



# Effective Health Care Program

## Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression: An Update of the 2007 Comparative Effectiveness Review

### Executive Summary

#### Background

Depressive disorders such as major depressive disorder (MDD), dysthymia, and subsyndromal depression (including minor depression) may be serious disabling illnesses. MDD is the most prevalent, affecting more than 16 percent (lifetime) of U.S. adults. In 2000, the U.S. economic burden of depressive disorders was estimated to be \$83.1 billion. Likely, this number has increased during the past 10 years. More than 30 percent of these costs are attributable to direct medical expenses.

Pharmacotherapy dominates the medical management of depressive disorders and may include first-generation antidepressants (tricyclic antidepressants and monoamine oxidase inhibitors) and more recently developed second-generation antidepressants. These second-generation treatments include selective serotonin reuptake inhibitors (SSRIs: citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, sertraline), selective serotonin and norepinephrine reuptake inhibitors (SSNRIs: duloxetine), serotonin and norepinephrine reuptake inhibitors (SNRIs: desvenlafaxine, mirtazapine, venlafaxine), and other second-generation antidepressants (bupropion, nefazodone, trazodone). The mechanism of action of most of these agents is poorly understood.

#### Effective Health Care Program

The Effective Health Care Program was initiated in 2005 to provide valid evidence about the comparative effectiveness of different medical interventions. The object is to help consumers, health care providers, and others in making informed choices among treatment alternatives. Through its Comparative Effectiveness Reviews, the program supports systematic appraisals of existing scientific evidence regarding treatments for high-priority health conditions. It also promotes and generates new scientific evidence by identifying gaps in existing scientific evidence and supporting new research. The program puts special emphasis on translating findings into a variety of useful formats for different stakeholders, including consumers.

The full report and this summary are available at [www.effectivehealthcare.ahrq.gov/reports/final.cfm](http://www.effectivehealthcare.ahrq.gov/reports/final.cfm).

These drugs work, at least in part, through their effects on neurotransmitters such as serotonin, norepinephrine, or dopamine in the central nervous system.



Agency for Healthcare Research and Quality

Advancing Excellence in Health Care • [www.ahrq.gov](http://www.ahrq.gov)

Effective  
Health Care

In general, the efficacy of first- and second-generation antidepressant medications is similar. However, first-generation antidepressants often produce multiple side effects that many patients find intolerable, and the risk for harm when taken in overdose or in combination with certain medications is high. Because of their relatively favorable side-effect profile, the second-generation antidepressants play a prominent role in the management of patients with MDD and are the focus of this review.

## Objectives

This report is an update by RTI–UNC (Research Triangle Institute International–University of North Carolina) Evidence-based Practice Center of the 2007 Comparative Effectiveness Review of second-generation antidepressants. It summarizes the available evidence on the comparative efficacy, effectiveness, and harms of 13 second-generation antidepressants—bupropion, citalopram, desvenlafaxine, duloxetine, escitalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, sertraline, trazodone, and venlafaxine—in treating patients with MDD, dysthymia, and subsyndromal depression. It also evaluates the comparative efficacy and effectiveness for maintaining remission and treating accompanying symptoms such as anxiety, insomnia, or neurovegetative symptoms.

Specifically, we address the following Key Questions (KQs) in this report:

- 1a. For adults with major depressive disorder (MDD), dysthymia, or subsyndromal depressive disorders, do commonly used medications for depression differ in efficacy or effectiveness in treating depressive symptoms?
- 1b. If a patient has responded to one agent in the past, is that agent better than current alternatives at treating depressive symptoms?
- 1c. Are there any differences in efficacy or effectiveness between immediate-release and extended-release formulations of second-generation antidepressants?
- 2a. For adults with a depressive syndrome that has responded to antidepressant treatment, do second-generation antidepressants differ in their efficacy or effectiveness for preventing relapse (i.e., continuation phase) or recurrence (i.e., maintenance phase) when a patient:
  - continues the drug they initially responded to, or
  - switches to a different antidepressant?
- 2b. For adults with a depressive syndrome that has not responded to acute antidepressant treatment or has relapsed (continuation phase) or recurred (maintenance phase), do alternative second-generation antidepressants differ in their efficacy or effectiveness?
3. In depressed patients with accompanying symptoms such as anxiety, insomnia, and neurovegetative symptoms, do medications or combinations of medications (including tricyclics in combination) differ in their efficacy or effectiveness for treating the depressive episode or for treating the accompanying symptoms?
- 4a. For adults with a depressive syndrome, do commonly used antidepressants differ in safety, adverse events, or adherence? Adverse effects of interest include but are not limited to nausea, diarrhea, headache, tremor, daytime sedation, decreased libido, failure to achieve orgasm, nervousness, insomnia, and more serious events including suicide.
- 4b. Are there any differences in safety, adverse events, or adherence between immediate-release and extended-release formulations of second-generation antidepressants?
5. How do the efficacy, effectiveness, or harms of treatment with antidepressants for a depressive syndrome differ for the following subpopulations?
  - elderly or very elderly patients
  - other demographic groups (defined by age, ethnic or racial groups, and sex)
  - patients with medical comorbidities (e.g., ischemic heart disease, cancer)
  - patients with psychiatric and behavioral comorbidities (e.g., substance abuse disorders)
  - patients taking other medications

## Methods

The topic of this report and preliminary KQs arose through a public process involving the public, the Scientific Resource Center (SRC), and various stakeholder groups ([www.effectivehealthcare.ahrq.gov/index.cfm/who-is-involved-in-the-effective-health-care-program/about-the-stakeholder-group/](http://www.effectivehealthcare.ahrq.gov/index.cfm/who-is-involved-in-the-effective-health-care-program/about-the-stakeholder-group/)).

To identify articles relevant to each KQ, we searched PubMed, Embase, the Cochrane Library, PsycInfo, and International Pharmaceutical Abstracts. We used either Medical Subject Headings (MeSH or MH) as search terms when available or keywords when appropriate. We combined terms for selected indications (major depressive disorder, dysthymia, minor depression, subsyndromal depressive disorder), drug interactions, and adverse events with a list of 13 specific second-generation antidepressants (bupropion, citalopram, desvenlafaxine, duloxetine, escitalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, sertraline, trazodone, and venlafaxine). We limited electronic searches to “human” and “English language.” We searched sources from 1980 to January 2011 to capture literature relevant to the scope of our topic. The SRC contacted pharmaceutical manufacturers and invited them to submit dossiers, including citations. We received dossiers from five pharmaceutical companies (AstraZeneca, Eli Lilly, GlaxoSmithKline, Warner Chilcott Pharmaceuticals, and Wyeth). The SRC also searched various sources for grey literature.

For this review, results from well-conducted, valid head-to-head trials provide the strongest evidence to compare drugs with respect to efficacy, effectiveness, and harms. Randomized controlled trials (RCTs) of at least 6 weeks’ duration and in adult study populations were eligible for inclusion. For quantitative analyses, we included all eligible studies without sample size limitations. In addition to head-to-head studies, we included placebo-controlled trials for mixed treatment comparisons or if no head-to-head trials were available for a particular KQ. If we concluded that we could not conduct any quantitative analyses, then we included studies only if they had sample sizes of 40 or larger.

For harms (i.e., evidence pertaining to safety, tolerability, and adverse events), we examined data from both experimental and observational studies. We included observational studies that had large sample sizes (1,000 patients or more), lasted at least 3 months, and reported an outcome of interest. Two people independently reviewed abstracts and full-text articles. If both reviewers agreed that the trial did not meet eligibility criteria, we excluded it. We obtained the full text of all remaining articles and used the same eligibility criteria to determine which, if any, to exclude at this stage.

To assess the quality (internal validity) of studies, we used predefined criteria based on those developed by the U.S. Preventive Services Task Force (ratings: good, fair, poor) and the National Health Service Centre for Reviews and Dissemination. Two people independently rated the quality of each included study.

We assessed statistically each of the 78 possible drug comparisons of second-generation antidepressants for the treatment of acute-phase MDD. We conducted meta-analyses of 6 direct comparisons; the remaining 72 analyses employed mixed treatment comparison meta-analyses to derive indirect comparisons.

We evaluated the strength of evidence based on methods guidance for the Evidence-based Practice Center program of the Agency for Healthcare Research and Quality. Strength of evidence is graded only for major comparisons and major outcomes for the topic at hand. The strength of evidence for each outcome or comparison that we graded incorporates scores on four domains: risk of bias, consistency, directness, and precision.

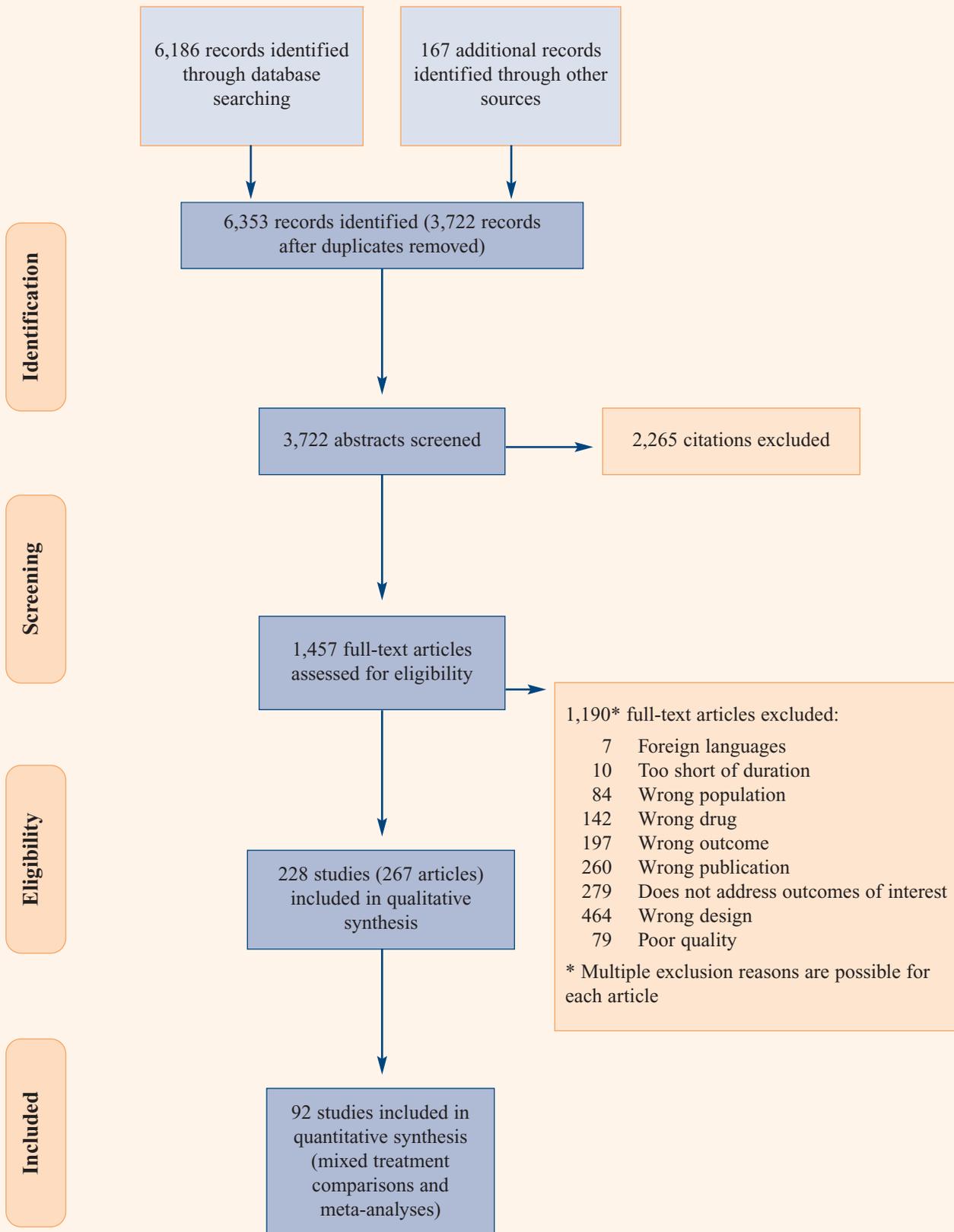
## Results

Overall, the new evidence (78 new studies, 87 articles) we found during the update of the 2007 report did not lead to changes in our main conclusion from that review—namely, that no substantial differences in efficacy exist among second-generation antidepressants for the treatment of MDD. Some results are now supported by better evidence than in 2007, which is reflected in a higher grade for the strength of the evidence for some outcomes. Our summary of evidence findings are presented in Tables A through I by KQ. The strength of evidence ratings for the main outcomes of each Key Question are detailed in Appendix G of the main report.

### Efficiency and Effectiveness

We identified 3,722 citations from searches and reviews of reference lists. Figure A documents the disposition of the 267 included articles in this review, working from 1,457 articles retrieved for full-text review and 1,190 excluded at this stage.

**Figure A. Results of literature search (PRISMA diagram)**



## Treatment of Major Depressive Disorder (KQ 1a)

Overall, 37 percent of patients did not respond during 6 to 12 weeks of treatment with second-generation antidepressants; 53 percent did not achieve remission. The evidence is insufficient to determine factors that can reliably predict response or nonresponse in individual patients.

Ninety-one head-to-head trials (i.e., comparisons between medications conducted within trials) provided data on 40 of the potential comparisons between the 13 second-generation antidepressants addressed in this report. Eight trials directly compared any non-SSRI second-generation antidepressant with any other; of these, only two comparisons were evaluated in more than one trial. Many efficacy trials were not powered to detect statistically or clinically significant differences, leading to inconclusive results.

Direct evidence from head-to-head trials was considered sufficient to conduct meta-analyses on the response to treatment (at least 50 percent improvement from baseline) for six drug–drug comparisons. Differences in efficacy reflected in some of these meta-analyses are of modest magnitude, and clinical implications remain to be determined.

- Citalopram versus escitalopram (5 published studies; 1,802 patients): For patients on escitalopram the odds ratio (OR) of response was statistically significantly higher than for patients on citalopram (OR, 1.47; 95% confidence interval [CI], 1.07 to 2.01). The number needed to treat (NNT) to gain 1 additional responder at week 8 with escitalopram compared with citalopram was 13 (95% CI, 8 to 39). These results are based on meta-analyses of head-to-head trials. Results of mixed-treatment comparisons, taking the entire evidence base on second-generation antidepressants into consideration, did not confirm these findings (OR, 0.51; 95% credible interval, 0.13 to 4.14).
- Fluoxetine versus paroxetine (5 studies; 690 patients): We did not find any statistically significant differences in response rates (OR, 1.08; 95% CI, 0.79 to 1.47) between fluoxetine and paroxetine.

- Fluoxetine versus sertraline (4 studies; 940 patients): The odds ratio of response was statistically significantly higher for sertraline than for fluoxetine (OR, 1.42; 95% CI, 1.08 to 1.85). The NNT to gain 1 additional responder at 6 to 12 weeks with sertraline was 13 (95% CI, 8 to 58).
- Fluoxetine versus venlafaxine (6 studies; 1,197 patients): The odds ratio of response was statistically significantly higher for patients on venlafaxine than on fluoxetine (OR, 1.47; 95% CI, 1.16 to 1.86).
- Paroxetine versus duloxetine (3 studies; 849 patients): Pooled response rates were similar for patients on paroxetine or duloxetine (OR, 0.84; 95% CI, 0.63 to 1.12).
- Sertraline versus venlafaxine (3 studies; 470 patients): Pooled response rates were similar for patients on sertraline or venlafaxine (OR, 1.18; 95% CI, 0.81 to 1.72).

Most trials were efficacy RCTs conducted in carefully selected populations under carefully controlled conditions. Only three trials met criteria for being an effectiveness trial, which is intended to have greater applicability to typical practice. Of these trials, two were conducted in French primary-care settings and one in primary-care clinics in the United States. Findings were generally consistent with efficacy trials and did not reflect any substantial differences in comparative effectiveness in adults.

Findings from indirect comparisons yielded some statistically significant differences in response rates. The magnitudes of these differences, however, were small and are likely not to be clinically significant. Overall, we graded the strength of the evidence supporting no substantial differences in efficacy and effectiveness among second-generation antidepressants for the treatment of MDD in adults as moderate.

### Quality of Life

Quality of life or functional capacity was infrequently assessed, usually as a secondary outcome. Seventeen studies (3,960 patients), mostly of fair quality, indicated no statistical differences in efficacy with respect to health-related quality of life. The strength of evidence is moderate.

## Speed of Response

Seven studies, all of fair quality and funded by the maker of mirtazapine, reported that mirtazapine had a significantly faster onset of action than citalopram, fluoxetine, paroxetine, and sertraline. The pooled NNT to yield one additional responder after 1 or 2 weeks of treatment is seven (95% CI, 5 to 12); after 4 weeks of treatment, however, most response rates were similar. The strength of evidence is moderate.

## Treatment of Dysthymia (KQ 1a)

### Efficacy and Effectiveness

We identified no head-to-head trial comparing different medications in a population with dysthymia. One good-quality and four fair-quality placebo-controlled trials provide mixed evidence on the general efficacy and effectiveness of fluoxetine, paroxetine, and sertraline for the treatment of dysthymia. A fair-quality effectiveness study provides mixed evidence on the effectiveness of paroxetine compared with placebo. A subgroup of patients older than 60 years old showed a significantly greater improvement than those on placebo; a subgroup of patients younger than 60 years old did not show any difference in effectiveness between paroxetine and placebo. The strength of evidence is insufficient.

## Treatment of Subsyndromal Depression (KQ 1a)

### Efficacy and Effectiveness

The only head-to-head evidence for treating patients with subsyndromal depression came from a nonrandomized, open-label trial comparing citalopram with sertraline. This study did not detect any differences in efficacy. Findings from two placebo-controlled trials (both fair quality) were insufficient to draw any conclusions about the comparative efficacy and effectiveness of second-generation antidepressants for the treatment of subsyndromal depression. The strength of evidence is low.

## Response to Antidepressant Agents After Successful Response in the Past (KQ 1b)

We did not find any evidence to answer this KQ.

## Difference in Efficacy Between Immediate- and Extended-Release Formulations (KQ 1c)

Four RCTs and one pooled analysis of two identical RCTs provide mixed results about differences in efficacy between immediate- and extended-release formulations of various drugs.

Two RCTs reported similar rates of maintenance of response and relapse for patients treated with fluoxetine daily or fluoxetine weekly during the continuation phase. Similarly, one RCT and a pooled analysis of two identical RCTs did not find any differences in response rates in patients treated with paroxetine IR (immediate release) or paroxetine CR (controlled release) for acute-phase MDD. The strength of evidence is moderate.

By contrast, one RCT reported higher response rates for patients on venlafaxine IR than venlafaxine XR (extended release).

We could not find any studies on other medications, such as bupropion or fluvoxamine, that are available as both immediate- and extended-release formulations.

## Maintenance of Response or Remission (KQ 2a)

### Efficacy and Effectiveness

Six head-to-head RCTs suggest that no substantial differences exist between escitalopram and desvenlafaxine, escitalopram and paroxetine, fluoxetine and sertraline, fluoxetine and venlafaxine, fluvoxamine and sertraline, or trazodone and venlafaxine for maintaining response or remission (i.e., preventing relapse or recurrence of MDD). One naturalistic study provides fair-quality evidence that rehospitalization rates do not differ between groups of patients continuing fluoxetine versus venlafaxine. The strength of the evidence is moderate. Thirty-one placebo-controlled trials support the general efficacy and effectiveness of most second-generation antidepressants for preventing relapse or recurrence. The overall

strength of this evidence is moderate.

No evidence addressed how second-generation antidepressants compare when a patient responds to one agent and then is required to switch to a different agent (e.g., because of changes in insurance benefit).

## **Achieving Response in Unresponsive or Recurrent Disease (KQ 2b)**

### **Efficacy and Effectiveness**

Four head-to-head studies and two effectiveness studies provide conflicting evidence on differences among second-generation antidepressants in treatment-resistant depression. A good-quality effectiveness study suggests that no substantial differences exist among bupropion SR (sustained release), sertraline, and venlafaxine XR, but a fair-quality effectiveness study suggests that venlafaxine is modestly more effective than citalopram, fluoxetine, mirtazapine, paroxetine, and sertraline. Three of four efficacy studies (all fair quality) suggest that venlafaxine trended toward being more effective than citalopram, fluoxetine, and paroxetine, although only the comparison with paroxetine was statistically significant. Given the conflicting results, the overall strength of the evidence is low.

Although several comparative studies included patients who had relapsed or who were experiencing a recurrent depressive episode, no study specifically compared one second-generation antidepressant with another as a second-step treatment in such patients.

## **Treatment of Depression in Patients With Accompanying Symptom Clusters (KQ 3)**

### **Anxiety**

Evidence from seven head-to-head trials (all fair quality) suggests that antidepressant medications do not differ substantially in antidepressive efficacy for patients with MDD and anxiety symptoms. The trials found no substantial differences in efficacy among fluoxetine, paroxetine, and sertraline or between citalopram and sertraline, bupropion and sertraline, or venlafaxine and sertraline. One trial found statistically significant superiority of venlafaxine over fluoxetine. Two trials provided inconsistent evidence regarding the

superiority of escitalopram over paroxetine. The strength of evidence is moderate.

### **Insomnia**

One head-to-head study provided evidence regarding comparative efficacy of medications for treatment of depression in patients with accompanying insomnia. The study showed no statistically significant differences in depressive outcomes for fluoxetine compared with paroxetine and sertraline. One trial of fluoxetine supplemented with eszopiclone compared with fluoxetine alone showed no statistically significant difference between the groups for depression scores when the sleep items were excluded from the analysis. The strength of evidence is low.

### **Low Energy**

One placebo-controlled RCT showed that bupropion XR is superior to placebo for treating depression in patients with low energy. The strength of evidence is insufficient.

### **Melancholia**

Two head-to-head trials provide limited evidence on the comparative effects of medication for treating depression in patients with melancholia. In one, depression response rates for sertraline were superior to those for fluoxetine; in another, depression scores improved similarly for venlafaxine and fluoxetine. The strength of evidence is insufficient.

### **Pain**

Two fair-quality trials that required baseline pain for inclusion produced conflicting evidence regarding the superiority of duloxetine compared with placebo for treating depression in patients with pain of at least mild intensity. The strength of evidence is insufficient.

### **Psychomotor Changes**

One fair-quality head-to-head trial reported no statistically significant difference between fluoxetine and sertraline for treating depression in patients with psychomotor retardation. The same study found that sertraline was more efficacious than fluoxetine for treating depression in patients with psychomotor agitation. The strength of evidence is insufficient.

## **Somatization**

We identified no relevant studies.

## **Treatment of Symptom Clusters in Patients With Accompanying Depression (KQ 3)**

### **Anxiety**

Twelve head-to-head trials and two placebo-controlled trials (all fair quality) provide evidence that antidepressant medications do not differ substantially in efficacy for treatment of anxiety associated with MDD. Trials found no substantial differences in efficacy for the following: fluoxetine, paroxetine, and sertraline; sertraline and bupropion; sertraline and venlafaxine; citalopram and mirtazapine; escitalopram and fluoxetine; and paroxetine and nefazodone. One trial found that venlafaxine was statistically significantly superior to fluoxetine and one trial found that escitalopram was superior to paroxetine. The strength of evidence is moderate.

### **Insomnia**

Six head-to-head trials (all fair quality) and one placebo-controlled trial provide limited evidence about comparative effects of antidepressants on insomnia in patients with depression. Three trials indicated similar efficacy for improving sleep for the following comparisons: fluoxetine, paroxetine, and sertraline; escitalopram and fluoxetine; and fluoxetine and mirtazapine. One trial suggested that trazodone was superior to fluoxetine and one trial suggested that trazodone is superior to venlafaxine in improving sleep scores in depressed patients. One trial showed that supplementing fluoxetine therapy with eszopiclone leads to improved sleep. The strength of evidence is low.

### **Low Energy**

One placebo-controlled RCT showed that bupropion XR is superior to placebo for treating low energy in depressed patients. The strength of evidence is insufficient.

### **Melancholia**

We identified no relevant study.

## **Pain**

One fair-quality systematic review showed that improvement in pain scores was similar for duloxetine and paroxetine. Six studies provided mixed evidence for the superiority of duloxetine or paroxetine compared with placebo for treatment of accompanying pain. The strength of evidence is moderate.

### **Psychomotor Changes**

We identified no relevant study.

### **Somatization**

One head-to-head trial of escitalopram and fluoxetine and one open-label effectiveness trial of fluoxetine, paroxetine, and setraline found no statistically significant difference for treating somatization in patients with depression. The strength of evidence is insufficient.

## **Differences in Harms (Adverse Events) (KQ 4a)**

We analyzed adverse-events data from 92 head-to-head efficacy studies on 22,586 patients, along with data from 48 additional studies of both experimental and observational design. Only five RCTs were designed primarily to detect differences in adverse events. Methods of adverse-events assessment in efficacy trials differed greatly. Few studies used objective scales. Determining whether assessment methods were unbiased and adequate was often difficult.

### **General Tolerability**

Constipation, diarrhea, dizziness, headache, insomnia, nausea, sexual adverse events, and somnolence were commonly and consistently reported adverse events. On average, 63 percent of patients in efficacy trials experienced at least one adverse event. Nausea and vomiting were found to be the most common reasons for discontinuation in efficacy studies. Overall, second-generation antidepressants have similar adverse-events profiles, and the strength of evidence is high.

However, some differences in the incidence of specific adverse events exist as follows:

- Venlafaxine was associated with an approximately 52 percent (95% CI, 25 to 84 percent) higher incidence of nausea and vomiting than SSRIs as a class. The strength of evidence is high.
- Mirtazapine led to higher weight gains than comparator drugs. Mean weight gains relative to pretreatment weights ranged from 0.8 kg to 3.0 kg after 6–8 weeks of treatment. The strength of evidence for higher risks of weight gain with mirtazapine than with other antidepressants is high.
- Sertraline led to higher rates of diarrhea than comparator drugs (bupropion, citalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, venlafaxine) in most studies. The incidence was 8 percent (95% CI, 3 to 11 percent) higher than with comparator drugs. Whether this finding can be extrapolated to comparisons of sertraline with other second-generation antidepressants remains unclear. The strength of evidence that sertraline has a higher risk of diarrhea than other antidepressants is moderate.
- Trazodone was associated with an approximately 16 percent (3 percent less to 36 percent higher) higher incidence of somnolence than comparator drugs (bupropion, fluoxetine, mirtazapine, paroxetine, venlafaxine). Whether this finding can be extrapolated to comparisons of trazodone with other second-generation antidepressants remains unclear. The strength of evidence that trazodone leads to higher rates of somnolence than comparator drugs is moderate.

### Discontinuation Rates

Overall discontinuation rates were similar between SSRIs as a class and other second-generation antidepressants. The strength of evidence is high.

Discontinuation rates because of adverse events were also similar between SSRIs as a class and bupropion, mirtazapine, nefazodone, and trazodone. The strength of evidence is high. Duloxetine had a 67 percent (95% CI, 17 to 139) higher and venlafaxine an approximately 40 percent (95% CI, 16 to 73) higher risk for discontinuation because of adverse events than SSRIs as a class. The strength of evidence is high.

Discontinuation rates because of lack of efficacy were similar between SSRIs as a class and bupropion, duloxetine, mirtazapine, nefazodone, and trazodone. Venlafaxine had a 34 percent (95% CI, 47 to 93) lower

risk of discontinuation because of lack of efficacy than SSRIs as a class. The strength of evidence is high.

### Severe Adverse Events

The strength of the evidence on the comparative risks of second-generation antidepressants on most serious adverse events is insufficient to draw firm conclusions. In general, trials and observational studies were too small and study durations too short to assess the comparative risks of rare but serious adverse events such as suicidality, seizures, cardiovascular adverse events, serotonin syndrome, hyponatremia, or hepatotoxicity. Long-term observational evidence is often lacking or prone to bias.

### Sexual Dysfunction

Six trials and a pooled analysis of two identical RCTs provide evidence that bupropion causes lower rates of sexual dysfunction than escitalopram, fluoxetine, paroxetine, and sertraline. The NNT to yield one additional person with a high overall satisfaction of sexual functioning is seven. This treatment effect was consistent across all studies. The strength of evidence that bupropion has lower rates of sexual dysfunction than comparator drugs is high.

Compared with other second-generation antidepressants (fluoxetine, fluvoxamine, nefazodone, and sertraline), paroxetine frequently led to higher rates of sexual dysfunction (16 percent vs. 6 percent). The strength of evidence is moderate.

### Other Severe Adverse Events

The existing evidence on the comparative risk for rare but severe adverse events such as suicidality, seizures, cardiovascular events, hyponatremia, hepatotoxicity, and serotonin syndrome is insufficient to draw firm conclusions. The strength of evidence is insufficient. Clinicians should keep in mind the risk of such harms during any course of treatment with a second-generation antidepressant.

### Adherence

Efficacy studies do not indicate any substantial differences in adherence among second-generation antidepressants. The strength of evidence is moderate.

To what extent findings from highly controlled efficacy trials can be extrapolated to “real-world” settings

remains uncertain. The evidence is insufficient to reach any conclusions about differences in adherence in effectiveness studies.

## **Comparative Harms and Adherence of Immediate- Versus Extended-Release Formulations (KQ 4b)**

Overall, adverse-event rates were similar between fluoxetine daily and fluoxetine weekly dosing regimens. Likewise, adverse-event rates were similar between paroxetine IR and paroxetine CR, as well as venlafaxine IR and venlafaxine XR, except for higher rates of nausea in patients treated with paroxetine IR than paroxetine CR.

We could not find any studies on bupropion and fluvoxamine immediate- and extended-release formulations.

The strength of evidence is moderate that no differences in adverse events exist between daily and weekly formulations of fluoxetine. The strength of evidence is low that paroxetine IR leads to higher rates of nausea than paroxetine CR.

Based on one double-blinded RCT, no differences in adherence between patients treated with paroxetine IR and paroxetine CR (93 percent vs. 96 percent) appear to exist. The strength of evidence is moderate.

A retrospective cohort study, based on U.S. prescription data, showed higher refill adherence for prescriptions of bupropion XL (extended release) than bupropion SR. The strength of evidence is low.

Based on an open-label RCT, adherence to fluoxetine weekly was higher than to fluoxetine daily. The strength of evidence is low.

## **Efficacy, Effectiveness, and Harms for Selected Populations (KQ 5)**

### **Age**

Eleven head-to-head trials in older adult patients with MDD indicate that efficacy does not differ substantially among second-generation antidepressants. The strength of the evidence is moderate. We found no head-to-head studies addressing differences in efficacy or harms in older patients with dysthymia or subsyndromal depression.

Head-to-head trials suggest some differences in adverse events among older adults. The strength of the evidence is low.

### **Sex**

We found no head-to-head trials comparing the efficacy of antidepressants in men and women; the strength of evidence is insufficient. Evidence from one RCT comparing paroxetine with sertraline and one RCT comparing paroxetine with bupropion SR suggests differences in sexual side effects between men and women. The strength of evidence is low.

### **Race or Ethnicity**

We found no head-to-head trials comparing the efficacy of second-generation antidepressants in different racial or ethnic groups. One fair-quality trial found no significant differences in efficacy or quality of life between sertraline and placebo in low-income Latino and black patients. The remaining evidence is limited to a handful of poor-quality studies assessing the general efficacy of duloxetine or fluoxetine. The strength of the evidence is insufficient.

### **Comorbidities**

The evidence for various comorbidities (e.g., alcohol and substance abuse, Alzheimer's disease or other dementia, arthritis, cancer, coronary artery disease, diabetes, or stroke) is limited to subgroup analyses of head-to-head studies in MDD patients with co-occurring generalized anxiety disorder, a number of placebo-controlled trials across various comorbidities, and one systematic review of SSRIs for depression and comorbid myocardial infarction. These trials provide inadequate comparative evidence on the efficacy of second-generation antidepressants in subgroups with different coexisting conditions. The strength of the evidence is insufficient.

## **Discussion**

Overall, the new evidence (78 new studies, 87 articles) we found during the update of our 2007 report did not lead to changes in our main conclusion from that review—namely, that no substantial differences in efficacy exist among second-generation antidepressants for the treatment of MDD. Some results are now supported by better evidence than in 2007, which is reflected in a higher grade for the strength of the

evidence for some outcomes. In addition, the more advanced statistical analysis that we were able to do for indirect comparisons of second-generation antidepressants when no or only insufficient head-to-head evidence was available also confirmed that conclusion.

Therefore, our findings indicate that the existing evidence does not warrant the choice of one second-generation antidepressant over another based on either greater efficacy or greater effectiveness. Some of the comparisons rendered statistically significant results; the magnitudes of the differences, however, are small and likely not clinically significant. Furthermore, because we had 78 pairwise comparisons, some are expected to be statistically significant by chance alone.

Although second-generation antidepressants are similar in efficacy, they cannot be considered identical drugs. Evidence of high and moderate strength supports some differences among individual drugs with respect to onset of action, adverse events, and some measures of health-related quality of life; these differences are of modest magnitude but statistically significant. Specifically, consistent evidence from multiple trials demonstrates that mirtazapine has a faster onset of action than citalopram, fluoxetine, paroxetine, and sertraline and that bupropion has fewer sexual side effects than escitalopram, fluoxetine, paroxetine, and sertraline.

Some of these differences are small and might be offset by adverse events. For example, a faster onset of mirtazapine must be weighed against possible decreased adherence because of long-term weight gain. Nonetheless, some of these differences may be clinically significant and influence the choice of a medication for specific patients.

The evidence is sparse (strength of evidence for comparative efficacy is insufficient for dysthymia and subsyndromal depression). No conclusions can be drawn about comparative efficacy or effectiveness.

A considerable limitation of our conclusions is that they have been derived primarily from efficacy trials. For example, for acute-phase MDD we found only 3 effectiveness studies out of 92 head-to-head RCTs. Two of these effectiveness studies were conducted in Europe, and the 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. Whether, for acute-phase MDD, such findings can be further extrapolated to other second-generation antidepressants remains unclear.

Given that almost two in five patients do not respond to initial treatment and that several other systematic reviews have concluded that no one antidepressant performs better than any other, an important future pharmacologic research agenda item is to focus on making the initial treatment strategy more effective. Potential approaches include looking at ways to better predict the treatment response to optimize initial treatment selections (e.g., through genetic analysis) and to explore whether combinations of antidepressants at treatment initiation would improve response rates. Furthermore, studies need to explore patient preferences about dosing regimens and the level of acceptance that individual patients have for various adverse events.

In addition, more evidence is needed regarding the most appropriate duration of antidepressant treatment for maintaining response and remission. Such studies should also evaluate further whether different formulations (i.e., controlled release vs. immediate release) lead to differences in adherence and subsequently to differences in relapse or recurrence. Additionally, although most trials maintained the dose used in acute-phase treatment throughout continuation and maintenance treatment, little is known about the effect of drug dose on the risk of relapse or recurrence.

More research is also needed to evaluate whether the benefits or harms of second-generation antidepressants differ in populations with accompanying symptoms such as anxiety, insomnia, pain, or fatigue. This research should identify and use a common core of accurate measures to identify these subgroups. Likewise, future research has to clarify differences of second-generation antidepressants in subgroups based on age, sex, race or ethnicity, and common comorbidities.

Finally, no evidence addressed how second-generation antidepressants compare when a patient responds to one agent and then is required to switch to a different agent (e.g., because of changes in insurance benefit). Because these circumstances may be relevant for many patients, future studies should consider this question.

**Table A. Summary of findings with strength of evidence, Key Question 1a: Comparative efficacy and effectiveness of second-generation antidepressants**

<b>Disorder, and Outcome of Interest</b>	<b>Strength of Evidence<sup>a</sup></b>	<b>Findings<sup>b</sup></b>
<b>Major depressive disorder</b>		
Comparative efficacy	Moderate	Results from direct and indirect comparisons based on 61 head-to-head trials and 31 placebo-controlled trials indicate that no substantial differences in efficacy exist among second-generation antidepressants.
Comparative effectiveness	Moderate	Direct evidence from three effectiveness trials (one good) and indirect evidence from efficacy trials indicate that no substantial differences in effectiveness exist among second-generation antidepressants.
Quality of life	Moderate	Consistent results from 18 trials indicate that the efficacy of second-generation antidepressants with respect to quality of life does not differ among drugs.
Onset of action	Moderate	Consistent results from seven trials suggest that mirtazapine has a significantly faster onset of action than citalopram, fluoxetine, paroxetine, and sertraline. Whether this difference can be extrapolated to other second-generation antidepressants is unclear. Most other trials do not indicate a faster onset of action of one second-generation antidepressant compared with another.
<b>Dysthymia</b>		
Comparative efficacy	Insufficient	No head-to-head evidence exists. Results from five placebo-controlled trials were insufficient to draw conclusions about comparative efficacy.
Comparative effectiveness	Insufficient	No head-to-head evidence exists. One effectiveness trial provides mixed evidence about paroxetine versus placebo; patients older than 60 showed greater improvement on paroxetine; those younger than 50 did not show any difference.
Quality of life	Insufficient	No evidence
Onset of action	Insufficient	No evidence
<b>Subsyndromal depression</b>		
Comparative efficacy	Low	One nonrandomized, open-label trial did not detect any difference between citalopram and sertraline. Results from two placebo-controlled trials were insufficient to draw conclusions.
Comparative effectiveness	Insufficient	No evidence
Quality of life	Insufficient	No evidence
Onset of action	Insufficient	No evidence

<sup>a</sup>Strength of evidence grades (high, moderate, low, or insufficient) are based on methods guidance for the EPC program; outcomes for which we have no studies are designated no evidence.

<sup>b</sup>Good, fair, or poor designations relate to quality grades given to each study; see Methods chapter of main report. We provide the designations only for good (or poor) studies; the remaining studies are all of fair quality.

**Table B. Summary of findings with strength of evidence, Key Question 1b: Greater efficacy and effectiveness with previously effective medications**

<b>Disorder, and Outcome of Interest</b>	<b>Strength of Evidence<sup>a</sup></b>	<b>Findings<sup>b</sup></b>
Major depressive disorder	Insufficient	No evidence
Dysthymia	Insufficient	No evidence
Subsyndromal depression	Insufficient	No evidence

<sup>a</sup>Strength of evidence grades (high, moderate, low, or insufficient) are based on methods guidance for the EPC program; outcomes for which we have no studies are designated no evidence.

<sup>b</sup>Good, fair, or poor designations relate to quality grades given to each study; see Methods chapter of main report. We provide the designations only for good (or poor) studies; the remaining studies are all of fair quality.

**Table C. Summary of findings with strength of evidence, Key Question 1c: Differences in efficacy and effectiveness between immediate- and extended-release formulations**

<b>Disorder, and Outcome of Interest</b>	<b>Strength of Evidence<sup>a</sup></b>	<b>Findings<sup>b</sup></b>
Major depressive disorder	Moderate	Results from two trials indicate that no differences in response to treatment exist between paroxetine IR and paroxetine CR. Two trials did not detect significant differences in maintenance of response and remission between fluoxetine daily and fluoxetine weekly.
	Low	One trial reported higher response rates for venlafaxine XR than venlafaxine IR.
Dysthymia	Insufficient	No evidence
Subsyndromal depression	Insufficient	No evidence

CR = controlled release; IR = immediate release; XR = extended release

<sup>a</sup>Strength of evidence grades (high, moderate, low, or insufficient) are based on methods guidance for the EPC program; outcomes for which we have no studies are designated no evidence.

<sup>b</sup>Good, fair, or poor designations relate to quality grades given to each study; see Methods chapter of main report. We provide the designations only for good (or poor) studies; the remaining studies are all of fair quality.

**Table D. Summary of findings with strength of evidence, Key Question 2a: Efficacy and effectiveness of second-generation antidepressants for maintaining response or remission (i.e., preventing relapse or recurrence)**

<b>Disorder, and Outcome of Interest</b>	<b>Strength of Evidence<sup>a</sup></b>	<b>Findings<sup>b</sup></b>
<b>Continuing initial medications</b>		
Comparative efficacy	Moderate	Based on results from six efficacy trials and one naturalistic study, no significant differences exist between escitalopram and desvenlafaxine, escitalopram and paroxetine, fluoxetine and sertraline, fluoxetine and venlafaxine, fluvoxamine and sertraline, and trazodone and venlafaxine for preventing relapse or recurrence.
Comparative effectiveness	Insufficient	No evidence
<b>Switching medications</b>		
Comparative efficacy	Insufficient	No evidence
Comparative effectiveness	Insufficient	No evidence

<sup>a</sup>Strength of evidence grades (high, moderate, low, or insufficient) are based on methods guidance for the EPC program; outcomes for which we have no studies are designated no evidence.

<sup>b</sup>Good, fair, or poor designations relate to quality grades given to each study; see Methods chapter of main report. We provide the designations only for good (or poor) studies; the remaining studies are all of fair quality.

**Table E. Summary of findings with strength of evidence, Key Question 2b: Efficacy and effectiveness of second-generation antidepressants in managing treatment-resistant depression syndrome or treating recurrent depression**

<b>Disorder, and Outcome of Interest</b>	<b>Strength of Evidence<sup>a</sup></b>	<b>Findings<sup>b</sup></b>
Comparative efficacy	Low	Results from four trials suggest no differences or only modest differences between SSRIs and venlafaxine. Numerical trends favored venlafaxine over comparator drugs in three of these trials, but differences were statistically significant in only one trial, which compared venlafaxine with paroxetine.
Comparative effectiveness	Low	Results from two effectiveness studies are conflicting. Based on one trial rated good, no significant differences in effectiveness exist among bupropion SR, sertraline, and venlafaxine XR. One effectiveness trial found venlafaxine to be modestly superior to citalopram, fluoxetine, mirtazapine, paroxetine, and sertraline.

SR = slow release; SSRI = selective serotonin reuptake inhibitor; XR = extended release

<sup>a</sup>Strength of evidence grades (high, moderate, low, or insufficient) are based on methods guidance for the EPC program; outcomes for which we have no studies are designated no evidence.

<sup>b</sup>Good, fair, or poor designations relate to quality grades given to each study; see Methods chapter of main report. We provide the designations only for good (or poor) studies; the remaining studies are all of fair quality.

**Table F. Summary of findings with strength of evidence, Key Question 3: Comparative efficacy and effectiveness of second-generation antidepressants for treatment of depression in patients with accompanying symptom clusters**

<b>Accompanying Symptom, and Outcome of Interest</b>	<b>Strength of Evidence<sup>a</sup></b>	<b>Findings<sup>b</sup></b>
<b>Anxiety</b>		
Comparative efficacy for depression	Moderate	Results from five head-to-head trials suggest that efficacy does not differ substantially for treatment of depression in patients with accompanying anxiety.
Comparative effectiveness for depression	Insufficient	No evidence
Comparative efficacy for anxiety	Moderate	Results from eight head-to-head trials and three placebo-controlled trials suggest that no substantial differences in efficacy exist among second-generation antidepressants for treatment of accompanying anxiety symptoms.
Comparative effectiveness for anxiety	Insufficient	No evidence
<b>Insomnia</b>		
Comparative efficacy for depression	Insufficient	Results from one head-to-head study are insufficient to draw conclusions about the comparative efficacy for treating depression in patients with coexisting insomnia.
Comparative effectiveness for depression	Insufficient	No evidence
Comparative efficacy for insomnia	Low	Results from five head-to-head trials suggest 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.
Comparative effectiveness for insomnia	Insufficient	No evidence
<b>Low energy</b>		
Comparative efficacy for depression	Insufficient	Results from one placebo-controlled trial of bupropion XL are insufficient to draw conclusions about treating depression in patients with coexisting low energy. Results from head-to-head trials are not available.
Comparative effectiveness for depression	Insufficient	No evidence
Comparative efficacy for low energy	Insufficient	Results from one placebo-controlled trial of bupropion XL are insufficient to draw conclusions about treating low energy in depressed patients. Results from head-to-head trials are not available.

**Table F. Summary of findings with strength of evidence, Key Question 3: Comparative efficacy and effectiveness of second-generation antidepressants for treatment of depression in patients with accompanying symptom clusters (continued)**

<b>Accompanying Symptom, and Outcome of Interest</b>	<b>Strength of Evidence<sup>a</sup></b>	<b>Findings<sup>b</sup></b>
<b>Low energy (continued)</b>		
Comparative effectiveness for low energy	Insufficient	No evidence
<b>Melancholia</b>		
Comparative efficacy for depression	Insufficient	Results from two head-to-head trials are insufficient to draw conclusions about treating depression in patients with coexisting melancholia. Results are inconsistent across studies.
Comparative effectiveness for depression	Insufficient	No evidence
Comparative efficacy for melancholia	Insufficient	No evidence
Comparative effectiveness for melancholia	Insufficient	No evidence
<b>Pain</b>		
Comparative efficacy for depression	Insufficient	Results from two placebo-controlled trials are conflicting regarding the superiority of duloxetine over placebo. Results from head-to-head trials are not available.
Comparative effectiveness for depression	Insufficient	No evidence
Comparative efficacy for pain	Moderate	Evidence from one systematic review, two head-to-head trials (one poor), and five placebo-controlled trials indicate no difference in efficacy between paroxetine and duloxetine.
Comparative effectiveness for pain	Insufficient	No evidence
<b>Psychomotor change</b>		
Comparative efficacy for depression	Insufficient	Results from one head-to-head trial are insufficient to draw conclusions about the comparative efficacy for treating depression in patients with coexisting psychomotor change.
Comparative effectiveness for depression	Insufficient	No evidence
Comparative efficacy for psychomotor change	Insufficient	No evidence
Comparative effectiveness for psychomotor change	Insufficient	No evidence

**Table F. Summary of findings with strength of evidence, Key Question 3: Comparative efficacy and effectiveness of second-generation antidepressants for treatment of depression in patients with accompanying symptom clusters (continued)**

<b>Accompanying Symptom, and Outcome of Interest</b>	<b>Strength of Evidence<sup>a</sup></b>	<b>Findings<sup>b</sup></b>
<b>Somatization</b>		
Comparative efficacy for depression	Insufficient	No evidence
Comparative effectiveness for depression	Insufficient	No evidence
Comparative efficacy for somatization	Insufficient	Results from one head-to-head trial are insufficient to draw conclusions about the comparative efficacy for treating somatization in depressed patients. Results indicate similar improvement in somatization.
Comparative effectiveness for somatization	Insufficient	Evidence from one open-label head-to-head trial is insufficient to draw conclusions about the comparative efficacy for treating coexisting somatization in depressed patients. Results indicate no difference in effectiveness.

XL = extended release

aStrength of evidence grades (high, moderate, low, or insufficient) are based on methods guidance for the EPC program; outcomes for which we have no studies are designated no evidence.

b Good, fair, or poor designations relate to quality grades given to each study; see Methods chapter of main report. We provide the designations only for good (or poor) studies; the remaining studies are all of fair quality.

**Table G. Summary of findings with strength of evidence, Key Question 4a: Comparative risk of harms (safety, adverse events), adherence, and persistence**

<b>Outcome of Interest</b>	<b>Strength of Evidence<sup>a</sup></b>	<b>Findings<sup>b</sup></b>
<b>General tolerability</b>		
Adverse-events profiles	High	Adverse-events profiles, based on 92 efficacy trials and 48 studies of experimental or observational design, are similar among second-generation antidepressants. The incidence of specific adverse events differs across antidepressants
Comparative risk of nausea and vomiting	High	Meta-analysis of 15 studies indicates that venlafaxine has a higher rate of nausea and vomiting than SSRIs as a class.
Comparative risk of weight change	High	Results from seven trials indicate that mirtazapine leads to higher weight gains than citalopram, fluoxetine, paroxetine, and sertraline.
Comparative risk of gastrointestinal adverse events	Moderate	Results from 15 studies indicate that sertraline has a higher incidence of diarrhea than bupropion, citalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, and venlafaxine. Results from one systematic review confirm some of these findings.

**Table G. Summary of findings with strength of evidence, Key Question 4a: Comparative risk of harms (safety, adverse events), adherence, and persistence (continued)**

<b>Outcome of Interest</b>	<b>Strength of Evidence<sup>a</sup></b>	<b>Findings<sup>b</sup></b>
<b>General tolerability (continued)</b>		
Comparative risk of somnolence	Moderate	Results from six trials indicate that trazodone has a higher rate of somnolence than bupropion, fluoxetine, mirtazapine, paroxetine, and venlafaxine.
Comparative risk of discontinuation syndrome	Moderate	A good systematic review indicates that paroxetine and venlafaxine have the highest rates of discontinuation syndrome; fluoxetine has the lowest.
Comparative risk of discontinuation of treatment	High	Meta-analyses of numerous efficacy trials indicate that overall discontinuation rates are similar. Duloxetine and venlafaxine have a higher rate of discontinuations because of adverse events than SSRIs as a class. Venlafaxine has a lower rate of discontinuations because of lack of efficacy than SSRIs as a class.
<b>Severe adverse events</b>		
Comparative risk of suicidality (suicidal thoughts and behavior)	Insufficient	Results from 11 observational studies (two good quality), five meta-analyses or systematic reviews (four good), and one systematic review yield conflicting information about the comparative risk of suicidality.
Comparative risk of sexual dysfunction	High	Results from six trials indicate that bupropion causes significantly less sexual dysfunction than escitalopram, fluoxetine, paroxetine, and sertraline.
	Moderate	Among SSRIs, paroxetine has the highest rates of sexual dysfunction.
Comparative risk of seizures	Insufficient	Results from three studies (one good observational design) yield conflicting information about the comparative risk of seizures.
Cardiovascular events	Insufficient	Results from one good observational study and one pooled analysis yield noncomparative or conflicting information about the comparative risk of cardiovascular events.
Comparative risk of hyponatremia	Insufficient	No trials or observational studies assessing hyponatremia met criteria for inclusion in this review. One cohort study not meeting inclusion criteria suggested that hyponatremia was more common in elderly patients treated with various antidepressants than in placebo-treated patients.
Comparative risk of hepatotoxicity	Insufficient	Evidence from existing studies is insufficient to draw conclusions about the comparative risk of hepatotoxicity. Weak evidence indicates that nefazodone might have an increased risk of hepatotoxicity.
Comparative risk of serotonin syndrome	Insufficient	No trials or observational studies assessing serotonin syndrome were included in this review. Numerous case reports of this syndrome exist but were not included in this review.

**Table G. Summary of findings with strength of evidence, Key Question 4a: Comparative risk of harms (safety, adverse events), adherence, and persistence (continued)**

Outcome of Interest	Strength of Evidence <sup>a</sup>	Findings <sup>b</sup>
<b>Adherence</b>		
Comparative adherence in efficacy studies	Moderate	Efficacy studies indicate no differences in adherence.
Comparative adherence in effectiveness studies	Insufficient	Evidence from existing studies is insufficient to draw conclusions about adherence in real-world settings.
Comparative persistence	Insufficient	No evidence

SSRI = selective serotonin reuptake inhibitor

<sup>a</sup>Strength of evidence grades (high, moderate, low, or insufficient) are based on methods guidance for the EPC program; outcomes for which we have no studies are designated no evidence.

<sup>b</sup>Good, fair, or poor designations relate to quality grades given to each study; see Methods chapter of main report. We provide the designations only for good (or poor) studies; the remaining studies are all of fair quality.

**Table H. Summary of findings with strength of evidence, Key Question 4b: Differences in harms, adherence, and persistence between immediate- and extended-release formulations**

Outcome of Interest	Strength of Evidence <sup>a</sup>	Findings <sup>b</sup>
<b>Major depressive disorder</b>		
Comparative risk of harms	Moderate	Findings from one trial each indicate that no differences in harms exist between fluoxetine daily and fluoxetine weekly or between venlafaxine IR and venlafaxine XR.
	Low	One trial provides evidence that paroxetine IR leads to higher rates of nausea than paroxetine CR.
Comparative adherence	Low	One trial provides evidence that fluoxetine weekly has better adherence rates than fluoxetine daily.
Comparative persistence	Low	Evidence from one observational study indicates that prescription refills are more common with the extended-release than the immediate-release formulation of bupropion.
Dysthymia	Insufficient	No evidence
Subsyndromal depression	Insufficient	No evidence

CR = controlled release; IR = immediate release; XR = extended release

<sup>a</sup>Strength of evidence grades (high, moderate, low, or insufficient) are based on methods guidance for the EPC program; outcomes for which we have no studies are designated no evidence.

<sup>b</sup> Good, fair, or poor designations relate to quality grades given to each study; see Methods chapter of main report. We provide the designations only for good (or poor) studies; the remaining studies are all of fair quality.

**Table I. Summary of findings with strength of evidence, Key Question 5: Subgroups**

<b>Subpopulation of Interest, and Outcome of Interest</b>	<b>Strength of Evidence<sup>a</sup></b>	<b>Findings<sup>b</sup></b>
<b>Age</b>		
Comparative efficacy	Moderate	Evidence from 11 trials indicates that efficacy does not differ substantially among second-generation antidepressants for treating MDD in patients age 60 years or older.
	Insufficient	No head-to-head evidence found for dysthymia or subsyndromal depression. Results from one good placebo-controlled trial showed no difference between fluoxetine and placebo.
Comparative effectiveness	Insufficient	No evidence in older patients with MDD.
	Insufficient	One effectiveness study showed greater improvement with paroxetine versus placebo in dysthymia patients older than 60 years; insufficient evidence to draw conclusions on comparative effectiveness.
Comparative harms	Low	Results from six studies indicate that adverse events may differ somewhat across second-generation antidepressants in older adults.
	Insufficient	No head-to-head studies were found for dysthymia or subsyndromal depression.
<b>Sex</b>		
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.
<b>Race or ethnicity</b>		
Comparative efficacy	Insufficient	No evidence
Comparative effectiveness	Insufficient	No evidence
Comparative harms	Insufficient	No evidence
<b>Comorbidities</b>		
Comparative efficacy	Low	Results from a subgroup analysis of one trial indicate significantly greater response with venlafaxine XR than fluoxetine in patients with MDD and comorbid generalized anxiety disorder.
	Insufficient	Placebo-controlled trials assessed efficacy in patients with the following comorbidities: alcohol/substance abuse, Alzheimer's disease/dementia, arthritis, diabetes, HIV/AIDS, multiple sclerosis, stroke, and vascular disease. No head-to-head evidence exists on comparative efficacy.
Comparative effectiveness	Insufficient	No evidence
Comparative harms	Insufficient	No evidence

MDD = major depressive disorder; XR = extended release

<sup>a</sup>Strength of evidence grades (high, moderate, low, or insufficient) are based on methods guidance for the EPC program; outcomes for which we have no studies are designated no evidence.

<sup>b</sup>Good or fair designations relate to quality grades given to each study; see Methods chapter of main report. We provide the designations only for good (or poor) studies; the remaining studies are all of fair

## Glossary

CI	Confidence interval
CR	Controlled release
KQ	Key Question
IR	Immediate release
MDD	Major depressive disorder
NNT	Number needed to treat
RCT	Randomized controlled trial
SRC	Scientific Resource Center
SR	Sustained release
SSNRI	Selective serotonin norepinephrine reuptake inhibitor
SSRI	Selective serotonin reuptake inhibitor
OR	Odds ratio
XR	Extended release

## Full Report

This executive summary is part of the following document: Gartlehner G, Hansen RA, Morgan LC, Thaler K, Lux LJ, Van Noord M, Mager U, Gaynes BN, Thieda P, Strobelberger M, Lloyd S, Reichenpfader U, Lohr KN. Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression: An Update of the 2007 Comparative Effectiveness Review. Comparative Effectiveness Review No. 46. (Prepared by the RTI International–University of North Carolina Evidence-based Practice Center, Contract No. 290-2007-10056-I.) AHRQ Publication No. 12-EHC012-EF. Rockville, MD: Agency for Healthcare Research and Quality. December 2011.  
[www.effectivehealthcare.ahrq.gov/reports/final.cfm](http://www.effectivehealthcare.ahrq.gov/reports/final.cfm).

## For More Copies

For more copies of *Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression: An Update of the 2007 Comparative Effectiveness Review: Comparative Effectiveness Review Executive Summary No. 46* (AHRQ Pub. No. 12-EHC012-1), please call the AHRQ Clearinghouse at 1–800–358–9295 or email [ahrqpubs@ahrq.gov](mailto:ahrqpubs@ahrq.gov).





