficacy. Differences in onset of action and adverse events may be considered when choosing a medication.

Primary Funding Source: Agency for Healthcare Research and Quality.

Ann Intern Med. 2011;155:772-785.

See also:
Print
Editors’ Notes

Web-Only
Appendix Tables
Appendix Figures
CME quiz (preview on page I-27)
Conversion of graphics into slides
Methods
An open process involving the public (described at www.effectivehealthcare.ahrq.gov/index.cfm/what-is-comparative-effectiveness-research1/what-is-the-research-process), the Scientific Resource Center for the Effective Health Care Program of the AHRQ, and various stakeholder groups produced key questions. We followed a standardized protocol for all review steps (17).

Data Sources and Searches
We searched PubMed, Embase, PsycINFO, the Cochrane Library, and International Pharmaceutical Abstracts from 1980 to August 2011. We used Medical Subject Heading terms as search terms when available or keywords when appropriate. We combined terms for MDD with a list of 13 second-generation antidepressants (bupropion, citalopram, desvenlafaxine, duloxetine, escitalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, sertraline, trazodone, and venlafaxine) and their trade names. We limited electronic searches to “adult 19 + years,” “human,” and “English language.” We also performed semiautomated manual searches of reference lists of pertinent review articles and letters to the editor by using Scopus (18).

Table 1. Second-Generation Antidepressants Approved for Use in the United States

| Generic Name | U.S. Trade Name* | Dosage Forms | Therapeutic Classification | Labeled Uses | Cost, $† |
|--------------|------------------|--------------|----------------------------|--------------|---------|
| Bupropion‡   | Wellbutrin, Wellbutrin SR, Wellbutrin XL | 75- or 100-mg tablets; 100-, 150-, or 200-mg SR tablets; 150- or 300-mg XL tablets | Other | MDD, seasonal affective disorder | 53–166 |
| Citalopram‡  | Celexa            | 10-, 20-, or 40-mg tablets; 2-mg/mL solution | SSRI | MDD | 31–38 |
| Desvenlafaxine| Pristiq          | 50- or 100-mg tablets | SNRI | MDD | – |
| Duloxetine   | Cymbalta          | 20-, 30-, or 60-mg capsules | SNRI | MDD, GAD, neuropathic pain, fibromyalgia | – |
| Escitalopram | Lexapro           | 5-, 10-, or 20-mg tablets; 1-mg/mL solution | SSRI | MDD, OCD, PMDD, panic disorder, bulimia nervosa | – |
| Fluoxetine‡  | Prozac, Prozac Weekly | 10-, 20-, 40-, or 90-mg capsules; 4-mg/mL solution | SSRI | MDD | 22–136 |
| Fluvoxamine‡ | Luvox             | 25-, 50-, or 100-mg tablets | SSRI | OCD | 99–106 |
| Mirtazapine‡ | Remeron, Remeron SolTab | 15-, 30-, or 45-mg tablets; 15-, 30-, or 45-mg orally disintegrating tablets | Other | MDD | 44–77 |
| Nefazodone‡  | Serzone           | 50-, 100-, 150-, 200-, or 250-mg tablets | Other | MDD | – |
| Paroxetine‡  | Paxil, Paxil CR  | 10-, 20-, 30-, or 40-mg tablets; 2-mg/mL solution; 12.5-, 25-, or 37.5-mg CR tablets | SSRI | MDD, OCD, panic disorder, social anxiety disorder, GAD, PTSD, PMDD§ | 20–115 |
| Sertraline‡  | Zoloft            | 25-, 50-, or 100-mg tablets; 20-mg/mL solution | SSRI | MDD, OCD, panic disorder, PTSD, PMDD, social anxiety disorder | 28–29 |
| Trazodone‡   | Desyrel           | 50-, 100-, 150-, or 300-mg tablets | Other | MDD | NR |
| Venlafaxine‡ | Effexor, Effexor XR | 25-, 37.5-, 50-, 75-, or 100-mg tablets; 37.5-, 75-, or 150-mg XR capsules | SNRI | MDD, GAD, panic disorder, social anxiety disorder| 88–129 |

CR = controlled release; GAD = generalized anxiety disorder; MDD = major depressive disorder; NR = not reported; OCD = obsessive-compulsive disorder; PMDD = premenstrual dysphoric disorder; PTSD = posttraumatic stress disorder; SNRI = serotonin and norepinephrine reuptake inhibitor; SR = sustained release; SSRI = selective serotonin reuptake inhibitor; XL = extended release; XR = extended release.
* Wellbutrin, Wellbutrin SR, Wellbutrin XL, Paxil, and Paxil CR (GlaxoSmithKline, Middlesex, United Kingdom); Celexa and Lexapro (Forest Laboratories, New York, New York); Pristiq, Zoloft, Effexor, and Effexor XR (Pfizer, New York, New York); Cymbalta, Prozac, and Prozac Weekly (Eli Lilly, Indianapolis, Indiana); Luvox (Jazz Pharmaceuticals, Palo Alto, California); Remeron and Remeron SolTab (Merck, Whitehouse Station, New Jersey); and Serzone and Desyrel (Bristol-Myers Squibb, Princeton, New Jersey).
† Cost estimates are ranges for various formulations and dosages of the same drug. From Consumer Reports Best Buy Drugs. The Antidepressants: Treating Depression. Comparing Effectiveness, Safety, and Price. April 2011. Accessed at www.consumerreports.org/health/resources/pdf/best-buy-drugs/Antidepressants_update.pdf.
‡ Generic available for some dosage forms.
§ Only Paxil CR (not Paxil) is approved for the treatment of PMDD.
¶ Only Effexor XR (not Effexor) is approved for the treatment of GAD and social anxiety disorder.

Context
Multiple second-generation antidepressants with different pharmacologic actions are available for treating major depressive disorder in adults.

Contribution
This comparative effectiveness review of 234 studies found no clinically important differences in treatment response among second-generation antidepressants. Differences among agents did exist in onset of action, dosing regimens, and adverse effects.

Caution
Most studies were efficacy trials conducted in selected populations.

Implication
Possible side effects, convenience of dosing regimens, and costs may best guide the choice of a second-generation antidepressant for treating major depression in adults, because these agents probably have similar efficacy.

—The Editors
The Scientific Resource Center searched the following sources for potentially relevant unpublished literature: the U.S. Food and Drug Administration (FDA) Web site, Health Canada, Authorized Medicines for the European Union, ClinicalTrials.gov, Current Controlled Trials, Clinical Study Results, the World Health Organization International Clinical Trials Registry Platform, Conference Papers Index, the National Institutes of Health Research Portfolio Online Reporting Tools, the U.S. National Library of Medicine Health Services Research Projects in Process, the Hayes Health Technology Assessment, and the New York Academy of Medicine’s gray literature index. The Scientific Resource Center also invited pharmaceutical manufacturers to submit dossiers on completed research for each drug. We received dossiers from AstraZeneca (London, United Kingdom) and Warner Chilcott (Dublin, Ireland).

**Study Selection**

Two persons independently reviewed abstracts and full-text articles. Studies reported only in abstract form were excluded. To assess efficacy or effectiveness, we included head-to-head randomized, controlled trials (RCTs) of at least 6 weeks’ duration that compared 2 drugs. Because many comparisons lacked head-to-head evidence, we included placebo-controlled trials for indirect comparisons. All outcomes of interest were health-related (for example, response, remission and quality of life).

To specifically assess harms, we examined RCTs as well as data from observational studies with 1000 participants or more and a follow-up of 12 weeks or more. To determine the differences of benefits and harms in subgroups and participants with accompanying symptoms, we reviewed head-to-head and placebo-controlled trials. We included meta-analyses if we believed them to be relevant for a key question and of good or fair methodological quality.

We excluded studies that both reviewers agreed did not meet eligibility criteria. Investigators resolved disagreements about inclusion or exclusion by consensus or by involving a third reviewer.

**Data Extraction and Quality Assessment**

Trained reviewers abstracted data from each study and assigned an initial quality rating by using the Web-based data abstraction form SRSNexus, version 4.0 (Mobius Analytics, Ottawa, Ontario, Canada). A senior reviewer evaluated completeness of data abstraction and confirmed the quality rating.

To assess trial quality (risk for bias), we used predefined criteria based on those developed by the U.S. Preventive Services Task Force (ratings of good, fair, or poor) (20) and the National Health Service Centre for Reviews and Dissemination (21). To assess the quality of observational studies, we used criteria outlined by Deeks and colleagues (22). We rated studies with a high risk for bias in 1 or more categories as “poor” quality and excluded them from the analyses.

To identify effectiveness studies, we used a tool that distinguishes them from efficacy trials on the basis of certain elements of study design (23). To evaluate the comparability of drug doses, we considered a large range of doses within and across studies. Because no reference standard exists for comparing doses among drugs, we had previously created a comparative dose classification system to identify gross inequities in comparisons of drug doses (24). We used this roster, which does not indicate dosing equivalence, to detect inequalities in dosing that could affect comparative efficacy and effectiveness.

**Data Synthesis and Analysis**

We conducted meta-analyses of head-to-head comparisons if 3 or more studies provided data to calculate either the odds ratio (OR) of achieving response (defined as ≥50% improvement from baseline) or the weighted mean difference of changes on the Hamilton Rating Scale for Depression (HAM-D) or the Montgomery-Asberg Depression Rating Scale (MADRS).

For each meta-analysis, we tested for heterogeneity by using the Cochran Q test and estimated the extent of heterogeneity by using the I² statistic. If heterogeneity was high (>60%), we explored differences in clinical and methodological characteristics among studies considered for meta-analyses. We assessed publication bias by using funnel plots and Kendall τ rank correlation.

Lacking head-to-head evidence for many drug comparisons, we conducted mixed-treatment comparisons of head-to-head and placebo-controlled trials by using Bayesian methods (25, 26). Because of clinical heterogeneity, we did not include studies conducted in patients older than 65 years. Our outcome measure of choice was the rate of response on the HAM-D. We recalculated response rates for each study by using the number of all randomly assigned patients as the denominator.

We gave all drug effect parameters flat normal (0, 1000) priors and gave the between-study SD flat, uniform distributions with a large range. We discarded a burn-in of 20 000 simulations. All results are based on a further sample of 80 000 simulations. We calculated the OR and 95% credible interval (CrI) for all possible comparisons among our drugs of interest.

All statistical analyses were performed by using StatsDirect Statistical Software, version 2.7.7 (StatsDirect, Cheshire, United Kingdom). We computed Bayesian inferences by using a Markov-chain Monte Carlo simulation with WinBUGS, version 1.4.3 (Medical Research Council Biostatistical Unit, Cambridge, United Kingdom). We evaluated the strength of evidence for major comparisons and outcomes by using a modified Grading of Recommendations Assessment, Development and Evaluation approach (27).
Role of the Funding Source

The AHRQ participated in formulating the key questions and reviewed planned methods and data analyses, as well as interim and final evidence reports. The AHRQ had no role in study selection, quality ratings, and interpretation in or synthesis of the evidence.

RESULTS

Our searches identified 3927 citations (Appendix Figure 1, available at www.annals.org). We included 234 studies of good or fair quality, of which 118 were head-to-head RCTs presented in report form at www.effectivehealthcare.ahrq.gov. Pharmaceutical companies financially supported most of the studies (77%), governmental agencies or independent funds supported 7%, and undetermined sources funded 16%. Funnel plots of head-to-head trials did not indicate publication bias.

Overall, comparative efficacy and effectiveness of second-generation antidepressants did not differ substantially for treating patients with MDD. These findings pertain to patients in the acute, continuation, and maintenance phases of this condition; those with accompanying symptom clusters; and subgroups defined by age, sex, ethnicity, or comorbid conditions, although only sparse evidence for these findings exists for subgroups. Overall, 37% of patients with acute-phase MDD who received first-line treatment did not achieve response within 6 to 12 weeks, and 53% did not achieve remission.

Comparative Efficacy for Acute-Phase Treatment of MDD

Ninety-three good- or fair-quality head-to-head trials that included more than 20,000 patients compared the efficacy or effectiveness of the treatment of acute-phase MDD. These studies provided direct evidence for 40 of 78 possible comparisons among these drugs. Direct evidence from head-to-head trials was sufficient to conduct meta-analyses for 6 drug–drug comparisons. In addition, we conducted mixed-treatment comparisons of response rates for all comparisons, incorporating 64 placebo-controlled or head-to-head trials.

Overall, treatment effects were similar among second-generation antidepressants (Table 2). Some analyses yielded statistically significant differences among treatments, but the magnitudes of differences were modest and probably not clinically relevant.

Meta-analyses of head-to-head trials showed statistically significantly greater response rates for escitalopram than citalopram (1 unpublished study [28] and 5 published studies [29–33] involving 1802 patients [OR, 1.49 [95% CI, 1.07 to 2.01]]) (Figure 1), and venlafaxine than fluoxetine (6 studies [34–37] involving 960 patients [OR, 1.42 [CI, 1.08 to 1.85]]) (Figure 1), and venlafaxine than fluoxetine (6 studies [38–43] involving 1197 patients [OR, 1.47 [CI, 1.16 to 1.86]]) (Figure 2).

The 2 largest relative differences in response rates were between escitalopram and citalopram and fluoxetine and venlafaxine, but absolute differences were modest. On average, 62% of patients receiving escitalopram and 56% receiving citalopram achieved a response. The pooled difference of the reduction of points on the MADRS scale was 1.52 in favor of escitalopram (CI, 0.59 to 2.45 points), which is approximately one sixth of the average SD of change on the MADRS scale in trials.

The additional benefit of venlafaxine versus fluoxetine was similarly modest. On average, 65% of patients receiving venlafaxine and 60% receiving fluoxetine achieved a response. Pooled results of reductions of points on the HAM-D showed a non–statistically significant 1.30-point greater reduction for patients receiving venlafaxine versus fluoxetine (CI, 0.32 lower reduction to 2.92 greater reduction).

Mixed-treatment comparisons of drugs (Figures 1 to 3 and Appendix Figure 2, available at www.annals.org) did not show statistically significant differences in response rates except for escitalopram over duloxetine (OR, 0.74 [95% CrI, 0.56 to 0.98]) and escitalopram over fluoxetine (OR, 0.66 [CrI, 0.49 to 0.89]).

Seventeen studies (n = 3960) indicated no differences in health-related quality of life (Table 2) (30, 37, 41, 44–47, 49–58) Seven studies, all funded by the maker of mirtazapine, reported that this agent has a significantly faster onset of action than some comparators (49, 50, 55, 59–62). After 4 weeks of treatment, most response rates among the drugs studied were similar. In 1 trial, mirtazapine and venlafaxine did not differ in speed of action (52).

Achieving Response in Unresponsive or Recurrent Disease

Overall, 37% of patients did not achieve a treatment response during 6 to 12 weeks of treatment with second-generation antidepressants, and 53% did not achieve remission. The STAR*D (Sequenced Treatment Alternatives to Relieve Depression) trial (63) provides the best evidence for assessing alternative medications among patients in whom initial therapy has failed. Approximately 1 in 4 of the 727 participants who switched medications after initial treatment failure became symptom-free; however, no statistically significant difference was seen in patients who switched to sustained-release bupropion, sertraline, or extended-release venlafaxine. In 3 additional head-to-head trials involving patients with treatment-resistant depression, response and remission rates were numerically better with venlafaxine than with comparators (64–67), but differences generally were not statistically significant.

Maintaining Response or Remission After Successful Treatment

In several head-to-head trials (68–75), overall efficacy in maintaining remission did not significantly differ between escitalopram and desvenlafaxine (74), escitalopram and paroxetine (72), fluoxetine and sertraline (68), fluox-
Comparative Benefits and Harms of Second-Generation Antidepressants

Table 2. Comparative Efficacy and Effectiveness of Second-Generation Antidepressants: Findings and Strength of Evidence

| Outcome | Strength of Evidence | Findings |
|---------|----------------------|----------|
| **Treating acute-phase MDD** | | |
| Comparative efficacy | Moderate | Results from direct and indirect comparisons indicate that clinical response and remission rates are similar among second-generation antidepressants. |
| Comparative effectiveness | Moderate | One good-quality and 2 fair-quality effectiveness studies indicate that no substantial differences in effectiveness exist among second-generation antidepressants. |
| Quality of life | Moderate | Consistent results from 17 mostly fair-quality studies indicate that the efficacy of second-generation antidepressants regarding quality of life does not differ among drugs. |
| Onset of action | Moderate | Consistent results from 7 fair-quality trials suggest that mirtazapine has a statistically significantly faster onset of action than citalopram, fluoxetine, paroxetine, and sertraline. Whether this difference favoring mirtazapine can be extrapolated to other second-generation antidepressants is unclear. Most other trials do not indicate a faster onset of action of a particular second-generation antidepressant compared with another. |
| **Maintaining response or remission†** | | |
| Comparative efficacy | Moderate | Findings from 5 efficacy trials and 1 naturalistic study show no statistically significant differences in preventing relapse or recurrence between escitalopram and paroxetine, fluoxetine and sertraline, fluoxetine and venlafaxine, fluvoxamine and sertraline, and trazodone and venlafaxine. |
| **Managing treatment-resistant depression** | | |
| Comparative efficacy | Low | Results from 3 trials support modestly better efficacy for venlafaxine compared with citalopram, fluoxetine, and paroxetine. |
| Comparative effectiveness | Low | Results from 2 effectiveness studies are conflicting. One good-quality trial showed no statistically significant differences in effectiveness among sustained-release bupropion, sertraline, and extended-release venlafaxine. One fair-quality effectiveness trial found venlafaxine to be modestly superior to citalopram, fluoxetine, mirtazapine, paroxetine, and sertraline; however, differences may not be clinically relevant. |
| **Treating depression in patients with accompanying symptom clusters** | | |
| Anxiety | | |
| Comparative efficacy for depression | Moderate | Results from 5 fair-quality head-to-head trials suggest that efficacy does not differ substantially for treatment of depression in patients with accompanying anxiety. |
| Comparative efficacy for anxiety | Moderate | Results from 8 fair-quality head-to-head trials and 3 fair-quality placebo-controlled trials suggest that no substantial differences in efficacy exist among second-generation antidepressants for treatment of accompanying anxiety. |
| Insomnia | | |
| Comparative efficacy for depression | Insufficient | Evidence from 1 fair-quality head-to-head study is insufficient to draw conclusions about the comparative efficacy for treating depression in patients with coexisting insomnia. |
| Comparative efficacy for insomnia | Low | Evidence from 5 fair-quality head-to-head trials suggests that no substantial differences in efficacy exist among second-generation antidepressants for treatment of accompanying insomnia. Results are limited by study design; differences in outcomes are of unknown clinical significance. |
| Low energy | | |
| Comparative efficacy for depression | Insufficient | Evidence from 1 placebo-controlled trial of bupropion XL is insufficient to draw conclusions about treating depression in patients with coexisting low energy. |
| Comparative efficacy for low energy | Insufficient | Evidence from 1 placebo-controlled trial of bupropion XL is insufficient to draw conclusions about treating low energy in patients with depression. |
| Melancholia | | |
| Comparative efficacy for depression | Insufficient | Evidence from 2 fair-quality head-to-head studies is insufficient to draw conclusions about treating depression in patients with coexisting melancholia. Results are inconsistent across studies. |
| Comparative efficacy for melancholia | No evidence | – |
| Pain | | |
| Comparative efficacy for depression | Insufficient | Evidence from 2 fair-quality placebo-controlled studies is conflicting about the superiority of duloxetine over placebo. Results from head-to-head trials are not available. |
| Comparative efficacy for pain | Moderate | Evidence from 1 systematic review, 2 head-to-head trials (1 fair-quality trial and 1 poor-quality trial), and 5 placebo-controlled trials indicate no difference in efficacy between paroxetine and duloxetine. |
| Psychomotor changes | | |
| Comparative efficacy for depression | Insufficient | Evidence from 1 fair-quality head-to-head trial is insufficient to draw conclusions about the comparative efficacy for treating depression in patients with coexisting psychomotor change. Results indicate that comparative outcomes for psychomotor retardation and psychomotor change may differ. |
| Comparative efficacy for psychomotor change | No evidence | – |
| Somatization | | |
| Comparative efficacy for depression | No evidence | – |
| Comparative efficacy for somatization | Insufficient | Evidence from 1 randomized, head-to-head trial is insufficient to draw conclusions about the comparative efficacy for treating somatization in patients with depression. Results indicate similar improvement in somatization. |

MDD = major depressive disorder; XL = extended release.

* High strength of evidence indicates high confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of effect. Moderate strength of evidence indicates that the evidence reflects the true effect. Further research may change our confidence in the estimate of effect and may change the estimate. Low strength of evidence indicates that the evidence reflects the true effect. Further research is likely to change both the confidence in the estimate of effect and the estimate. Insufficient strength of evidence indicates that evidence is either unavailable or does not permit a conclusion.

† Preventing relapse or recurrence.
etine and venlafaxine (73, 75), fluvoxamine and sertraline (69, 70), and trazodone and venlafaxine (71). One of these studies reported a significantly shorter time to recurrence with fluoxetine than with venlafaxine during 2 years of maintenance treatment (75). In one naturalistic study, rehospitalization rates did not differ between patients continuing therapy with fluoxetine versus venlafaxine (76).

Efficacy or Effectiveness in Treating Depression or Accompanying Symptoms

Clinicians may use symptom clusters that accompany depression (for example, anxiety and insomnia) to guide antidepressant selection. We identified studies addressing 7 symptom clusters: anxiety, insomnia, low energy, pain, psychomotor change (retardation or agitation), melancholia (a subtype of depression that is a severe form of MDD with characteristic somatic symptoms), and somatization (physical symptoms that are manifestations of depression rather than of an underlying physical illness). Table 2 summarizes these findings.

![Figure 1. Odds ratios of response rates comparing selective serotonin reuptake inhibitors.](image)

h-h = head-to-head; MA = meta-analysis; MTC = mixed-treatment comparison.

* The first number indicates the number of trials comparing 2 drugs; the second indicates the number of additional studies used to perform MTCs.

pressants have similar antidepressive efficacy. Two of these studies compared SSRIs (fluoxetine, paroxetine, and sertraline) (77, 78), 1 compared sertraline and sustained-release bupropion (79), and 1 compared sertraline and extended-release venlafaxine (45). One study reported a greater decrease in severity of depression and higher response rates with venlafaxine than with fluoxetine (75% vs. 49%) (39).

For other symptom clusters, such as insomnia (35), melancholia (78, 80), or psychomotor changes (78), most studies indicated similar treatment effects for depression among compared drugs. Because these studies were small or had conflicting results, the strength of the evidence is low.

Treatment of Accompanying Symptom Clusters in Patients With Depression

Results from 8 head-to-head trials suggested that antidepressant medications do not differ in efficacy for treating anxiety associated with MDD. Among these studies, 4 compared SSRIs (including escitalopram, fluoxetine, sertraline, and paroxetine) (77, 81–83); 3 compared paroxetine and nefazodone (84), citalopram and mirtazapine (50), and sertraline and sustained-release
bupropion (79); and 1 compared extended-release venlafaxine and sertraline (45). Only 1 trial (146 participants) reported that patients receiving venlafaxine had statistically significantly greater reductions in Covi Anxiety Scale scores (5.7 vs. 3.9) than those receiving fluoxetine (39).

For insomnia, 2 studies suggested greater improvement in sleep scores with trazodone than with fluoxetine (47) and venlafaxine (71). In 3 other studies, rates of insomnia did not significantly differ in patients receiving escitalopram or fluoxetine (83); fluoxetine, paroxetine, or sertraline (35); or fluoxetine or mirtazapine (55). A well-conducted meta-analysis (85) of 3 fair-quality head-to-head trials (86–88) and 1 poor-quality trial (89) (1466 participants) found no substantial difference between duloxetine and paroxetine in the relief of accompanying pain.

**Risk for Harms**

We analyzed 93 head-to-head studies and 48 additional studies of both experimental and observational design to assess the comparative risk for harm. We distinguished adverse events from serious adverse events on the basis of an FDA classification. A serious adverse event is any medical occurrence that results in death, is life-threatening, requires hospitalization, results in persistent or substantial disability or incapacity, or is a congenital birth defect (90). Table 3 summarizes these findings.
**Adverse Events and Discontinuation of Therapy**

In efficacy trials, an average of 63% of patients experienced at least 1 adverse event during treatment. Diarrhea, dizziness, dry mouth, fatigue, headache, nausea, sexual dysfunction, sweating, tremor, and weight gain were commonly reported. Overall, second-generation antidepressants caused similar adverse events; however, the frequency of specific events differed among some drugs ([Appendix Table 1](#), available at www.annals.org).

Overall discontinuation rates were similar between SSRIs and other second-generation antidepressants (range of means, 15% to 25%). Duloxetine had a 67% (CI, 17% to 139%) and venlafaxine had a 40% (CI, 16% to 73%) higher risk for discontinuation of therapy because of adverse events than SSRIs as a class did. Discontinuation rates due to lack of efficacy were similar between SSRIs and other second-generation antidepressants except for venlafaxine. Venlafaxine had a 34% (CI, 47 to 93) lower risk for discontinuation of therapy because of lack of efficacy than SSRIs did.

**Serious Adverse Events**

Except for sexual dysfunction, trials and observational studies were too small and than durations were too short to

---

**Figure 3. Odds ratios of response rates comparing SSRIs with other second-generation antidepressants.**

| Comparison                  | Odds Ratio (95% CI) | Study Type |
|-----------------------------|---------------------|------------|
| SSRIs vs. other second-generation antidepressants |                      |            |
| Citalopram vs. bupropion    | 0.56 (0.28–1.29)    | MTC of 0/64 trials* |
| Citalopram vs. mirtazapine  | 0.73 (0.35–1.83)    | MTC of 1/63 trials* |
| Citalopram vs. nefazodone   | 0.64 (0.30–1.63)    | MTC of 0/64 trials* |
| Citalopram vs. trazodone    | 0.49 (0.23–1.26)    | MTC of 0/64 trials* |
| Escitalopram vs. bupropion  | 0.74 (0.50–1.06)    | MTC of 0/64 trials* |
| Escitalopram vs. mirtazapine| 0.99 (0.57–1.59)    | MTC of 0/64 trials* |
| Escitalopram vs. nefazodone | 0.87 (0.48–1.45)    | MTC of 0/64 trials* |
| Escitalopram vs. trazodone  | 0.67 (0.38–1.10)    | MTC of 0/64 trials* |
| Fluoxetine vs. bupropion    | 1.11 (0.81–1.48)    | MTC of 2/62 trials* |
| Fluoxetine vs. mirtazapine  | 1.48 (0.91–2.28)    | MTC of 3/61 trials* |
| Fluoxetine vs. nefazodone   | 1.30 (0.76–2.10)    | MTC of 1/63 trials* |
| Fluoxetine vs. trazodone    | 1.00 (0.60–1.57)    | MTC of 2/62 trials* |
| Fluvoxamine vs. bupropion   | 0.59 (0.26–1.65)    | MTC of 0/64 trials* |
| Fluvoxamine vs. mirtazapine | 0.77 (0.33–2.32)    | MTC of 0/64 trials* |
| Fluvoxamine vs. nefazodone  | 0.67 (0.28–2.05)    | MTC of 0/64 trials* |
| Fluvoxamine vs. trazodone   | 0.52 (0.22–1.58)    | MTC of 0/64 trials* |
| Paroxetine vs. bupropion    | 0.93 (0.65–1.28)    | MTC of 2/62 trials* |
| Paroxetine vs. mirtazapine  | 1.24 (0.77–1.89)    | MTC of 3/61 trials* |
| Paroxetine vs. nefazodone   | 1.09 (0.62–1.78)    | MTC of 2/62 trials* |
| Paroxetine vs. trazodone    | 0.84 (0.50–1.32)    | MTC of 1/63 trials* |
| Sertraline vs. bupropion    | 0.90 (0.67–1.20)    | MTC of 3/61 trials* |
| Sertraline vs. mirtazapine  | 1.21 (0.73–1.88)    | MTC of 1/63 trials* |
| Sertraline vs. nefazodone   | 1.06 (0.63–1.69)    | MTC of 1/63 trials* |
| Sertraline vs. trazodone    | 0.82 (0.49–1.28)    | MTC of 1/63 trials* |

MTC = mixed-treatment comparison; SSRI = selective serotonin reuptake inhibitor.

* The first number indicates the number of trials directly comparing 2 drugs; the second indicates the number of additional studies used to perform MTCs.
assess the comparative risks for rare but serious adverse events, such as suicidality, seizures, cardiovascular events, the serotonin syndrome, hyponatremia, or hepatotoxicity.

**Sexual Dysfunction.** Five trials and a pooled analysis (2399 participants) of 2 identical RCTs provided evidence that bupropion causes lower rates of sexual dysfunction than escitalopram (91), fluoxetine (92), paroxetine (93), and sertraline (94–96). Compared with other second-generation antidepressants, paroxetine frequently caused higher rates of sexual dysfunction, particularly ejaculatory dysfunction. These differences, however, did not always reach statistical significance (35, 44, 60, 81, 97–101).

Underreporting of sexual dysfunction in efficacy studies is likely. A fair-quality Spanish prospective, observational study (102 participants) reported that 59% of patients treated with second-generation antidepressants experienced sexual dysfunction (102).

**Suicidality.** Although suicide is relatively rare and affects approximately 1 in 8000 psychiatric patients treated with second-generation antidepressants, 1 in 166 patients reported suicidal feelings while receiving treatment with a second-generation antidepressant (103).

Thirteen studies assessed the risk for suicidality (defined as suicidal thinking or behavior) in patients treated with second-generation antidepressants (104–116). Data on the comparative risk for suicidality among second-generation antidepressants were sparse. Results from existing studies did not indicate that any particular drug of interest had an excess risk compared with other second-generation antidepressants (106–109, 113, 116).

Several large observational studies determined that second-generation antidepressants cause a general increase in the risk for suicidality (106, 107, 116). A recent meta-analysis of observational studies in a combined population...
Treatment of Major Depressive Disorder in Subgroups

No study directly compared efficacy, effectiveness, and harms of second-generation antidepressants between subgroups and the general population for treatment of MDD. However, numerous studies conducted subgroup analyses or used subgroups as the study population (Appendix Table 2, available at www.annals.org).

Multiple head-to-head trials (36, 58, 117–125) indicated that the efficacy of second-generation antidepressants did not differ in participants aged 55 years or older. Efficacy trials usually did not address differences in efficacy or effectiveness between men and women. Two head-to-head RCTs provided limited evidence on adverse sexual effects of these agents; 1 reported a higher risk for sexual dysfunction in men than in women receiving paroxetine (93), and the other reported greater sexual dysfunction in women receiving paroxetine than in those receiving sertraline (44).

No head-to-head trials or other studies directly compared differences in efficacy, effectiveness, and harms among groups identified by race or ethnicity or between patients with depression and comorbid conditions and the general population. One recent RCT reported no differences between citalopram and fluoxetine in participants with type 2 diabetes and MDD (126).

Discussion

In this systematic review of data from 234 studies, direct and indirect comparisons of second-generation antidepressants showed no substantial differences in efficacy for the treatment of MDD. Statistically significant results were small and are unlikely to have clinical relevance. No differences in efficacy were seen in patients with accompanying symptoms or in subgroups based on age, sex, ethnicity, or comorbid conditions.

Although second-generation antidepressants are similar in efficacy, they cannot be considered identical drugs. Differences with respect to onset of action, adverse events, and some measures of health-related quality of life may be clinically relevant and influence the choice of a medication for a specific patient. For example, mirtazapine has a faster onset of action than citalopram, fluoxetine, paroxetine, and sertraline (49, 55, 60–62), whereas bupropion has fewer sexual side effects than escitalopram, fluoxetine, paroxetine, and sertraline (91, 92, 94, 96, 127).

Our findings are consistent with results of most other systematic reviews assessing the comparative efficacy and safety of second-generation antidepressants (8–14). Our conclusions contradict some findings of the 2009 MANGA study, which indicated that escitalopram and sertraline have the best efficacy–acceptability ratio compared with that of other agents (15). The MANGA study, however, has been criticized for methodological shortcomings (128–132). Specifically, the authors included studies with a high risk for bias and open-label designs, assumed that a response on the HAM-D equals a response on MADRS or the Clinical Global Inventory, excluded placebo-controlled trials in their network meta-analysis, and overstated the importance of statistically significant findings without considering clinical relevance. In particular, the assumption that responses on different scales are comparable is not evidence-based (133) and thus might introduce substantial bias in a mixed-treatment comparison model.

For the current update of our review, we used the same statistical methods as the authors of the MANGA study, although we retained more rigid systematic review methods. We specifically excluded studies with high risk for bias or open-label designs and limited mixed-treatment comparisons to ORs of response on a single diagnostic scale (HAM-D). Furthermore, whenever possible, we used meta-analyses of head-to-head trials to determine the relative efficacy.

Our study has several limitations. Most important, we primarily derived our conclusions from efficacy trials with highly selected populations. For example, for data on acute-phase MDD, we found only 3 effectiveness studies (37, 120, 134) out of 93 head-to-head RCTs. Two of these effectiveness studies were conducted in Europe, and their applicability to the U.S. health care system might be limited. Although findings from effectiveness studies are generally consistent with those from efficacy trials, the evidence is limited to a few comparisons.

Indirect comparisons have methodological limitations, most prominently the assumption that prognostic factors for a specific outcome (for example, response to treatment) are similar across study populations in the network meta-analyses. Nevertheless, they are a valuable additional analytic tool when available head-to-head evidence is insufficient.

Publication bias is a concern for all systematic reviews and has been empirically proven to be problematic for placebo-controlled trials of second-generation antidepressants (135, 136). Selective availability of studies with positive results can seriously bias conclusions, particularly when a pharmaceutical company compares 2 of its own drugs (as in the case of citalopram and escitalopram). The small number of studies for individual comparisons limits the validity of statistical methods to explore publication bias, such as funnel plots.
How do these findings that pharmacologic differences among second-generation antidepressants do not translate into substantial clinical differences, although tolerability may differ, inform the practicing clinician? Given the difficulty in predicting what medication will be both efficacious for and tolerated by an individual patient, familiarity with a broad spectrum of antidepressants is prudent. Existing evidence of efficacy, however, does not warrant choosing a particular second-generation antidepressant as first-line therapy for acute-phase MDD or as a subsequent treatment in patients who do not respond to therapy or experience remission. Because of differences in adverse events and dosing regimens, engaging in informed decision making can help physicians to take patient preferences into consideration.

From Danube University, Krems, Austria; Auburn University, Auburn, Alabama; RTI International, Research Triangle Park, North Carolina; the University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; and the Ludwig Boltzmann Institute for Health Promotion Research, Vienna, Austria.

Disclaimer: This manuscript and the work it is derived from was commissioned by the Agency for Healthcare Research and Quality (AHRQ), through a contract to the x2 Evidence-based Practice Center (contract 290200710056I#2). While AHRQ has approved the assertion of copyright by the authors, as noted in the attached letter from the AHRQ Contracting Officer, the government retains rights to the use of the manuscript according to the contract and the Federal Acquisition Regulations (FAR). In order to facilitate and meet the need of public access to research works and findings funded by the government, AHRQ will publish the full report from which this manuscript is derived. The original report from which this manuscript was derived has undergone rigorous peer and public review through the Effective Health Care Program.

Acknowledgment: The authors thank Shrinkant Bangdiwala, Visali Perival, Loraine Monroe, Irene Wild, and Evelyn Auer.

Financial Support: By contract 290200710056I#2 from the AHRQ to RTI International.

Potential Conflicts of Interest: Dr. Garthlehner: Grant (money to institution): AHRQ; Dr. Hansen: Grant (money to institution): AHRQ; Consultancy: Novartis and Takeda Pharmaceutical Company; Grants/Grants pending (money to institution): AHRQ, North Carolina Department of Health and Human Services, and the Foundation for the National Institutes of Health; Payment for development of educational presentations: WebMD. Ms. Morgan: Grant (money to institution): AHRQ; Dr. Thaler: Grant (money to institution): AHRQ; Ms. Lux: Grant (money to institution); AHRQ, Dr. Mager and Ms. Strobelberger: Grant (money to institution): Danube University. Ms. Thieda: Grant (money to institution): AHRQ; Dr. Gaynes: Grant (money to institution): AHRQ; Grants/Grants pending (money to institution): National Institute of Mental Health. Ms. Lloyd: Grant (money to institution): AHRQ. Disclosures can also be viewed at www.acponline.org/authors/scmej/ConflictOfInterestForms.do?msNum=M11-1590.

Requests for Single Reprints: Gerald Garthlehner, MD, MPH, Danube University, Karl Dorrek-Straße 30, 3500 Krems, Austria; e-mail, gerald.garthlehner@donau-uni.ac.at.

Current author addresses and author contributions are available at www.annals.org.

References

1. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry. 2005;62:593-602. [PMID: 15939837]

2. Wu E, Greenberg P, Yang E, Yu A, Ben-Hamadi R, Ezer MH. Comparison of treatment persistence, hospital utilization and costs among major depressive disorder geriatric patients treated with escitalopram versus other SSRI/SNRI antidepressants. Curr Med Res Opin. 2008;24:2805-13. [PMID: 18755054]

3. Birnbaum HG, Ben-Hamadi R, Greenberg PE, Hsieh M, Tang J, Reygrobellet C. Determinants of direct cost differences among US employees with major depressive disorders using antidepressants. Pharmacoconomics. 2009;27:507-17. [PMID: 19640013]

4. Olsson M, Marcus SC. National patterns in antidepressant medication treatment. Arch Gen Psychiatry. 2009;66:948-56. [PMID: 19652124]

5. Berkrot B. U.S. prescription drug sales hit $300 bn in 2009. Thomson Reuters. 1 April 2010. Accessed at www.reuters.com/article/2010/04/01/us-drug-sales-ISTRE6303CU20100401 on 10 April 2010.

6. Hansen RA, Garthlehner G, Lohr KN, Gaynes BN, Carey TS. Efficacy and safety of second-generation antidepressants in the treatment of major depressive disorder. Ann Intern Med. 2005;143:415-26. [PMID: 16172440]

7. Williams JW Jr, Murrow CD, Chiouette E, Noël PH, Aguilar C, Cornell J. A systematic review of newer pharmacotherapies for depression in adults: evidence report summary. Ann Intern Med. 2000;133:743-56. [PMID: 10787370]

8. Cipriani A, Furukawa TA, Geddes JR, Malvini L, Signoretti A, McGuire H, et al; MANGA Study Group. Does randomized evidence support sertraline as first-line antidepressant for adults with acute major depression? A systematic review and meta-analysis. J Clin Psychiatry. 2008;69:1732-42. [PMID: 19026250]

9. Omori IM, Watanabe N, Nakagawa A, Akechi T, Cipriani A, Barbui C, et al; Meta-Analysis of New Generation Antidepressants (MANGA) Study Group. Efficacy, tolerability and side-effect profile of fluoxamine for major depression: meta-analysis. J Psychopharmacol. 2009;23:539-50. [PMID: 18562407]

10. Watanabe N, Omori IM, Nakagawa A, Cipriani A, Barbui C, McGuire H, et al; Multiple Meta-Analyses of New Generation Antidepressants (MANGA) Study Group. Mirtazapine versus other antidepressants in the acute-phase treatment of adults with major depression: systematic review and meta-analysis. J Clin Psychiatry. 2008;69:1404-15. [PMID: 19193411]

11. Weinmann S, Becker T, Koesters M. Re-evaluation of the efficacy and tolerability of venlafaxine vs SSRI meta-analysis. Psychopharmacology (Berl). 2008;196:511-20; discussion 521-2. [PMID: 17955213]

12. Girardi P, Pompili M, Innamorati M, Mancini M, Serafini G, Mazzarini L, et al; Duloxetine in acute major depression: review of comparisons to placebo and standard antidepressants using dissimilar methods. Hum Psychopharmacol. 2009;24:177-90. [PMID: 19229839]

13. Eckert L, Falissard B. Using meta-regression in performing indirect-comparisons: comparing escitalopram with venlafaxine XR. Curr Med Res Opin. 2006;22:231-3. [PMID: 17076991]

14. Eckert L, Lançon C. Duloxetine compared with fluoxetine and venlafaxine: use of meta-regression analysis for indirect comparisons. BMC Psychiatry. 2006;6:30. [PMID: 16867188]

15. Cipriani A, Furukawa TA, Salanti G, Geddes JR, Higgins JP, Churchill R, et al. Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. Lancet. 2009;373:746-58. [PMID: 19185342]

16. Garthlehner G, Hansen RA, Thieda P, DeVeauagh-Geiss AM, Gaynes BN, Krebs EE, et al. Comparative Effectiveness of Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression. Rockville, MD: Agency for Healthcare Research and Quality; 2007.

17. Garthlehner G, Hansen RA, Morgan LC, Thaler K, Lux IJ, Van Noord M, et al. Comparative Effectiveness of Second Generation Antidepressants in the Pharmacologic Treatment of Adult Depression: An Update to a 2007 Report. Research Protocol. 13 July 2011. Rockville, MD: Agency for Healthcare Research and Quality; 2011. Accessed at www.effectivehealthcare.ahrq.gov/index.cfm?search=for-guides-reviews-and-reports?&pageaction=displayproduct&productId
24. Gaynes BN. Controlling for drug dose in systematic review and meta-analysis: a case study of the effect of antidepressant dose. Med Decis Making. 2009;29:91-103. [PMID: 19141788]

25. Jansen JP, Crawford B, Bergman G, Stam W. Bayesian meta-analysis of multiple treatment comparisons: an introduction to mixed treatment comparisons. Value Health. 2008;11:956-64. [PMID: 18489499]

26. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med. 2006;23:3105-24. [PMID: 15449338]

27. Owens DK, Lohr KN, Atkins D, Treadwell JR, Reston JT, Bass EB, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions—AHRQ and the effective health-care program. J Clin Epidemiol. 2001;64:510-23. [PMID: 11698043]

28. Hansen RA, Moore CG, Dusetzina SB, Leinwand BI, Gartlehner G, Raynes BN. Controlling for drug dose in systematic review and meta-analysis: a case study of the effect of antidepressant dose. Med Decis Making. 2009;29:91-103. [PMID: 19141788]

29. Jansen JP, Crawford B, Bergman G, Stam W. Bayesian meta-analysis of multiple treatment comparisons: an introduction to mixed treatment comparisons. Value Health. 2008;11:956-64. [PMID: 18489499]

30. Burke WJ, Gergel I, Buse A. Fixed-dose trial of the single isomer SSRI escitalopram in depressed outpatients. J Clin Psychiatry. 2002;63:316-6. [PMID: 12000207]

31. Colonna L, Andersen HF, Reines EH. A randomized, double-blind, 24-week study of escitalopram (10 mg/day) versus citalopram (20 mg/day) in primary care patients with major depressive disorder. Curr Med Res Opin. 2005;21:1659-68. [PMID: 16238906]

32. Lepola UM, Loft H, Reines EH. Escitalopram (10-20 mg/day) is effective and well tolerated in a placebo-controlled study in depression in primary care. Int Clin Psychopharmacol. 2003;18:211-7. [PMID: 12817155]

33. Morselli, Verdoux H, Fantino B. Prospective, multicentre, randomized, double-blind study of the efficacy of escitalopram versus citalopram in outpatient treatment of major depressive disorder. Int Clin Psychopharmacol. 2005;20:131-7. [PMID: 15812262]

34. Bennie EH, Mullin JM, Martindale JJ. A double-blind multicenter trial comparing sertraline and fluoxetine in outpatients with major depression. J Clin Psychiatry. 1995;56:229-37. [PMID: 7775364]

35. Fava M, Hoog SL, Judge RA, Kopp JB, Nilsson ME, Gonzales JS. Acute efficacy of escitalopram versus sertraline and paroxetine in major depressive disorder including effects of baseline insomnia. J Clin Psychopharmacol. 2002;22:137-47. [PMID: 11910258]

36. Newhouse PA, Krishnan KR, Doraiswamy PM, Richter EM, Batzar ED, Clary CM. A double-blind comparison of sertraline and fluoxetine in depressed elderly outpatients. J Clin Psychiatry. 2000;61:559-68. [PMID: 10982198]

37. Secher D, Spray H, Paternetti S, Boyer P. A double-blind comparison of sertraline and fluoxetine in the treatment of major depressive episode in outpa-
Comparative Benefits and Harms of Second-Generation Antidepressants

73. [PMID: 16528138] Dyscontrol of sertraline and fluoxetine in major depression. J Clin Psychiatry. 2006;67:736-46. [PMID: 17825800]

74. [PMID: 17825800] Venlafaxine and paroxetine in treatment-resistant depression. J Affect Disord. 2006;58:233-6. [PMID: 11030486]

75. [PMID: 15232330] Venlafaxine extended release versus citalopram in treatment-resistant depression. J Affect Disord. 2006;96:23-30. [PMID: 16589385]

76. [PMID: 16907153] Venlafaxine and paroxetine in the treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. Eur Neuropsychopharmacol. 2004;14:457-70. [PMID: 15589385]

77. [PMID: 16528138] Venlafaxine and paroxetine in treatment-resistant depression. J Affect Disord. 2006;96:23-30. [PMID: 16589385]

78. [PMID: 16528138] Venlafaxine and paroxetine in the treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. Eur Neuropsychopharmacol. 2004;14:457-70. [PMID: 15589385]

79. [PMID: 16528138] Venlafaxine and paroxetine in treatment-resistant depression. J Affect Disord. 2006;96:23-30. [PMID: 16589385]
Current Author Addresses: Drs. Gartlehner, Thaler, and Mager and Ms. Strobelberger: Danube University, Karl Dorrek-Straße 30, 3500 Krems, Austria.
Dr. Hansen: Department of Pharmacy Care Systems, Harrison School of Pharmacy, Auburn University, 207 Dunstan Hall, Auburn, AL 36849.
Ms. Morgan, Ms. Lux, Ms. Lloyd, and Dr. Lohr: RTI International, Research Triangle Park, 3040 Cornwallis Road, PO Box 12194, NC 27709.
Ms. Van Noord and Ms. Wilkins: Cecil G. Sheps Center for Health Services Research, The University of North Carolina at Chapel Hill, 725 Martin Luther King Jr. Boulevard, Chapel Hill, NC 27599.
Ms. Thieda: 226 Barclay Road, Chapel Hill, NC 27516.
Dr. Gaynes: Department of Psychiatry, University of North Carolina School of Medicine, Third Floor, Medical School Wing D, Campus Box 7160, Chapel Hill, NC 27599-7160.
Dr. Reichenpfader: Mühlgasse 61/36, 8020 Graz, Austria.

Author Contributions: Conception and design: G. Gartlehner, R.A. Hansen, L.C. Morgan, P. Thieda, B.N. Gaynes, K.N. Lohr.
Analysis and interpretation of the data: G. Gartlehner, R.A. Hansen, L.C. Morgan, K. Thaler, M. Van Noord, U. Mager, P. Thieda, B.N. Gaynes, T. Wilkins, S. Lloyd, U. Reichenpfader, K.N. Lohr.
Drafting of the article: G. Gartlehner, R.A. Hansen, L.C. Morgan, K. Thaler, P. Thieda, B.N. Gaynes, K.N. Lohr.
Critical revision for important intellectual content: R.A. Hansen, L.C. Morgan, K. Thaler, L. Lux, B.N. Gaynes, S. Lloyd, K.N. Lohr.
Final approval of the article: G. Gartlehner, R.A. Hansen, K. Thaler, U. Mager, B.N. Gaynes, S. Lloyd, K.N. Lohr.
Statistical expertise: G. Gartlehner, P. Thieda, B.N. Gaynes, T. Wilkins.
Obtaining of funding: G. Gartlehner, L. Lux, B.N. Gaynes, S. Lloyd, K.N. Lohr.
Administrative, technical, or logistic support: L.C. Morgan, L. Lux, M. Van Noord, P. Thieda, M. Strobelberger.
Collection and assembly of data: G. Gartlehner, R.A. Hansen, L.C. Morgan, K. Thaler, M. Van Noord, U. Mager, P. Thieda, B.N. Gaynes, M. Strobelberger, S. Lloyd.

www.annals.org

6 December 2011 | Annals of Internal Medicine | Volume 155 • Number 11 | W-245
Appendix Figure 1. Summary of evidence search and selection.

Records identified through database searches (n = 6375)

- Abstracts screened (n = 3927)
  - Duplicate entries removed (n = 2615)
  - Citations excluded (n = 2419)

Additional records identified through other sources (n = 167)

- Full-text articles assessed for eligibility (n = 1508)
  - Excluded (n = 1234)*
    - Foreign-language study: 7
    - Duration too short: 10
    - Wrong population: 85
    - Wrong drug: 143
    - Wrong outcome: 201
    - Wrong publication: 264
    - Did not address outcomes of interest: 281
    - Wrong design: 474
    - Poor quality: 79

Studies included in qualitative synthesis (n = 234; 274 articles)

Studies included in quantitative synthesis (mixed-treatment comparisons and meta-analyses) (n = 93)

* The number of included articles differs from the number of included studies because some studies have multiple publications.
Appendix Figure 2. Odds ratios of response rates comparing SNRIs with other second-generation antidepressants and comparing second-generation antidepressants with one another.

Comparison

| SNRIs vs. other second-generation antidepressants | Odds Ratio (95% CI) | Study Type |
|--------------------------------------------------|---------------------|------------|
| Desvenlafaxine vs. bupropion                      | 0.89 (0.58–1.30)    | MTC of 0/64 trials* |
| Desvenlafaxine vs. mirtazapine                    | 1.19 (0.67–1.97)    | MTC of 0/64 trials* |
| Desvenlafaxine vs. nefazodone                     | 1.05 (0.57–1.76)    | MTC of 0/64 trials* |
| Desvenlafaxine vs. trazodone                      | 0.81 (0.45–1.35)    | MTC of 0/64 trials* |
| Duloxetine vs. bupropion                          | 0.99 (0.69–1.37)    | MTC of 0/64 trials* |
| Duloxetine vs. mirtazapine                        | 1.32 (0.79–2.09)    | MTC of 0/64 trials* |
| Duloxetine vs. nefazodone                         | 1.16 (0.67–1.90)    | MTC of 0/64 trials* |
| Duloxetine vs. trazodone                          | 0.90 (0.52–1.44)    | MTC of 0/64 trials* |
| Venlafaxine vs. bupropion                         | 0.78 (0.54–1.10)    | MTC of 0/64 trials* |
| Venlafaxine vs. mirtazapine                       | 1.04 (0.61–1.67)    | MTC of 0/64 trials* |
| Venlafaxine vs. nefazodone                        | 0.92 (0.52–1.52)    | MTC of 0/64 trials* |
| Venlafaxine vs. trazodone                         | 0.71 (0.41–1.15)    | MTC of 0/64 trials* |

| Other vs. other second-generation antidepressants | Odds Ratio (95% CI) | Study Type |
|--------------------------------------------------|---------------------|------------|
| Bupropion vs. mirtazapine                         | 1.36 (0.79–2.17)    | MTC of 0/64 trials* |
| Bupropion vs. nefazodone                          | 1.20 (0.67–2.01)    | MTC of 0/64 trials* |
| Bupropion vs. trazodone                           | 0.92 (0.54–1.46)    | MTC of 0/64 trials* |
| Mirtazapine vs. nefazodone                        | 0.83 (0.45–1.69)    | MTC of 0/64 trials* |
| Mirtazapine vs. trazodone                         | 0.66 (0.43–1.07)    | MTC of 0/64 trials* |
| Nefazodone vs. trazodone                          | 0.82 (0.40–1.51)    | MTC of 0/64 trials* |

MTC = mixed-treatment comparison; SNRI = serotonin and norepinephrine reuptake inhibitor.

* The first number indicates the number of trials directly comparing 2 drugs; the second indicates the number of additional studies used to perform MTCs.

Appendix Table 1. Main Differences in Specific Adverse Events

| Drug                | Comparators                                    | Differences in Adverse Events                                                                 |
|---------------------|------------------------------------------------|---------------------------------------------------------------------------------------------|
| Bupropion           | Escitalopram, fluoxetine, paroxetine, sertraline| Lower incidence of sexual dysfunction than comparator drugs (6% vs. 16%)                     |
| Mirtazapine         | Fluoxetine, paroxetine, trazodone, venlafaxine | Greater weight gain than comparator drugs (mean, 0.8–3.0 kg after 6–8 wk)                   |
| Paroxetine          | Escitalopram, duloxetine, fluoxetine, mirtazapine, nefazodone, and sertraline | Higher incidence of sexual dysfunction, particularly ejaculatory dysfunction, than comparator drugs |
| Sertraline          | Bupropion, citalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, venlafaxine | Higher incidence of diarrhea than comparator drugs (mean, 16% [95% CI, 13%–20%] vs. 8% [CI, 4%–13%]) |
| Trazodone           | Bupropion, fluoxetine, mirtazapine, paroxetine, venlafaxine | Higher incidence of somnolence than comparator drugs (mean, 42% [CI, 19%–64%] vs. 25% [CI, 3%–46%]) |
| Venlafaxine         | SSRIs as a class                                | Higher incidence of nausea and vomiting than SSRIs as a class (mean, 33% [CI, 23%–43%] vs. 22% [CI, 16%–29%]) |

SSRI = selective serotonin reuptake inhibitor.
### Appendix Table 2. Comparative Efficacy or Effectiveness in Subgroups: Findings and Strength of Evidence

| Outcome                  | Strength of Evidence | Findings                                                                                   |
|--------------------------|----------------------|-------------------------------------------------------------------------------------------|
| **Age**                  |                      |                                            |                                                                 |
| Comparative efficacy     | Moderate             | Evidence from 11 trials indicates that efficacy does not differ substantially among second-generation antidepressants for treating MDD in patients aged ≥60 y. |
| Comparative harms        | Low                  | Results from 6 studies indicate that adverse events may differ somewhat across second-generation antidepressants in elderly patients. |
| **Sex**                  |                      |                                            |                                                                 |
| Comparative efficacy     | Insufficient         | No evidence                                                                               |
| Comparative effectiveness| Insufficient         | No evidence                                                                               |
| Comparative harms        | Low                  | Two trials suggest differences between men and women in sexual side effects.               |
| **Race or ethnicity**    |                      |                                            |                                                                 |
| Comparative efficacy     | Insufficient         | No evidence                                                                               |
| Comparative effectiveness| Insufficient         | No evidence                                                                               |
| Comparative harms        | Insufficient         | No evidence                                                                               |
| **Comorbid conditions**  |                      |                                            |                                                                 |
| Comparative efficacy     | Low                  | Results from a subgroup analysis of 1 trial indicate significantly greater response with extended-release venlafaxine than fluoxetine in patients with MDD and comorbid generalized anxiety disorder. |
| Comparative effectiveness| Insufficient         | No evidence                                                                               |
| Comparative harms        | Insufficient         | No evidence                                                                               |

MDD = major depressive disorder.

* High strength of evidence indicates high confidence that the evidence reflects the true effect. Further research is very unlikely to change our confidence in the estimate of effect. Moderate strength of evidence indicates that the evidence reflects the true effect. Further research may change our confidence in the estimate of effect and may change the estimate. Low strength of evidence indicates that the evidence reflects the true effect. Further research is likely to change both the confidence in the estimate of effect and the estimate. Insufficient strength of evidence indicates that evidence is either unavailable or does not permit a conclusion.