Number 7

Effective Health Care

Comparative Effectiveness of Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression

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. 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) and serotonin and norepinephrine reuptake inhibitors (SNRIs). The mechanism of action of most of these agents is poorly understood. These drugs work, at least in part, through their effects on neurotransmitters such as serotonin, norepinephrine, or dopamine in the central nervous system.

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

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**

AHRQ
Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

Effective Health Care

profile, the second-generation antidepressants play a prominent role in the management of patients with major depressive disorder and are the focus of this review.

This report summarizes the available evidence on the comparative efficacy, effectiveness, and harms of 12 second-generation antidepressants—bupropion, citalopram, 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 for treating accompanying symptoms such as anxiety, insomnia, or neurovegetative symptoms.

We rate the strength of evidence according to a modified GRADE approach. GRADE incorporates four key elements—study design, study quality, consistency, and directness—to characterize the strength of the body of evidence to answer key questions. We used three grades: high, moderate, and low (combining the GRADE category of very low with low). The quality of individual studies is denoted as good, fair, or poor. We assessed statistically each of the 66 possible drug comparisons of second-generation antidepressants. When data were sufficient, we did four direct comparisons; the remaining 62 analyses employed indirect comparison approaches.

Specifically, we address the following key questions (KQs) in this report:

1a. For adults with 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?

2a. For adults with a depressive syndrome, do antidepressants differ in their efficacy or effectiveness for maintaining response or remission (i.e., preventing relapse or recurrence)?

2b. For adults receiving antidepressant treatment for a depressive syndrome that either has not responded (acute phase) or has relapsed (continuation phase) or recurred (maintenance phase), do alternative antidepressants differ in their efficacy or effectiveness?

3. Do medications or combinations of medications (including tricyclics in combination) used to treat depression differ in their efficacy or effectiveness for treating accompanying symptoms, such as anxiety, insomnia, and neurovegetative symptoms?

   3a: Do medications differ in their efficacy and effectiveness in treating the depressive episode?

   3b: Do medications differ in their efficacy and effectiveness in treating the accompanying symptoms?

4. 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 severe events including suicide.

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)?

Table A summarizes the findings on second-generation antidepressants in the treatment of adult depression.

Conclusions

Treatment of Major Depressive Disorder (KQ 1)

Efficacy and effectiveness. From a total of 2,099 citations identified, we ultimately included 293 articles in this review, which represented 187 studies of good or fair quality. Of these, 89 were head-to-head randomized controlled trials (RCTs) and 57 were placebo-controlled RCTs; the remainder were observational or other types of studies or other qualitative or quantitative systematic reviews.

Of these 187 studies, 126 were financially supported by pharmaceutical companies and 17 by government agencies or independent funds; for 44 studies, we could not determine the funding source.

Overall, 38 percent of patients did not respond during 6 to 12 weeks of treatment with second-generation antidepressants; 54 percent did not achieve remission. The evidence is insufficient to determine factors that can reliably predict response or nonresponse in individual patients.

Seventy-two head-to-head comparisons (i.e., comparisons between medications conducted within trials) provided data on 35 of the potential comparisons between the 12 second-generation antidepressants addressed in this report. Five trials directly compared any non-SSRI second-generation antidepressant with any other non-SSRI second-generation antidepressant; of these, only one comparison was 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 for four 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 vs. escitalopram* (five studies; 1,545 patients): Patients on escitalopram had an additional treatment effect of a 1.25-point reduction (95-percent confidence interval [CI], 0.10-2.39) on the Montgomery-Asberg Depression Rating Scale (MADRS) compared with patients on citalopram. The relative risk (RR) of response was statistically significantly greater for escitalopram than for citalopram (RR: 1.14; 95-percent CI, 1.04-1.26). The number needed to treat (NNT) to gain one additional responder at week 8 with escitalopram was 14 (95-percent CI, 7-111). Both drugs are produced by the same manufacturer, which funded all available studies.
- *Fluoxetine vs. paroxetine* (seven studies; 950 patients): We did not find any statistically significant differences in effect sizes on the Hamilton Rating Scale for Depression (HAM-D) or response rates between fluoxetine and paroxetine. Fluoxetine had an additional reduction of 0.55 (95-percent CI, –1.4-0.36; P = 0.23) points on HAM-D compared with paroxetine; paroxetine led to a higher rate of responders than fluoxetine (RR: 1.09; 95-percent CI, 0.99-1.21).
- *Fluoxetine vs. sertraline* (four studies; 940 patients): Patients on sertraline had an additional, statistically nonsignificant treatment effect of a 0.75-point reduction (95-percent CI, –0.45-1.95) on the HAM-D scale compared with patients on fluoxetine. The relative risk of response was statistically significantly greater for sertraline than for fluoxetine (RR: 1.11; 95-percent CI, 1.01-1.21). The NNT to gain one additional responder at 6 to 12 weeks with sertraline was 14 (95-percent CI, 8-22).
- *Fluoxetine vs. venlafaxine* (eight studies; 1,814 patients): Patients on venlafaxine had an additional, statistically nonsignificant treatment effect of a 1.31-point reduction (95-percent CI, 0.10-2.39) on the HAM-D scale compared with patients on fluoxetine. The relative risk of response was statistically significantly greater for venlafaxine than for fluoxetine (RR: 1.12; 95-percent CI, 1.01-1.24). The NNT to gain one additional responder at 6 to 12 weeks with venlafaxine was 12 (95-percent CI, 7-50). All studies were funded by the makers of venlafaxine.

Most trials were efficacy trials 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 generalizability 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 (i.e., comparisons of medications conducted across placebo-controlled trials rather than within a single trial) yielded no statistically significant differences in response rates. The precision of some of these estimates was low, leading to inconclusive results with wide confidence intervals. Nevertheless, point estimates of treatment effects from these analyses were consistent with those from direct evidence trials in indicating no or minimal differences in efficacy among available comparisons.

Overall, we rated the strength of the evidence as moderate for both comparative efficacy and comparative effectiveness.

Although second-generation antidepressants appear similar in average efficacy and effectiveness, the studies were not designed to test variation among individuals in their responses to individual drugs. The second-generation antidepressants cannot be considered identical drugs. Evidence of moderate strength supports some differences among individual drugs with respect to onset of action and some measures (e.g., sexual functioning) that could affect health-related quality of life. These are statistically significant but of modest magnitude; potential benefits might be offset by specific adverse events. Nonetheless, some of these differences may influence the choice of a medication for specific patients.

Quality of life. Quality of life or functional capacity was infrequently assessed, usually as a secondary outcome. Eighteen studies (4,050 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 NNT to yield one additional responder after 1 or 2 weeks of treatment is 7 (95-percent CI, 5-12); after 4 weeks of treatment, however, most response rates were similar. Again, this treatment effect was consistent across all studies, but whether this difference can be extrapolated to other second-generation antidepressants remains unclear. The strength of evidence is moderate.

Response to a second agent. The Sequenced Treatment Alternatives to Relieve Depression (STAR-D) trial is the only well-done study looking at the question of response to a second agent among those failing initial therapy. Results show that about one in four of the 727 people who participated in the switch became symptom free, and this did not differ statistically significantly in each of the three medication groups—bupropion sustained release (SR), sertraline, and venlafaxine extended release (XR).

Treatment of Dysthymia

Efficacy and effectiveness. We identified no head-to-head trial comparing different medications in a population with dysthymia. In placebo-controlled trials, significant differences in population characteristics make the evidence insufficient to identify differences between treatments.

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 showed a significantly greater improvement than those on placebo; a subgroup of patients younger than 60 years did not show any difference in effectiveness between paroxetine and placebo. The strength of evidence is low.

Treatment of Subsyndromal Depression

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.

Maintenance of Response or Remission (KQ 2a)

Efficacy and effectiveness. Three head-to-head RCTs suggest that no substantial differences exist between fluoxetine and sertraline, fluvoxamine and sertraline, and trazodone and venlafaxine for maintaining response or remission (i.e., preventing relapse or recurrence of MDD). The strength of the evidence is moderate. Twenty-one placebo-controlled trials support the general efficacy and effectiveness of most second-generation antidepressants for preventing relapse or recurrence. No evidence exists for duloxetine. The overall strength of this evidence is moderate.

Treatment of Treatment-Resistant Depression Syndrome or Relapse or Recurrence (KQ 2b)

Efficacy and effectiveness. One head-to-head efficacy study and two effectiveness studies provide conflicting evidence on differences among second-generation antidepressants in treatment-resistant depression. The efficacy study (fair quality) suggests that venlafaxine is modestly more effective than paroxetine. A good-quality effectiveness study suggests that no substantial differences exist among bupropion SR, sertraline, and

venlafaxine XR, but a fair-quality effectiveness study suggests that venlafaxine is modestly more effective than citalopram, fluoxetine, mirtazapine, paroxetine, and sertraline. Given the conflicting results, the overall strength of the evidence is moderate.

Although several comparative studies included patients who had relapsed or who were experiencing a recurrent depressive episode, no study s*pecifically* 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 3a)

Anxiety. Evidence from six head-to-head trials and one placebo-controlled trial (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 between fluoxetine, paroxetine, and sertraline; sertraline and bupropion; and sertraline and venlafaxine. One trial found statistically significant superiority of venlafaxine over fluoxetine. The strength of evidence is moderate.

Insomnia. Three head-to-head trials that identified a specific insomnia group (all fair quality) provide limited evidence regarding comparative efficacy of medications for treating depression in patients with accompanying insomnia. One trial found statistically significant superiority for escitalopram over citalopram. The strength of evidence is low.

Melancholia. Two head-to-head trials (both fair quality), one poor-quality head-to-head trial, and one fair-quality placebo-controlled study 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 more for venlafaxine than for fluoxetine. The strength of evidence is low.

Pain. One fair-quality trial that required baseline pain for inclusion found no difference in efficacy for duloxetine compared with placebo for treating depression in patients with pain of at least mild intensity. The strength of evidence is low.

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 low.

Somatization. We identified no relevant study.

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

Anxiety. Ten 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 between fluoxetine, paroxetine, and sertraline; sertraline and bupropion; sertraline and venlafaxine; citalopram and mirtazapine; and paroxetine and nefazodone. One trial found that venlafaxine was statistically significantly superior to fluoxetine. The strength of evidence is moderate.

Insomnia. Six head-to-head trials (all fair quality) provide limited evidence about comparative effects of antidepressants on insomnia in patients with depression. The strength of evidence is low.

Melancholia. We identified no relevant study.

Pain. Two head-to-head trials (one of fair and the other of poor quality) and three placebo-controlled trials (all fair quality) provide limited evidence about effects of antidepressants on pain symptoms in depressed patients. Two trials found no substantial difference in efficacy between duloxetine and paroxetine. The strength of evidence is low.

Psychomotor changes. We identified no relevant study.

Somatization. One open-label effectiveness trial found no statistically significant difference among three SSRIs for treating somatization in patients with depression. The strength of evidence is low.

Differences in Harms (Adverse Events) (KQ 4)

We analyzed adverse events data from 72 head-to-head efficacy studies on 16,780 patients, along with data from 39 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.

Adverse events profiles. Constipation, diarrhea, dizziness, headache, insomnia, nausea, and somnolence were commonly and consistently reported adverse events. On average, 61 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 10-percent (95-percent CI, 4-17 percent) higher incidence of nausea and vomiting than SSRIs as a class. In addition, pooled discontinuation rates because of adverse events in efficacy trials are statistically significantly higher for venlafaxine than for SSRIs (RR: 1.50; 95-percent CI, 1.21-1.84). The strength of evidence is high.
- In most studies, sertraline led to higher rates of diarrhea than comparator drugs (bupropion, citalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, venlafaxine). The incidence was 8-percent (95-percent CI, 3-11 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 is moderate.
- Mirtazapine led to higher weight gains than comparator drugs (fluoxetine, paroxetine, venlafaxine, and trazodone). Mean weight gains compared to pretreatment ranged from 0.8 kg to 3.0 kg after 6 to 8 weeks of treatment. Paroxetine had higher weight gains than fluoxetine and sertraline. The strength of evidence 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 is moderate.
- Discontinuation syndromes (e.g., headache, dizziness, nausea) occurred in 0 to 86 percent of patients. Paroxetine and venlafaxine had the highest incidence of this problem, and fluoxetine the lowest incidence. The strength of evidence is moderate.

Discontinuation rates. Overall discontinuation rates did not differ significantly between SSRIs as a class and bupropion, mirtazapine, nefazodone, trazodone, and venlafaxine. In the case of venlafaxine compared with SSRIs, higher discontinuation rates because of adverse events (11.5 percent vs. 8.5 percent) appear to be balanced by lower discontinuation rates because of lack of efficacy (3.5 percent vs. 4.4 percent). The strength of evidence is high.

Severe adverse events.

Sexual dysfunction. Bupropion is associated with a lower incidence of sexual dysfunction than fluoxetine, paroxetine, and sertaline. The NNT to gain one additional person with high overall satisfaction of sexual functioning is 6 (95-percent CI, 4-9). In head-to-head trials, paroxetine consistently had higher rates of sexual dysfunction than comparators (fluoxetine, fluvoxamine, nefazodone, and sertraline; 16 percent vs. 6 percent). Underreporting of absolute rates of sexual dysfunction, however, is likely in these studies. Whether these findings can be extrapolated to comparisons of bupropion and paroxetine with other second-generation antidepressants is unclear. 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 (i.e., elevated systolic and diastolic blood pressure and elevated pulse/heart rate), hyponatremia, hepatotoxicity, and serotonin syndrome, is insufficient to draw firm conclusions. The strength of evidence is low. 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. One observational study indicated that extended-release formulations might have a better adherence rate than immediate-release medications.

This finding, however, is likely attributable more to differences in dosing regimens than to differences in efficacy and harms. 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. The strength of evidence is low.

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

Age. Twelve head-to-head trials (one an effectiveness study), nine placebo-controlled trials, one retrospective cohort study, and one set of meta-analyses suggest that no major differences in efficacy and effectiveness exist among second-generation antidepressants in elderly or very elderly populations. The strength of evidence is moderate.

Harms such as hyponatremia and weight loss may differ in elderly or very elderly patients on active treatment vs. placebo, but the evidence on these two adverse events is limited to one small RCT and one observational study (both fair quality). The strength of evidence is low.

Sex. Indirect evidence from one fair-quality pooled analysis of head-to-head RCTs suggests that efficacy among second-generation antidepressants does not differ between men and women. This conclusion is supported by observational evidence. One fair-quality observational study indicated that harms, specifically the rates of sexual dysfunction, might differ between men and women. The strength of evidence is low.

Race or ethnicity. One poor-quality RCT suggests that the efficacy of second-generation antidepressants does not differ for patients in different race or ethnic groups. This study, however, may not have been powered to detect a difference. The strength of evidence is low.

Comorbidities. The evidence for various comorbidities (e.g., HIV/AIDS, alcohol abuse, Alzheimer's disease or other dementia, breast cancer, cardiovascular disease, stroke, and substance abuse) is limited to one head-to head study, a small number of placebo-controlled trials, and one systematic review. They provide limited evidence on the comparative efficacy of second-generation antidepressants in subgroups with different coexisting conditions. The strength of evidence is low.

Remaining Issues

We found no studies that identified reliable predictors of individual responses to a specific drug based on patients' clinical, demographic, or genetic characteristics. Owing to a substantial nonresponse rate to individual drugs and generally high incidence of side effects, many patients try multiple antidepressant medications before finding an effective, well-tolerated drug, but predicting which drug will be most effective or best tolerated in any given individual is not yet possible. Studies of tailoring therapy would have been eligible for this review, but we did not find any. Most of the included studies looked only at average effectiveness, excluded subjects with comorbidities, and did not examine differences in effectiveness according to broad demographic characteristics.

Effectiveness studies that would be most applicable to the broad population of depressed patients are generally lacking for most drugs. Effectiveness trials with less stringent eligibility criteria, patient-centered health outcomes, long study durations, and populations representative of patients encountered in primary care would be valuable to determine whether existing differences of second-generation antidepressants are clinically meaningful in "real-world" settings. These trials should be powered to be able to assess minimal clinically significant differences. Furthermore, they could provide valuable information on differences in adherence among second-generation antidepressants.

Major Depressive Disorder

Although the strength of evidence is moderate for the comparative efficacy for treating MDD during the acute phase, more evidence is needed to resolve whether one second-generation antidepressant is better than another in patients who either did not respond or could not tolerate a first-line treatment. In efficacy trials, on average, 38 percent of patients did not achieve a treatment response, and 54 percent did not achieve remission. The STAR-D trial is the best available evidence so far, but its results are limited to bupropion SR, sertraline, and venlafaxine XR.

Given the fact 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 predict better 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.

In addition, more evidence is needed regarding the most appropriate duration of antidepressant treatment for maintaining response and remission. Such studies should also evaluate 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. The role of other depression treatments, such as psychotherapy, vagal nerve stimulations, light therapy, and complementary medicines, as substitutes or complements to pharmaceutical management also needs to be better understood.

More research is also needed to evaluate whether 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 more accurate measures to identify these subgroups. Likewise, future research has to clarify differences of second-generation antidepressants in subgroups based on age, race, and common comorbidities.

Dysthymia and Subsyndromal Depression

Future research has to establish reliably the general efficacy of second-generation antidepressants for the treatment of dysthymia and subsyndromal depression. Ideally, multiple-arm, head-to-head trials, including placebo groups, should evaluate the general and comparative efficacy of second-generation antidepressants for treating these two conditions. If general efficacy can be established reliably, differences in subgroups based on accompanying symptoms, demographic characteristics, or comorbidities should be explored.

Addendum

As this report was going to press, a relevant study addressing sequential treatment steps among patients who did not obtain remission with initial acute-phase treatment was published. We were unable to incorporate this study fully into this report, but we found its results important in light of the general lack of high-quality evidence for treating patients who do not obtain remission with initial treatments.

The STAR-D trial–described in detail in the discussion of Key Question 2b (in the main report)–consisted of a series of RCTs examining sequential treatment steps in patients who did not obtain remission or could not tolerate previous treatments. Key Question 2b detailed the medication switch arms of the second-step treatment in which all patients in the analysis had failed initial treatment with citalopram and were randomized to second-step treatment with bupropion SR (N = 239), sertraline (N = 238), or venlafaxine XR (N = 250); this analysis found no statistically significant differences in remission rates between second-step treatments.

The more recently published study describes the acute and longer term outcomes associated with all four treatment steps. Patients not achieving remission or unable to tolerate a treatment step were encouraged to move to the next step; patients achieving acceptable benefit could enter a 12-month followup phase. All patients (N = 3,671) received citalopram in Step 1. Step 2 and Step 3 treatments were randomly assigned using an equipoise stratified randomized design. In this, 1,439 patients were randomized in Step 2, which included seven possible treatment alternatives (bupropion SR, sertraline, venlafaxine XR, cognitive therapy, citalopram plus bupropion, citalopram plus buspirone, or citalopram plus cognitive therapy). Step 3 randomized 390 patients to switch to mirtazapine or nortriptyline or to receive augmentation with lithium or triiodothyronine (T3). Step 4 used only a single randomization; 123 patients were randomized to tranylcypromine or venlafaxine XR plus mirtazapine.

Overall, 67 percent of patients achieved remission. Remission rates were 36.8 percent for Step 1, 30.6 percent for Step 2, 13.7 percent for Step 3, and 13.0 percent for Step 4. For patients achieving acceptable benefits who continued on in the 12-month followup

study, relapse rates were 40.1 percent, 55.3 percent, 64.6 percent, and 71.1 percent for those achieving benefit in Steps 1, 2, 3, and 4, respectively. In all steps, patients achieving remission (Quick Inventory of Depressive Symptomatology–Self Report [QIDS-SR-16] ≤ 5) were less likely to relapse than patients not achieving remission (acceptable benefit but QIDS-SR-16 > 5).

Full Report

This executive summary is part of the following document: Gartlehner G, Hansen RA, Thieda P, DeVeaugh-Geiss AM, Gaynes BN, Krebs EE, Lux LJ, Morgan LC, Shumate JA, Monroe LG, Lohr KN. Comparative Effectiveness of Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression. Comparative Effectiveness Review No. 7. (Prepared by RTI International-University of North Carolina Evidence-based Practice Center under Contract No. 290-02-0016.) Rockville, MD: Agency for Healthcare Research and Quality. January 2007. Available at: www.effectivehealthcare.ahrq.gov/reports/final.cfm.

For More Copies

For more copies of Comparative Effectiveness of Second-Generation Antidepressants in the Pharmocologic Treatment of Adult Depression: Executive Summary. No. 7. (AHRQ Pub. No. 07-EHC007-1), please call the AHRQ Clearinghouse at 1-800-358-9295 or e-mail ahrqpubs@ahrq.gov.

Table A. Summary of findings on treatment of adult depression with strength of evidence

Key question, disorder, and outcome of interest	Strength of evidence[1]	Findings[2]
Key Question 1a. Comparative efficacy and effectiveness of second-generation antidepressants		
Major depressive disorders		
Comparative efficacy	Moderate	Results from direct and indirect comparisons indicate that no substantial differences in efficacy exist among second-generation antidepressants.
Comparative effectiveness	Moderate	Direct evidence from 1 good and 2 fair effectiveness studies 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 studies, most of fair quality, 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 7 fair 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.
Dysthymia		
Comparative efficacy	Low	No head-to head evidence exists. Findings from 5 placebo-controlled trials were insufficient to draw conclusions about comparative efficacy.
Comparative effectiveness	Low	One fair effectiveness study provides mixed evidence about paroxetine vs. placebo; patients older than 60 showed greater improvement on paroxetine; those younger than 50 did not show any difference.
Quality of life	No evidence	
Onset of action	No evidence	
Subsyndromal depression		
Comparative efficacy	Low	One nonrandomized, open-label trial did not detect any difference between citalopram and sertraline. Findings from 2 placebo-controlled trials were insufficient to draw conclusions.
Comparative effectiveness	No evidence	
Quality of life	No evidence	
Onset of action	No evidence	
Key Question 1b: Greater efficacy and effectiveness with previously effective medications		
Major depressive disorder	No evidence	
Dysthymia	No evidence	
Subsyndromal depression	No evidence	

Table A. Summary of findings on treatment of adult depression with strength of evidence (continued)

Key question, disorder, and outcome of interest	Strength of evidence[1]	Findings[2]
Key Question 2a: Efficacy and effectiveness of second-generation antidepressants for maintaining response or remission (i.e., preventing relapse or recurrence)		
Comparative efficacy	Moderate	Based on findings from 3 efficacy trials, no significant differences exist between fluoxetine and sertraline, fluvoxamine and sertraline, and trazodone and venlafaxine for preventing relapse or recurrence. Whether this finding can be extrapolated to other second-generation antidepressants is unclear.
Comparative effectiveness	No evidence	
General effectiveness/efficacy	Moderate	Based on findings from 21 placebo-controlled trials, second-generation antidepressants are effective for preventing relapse or recurrence.
Key Question 2b: Efficacy and effectiveness of second-generation antidepressants in managing treatment-resistant depression syndrome or treating recurrent depression		
Managing treatment-resistant depression		
Comparative efficacy	Low	Results from 1 fair trial support modestly better efficacy for venlafaxine compared with paroxetine.
Comparative effectiveness	Moderate	Results from 2 effectiveness studies are conflicting. Based on 1 good trial, no significant differences in effectiveness exist among bupropion SR, sertraline, and venlafaxine XR. One fair effectiveness trial found venlafaxine to be modestly superior to citalopram, fluoxetine, mirtazapine, paroxetine, and sertraline.
General effectiveness/efficacy	Low	No placebo-controlled evidence was identified. Uncontrolled, open-label evidence supports the general efficacy of second-generation antidepressants.
Treating recurrent depression		
Comparative efficacy	No evidence	
Comparative effectiveness	No evidence	
Key Question 3a: Comparative efficacy and effectiveness of second-generation antidepressants for treatment of depression in patients with accompanying symptom clusters		
Anxiety		
Comparative efficacy	Moderate	Results from 6 head-to-head trials and 1 placebo-controlled trial (all fair quality) suggest that efficacy does not differ substantially for treatment of depression in patients with accompanying anxiety.
Comparative effectiveness	No evidence	

Table A. Summary of findings on treatment of adult depression with strength of evidence (continued)

Key question, disorder, and outcome of interest	Strength of evidence[1]	Findings[2]
Key Question 3a: Comparative efficacy and effectiveness of second-generation antidepressants for treatment of depression in patients with accompanying symptom clusters (continued)		
Insomnia		
Comparative efficacy	Low	Evidence from 3 fair head-to-head studies is insufficient to draw conclusions about treating depression in patients with coexisting insomnia. Results are limited by study design.
Comparative effectiveness	No evidence	
Melancholia		
Comparative efficacy	Low	Evidence from 2 fair head-to-head studies, 1 poor head-to-head study, and 1 fair placebo-controlled trial is insufficient to draw conclusions about treating depression in patients with coexisting insomnia. Results are inconsistent across studies.
Comparative effectiveness	No evidence	
Pain		
Comparative efficacy	Low	Evidence from 1 fair placebo-controlled study is insufficient to draw conclusions about treating depression in patients with coexisting pain. Results from head-to-head trials are not available.
Comparative effectiveness	No evidence	
Psychomotor change		
Comparative efficacy	Low	Evidence from 1 fair 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 be different.
Comparative effectiveness	No evidence	
Somatization		
Comparative efficacy	No evidence	
Comparative effectiveness	No evidence	
Key Question 3b: Comparative efficacy and effectiveness of second-generation antidepressants for treatment of symptom clusters in patients with depression		
Anxiety		
Comparative efficacy	Moderate	Results from 10 fair head-to-head trials and 2 fair placebo-controlled trials suggest that no substantial differences in efficacy exist among second-generation antidepressants for treatment of accompanying anxiety symptoms.
Comparative effectiveness	No evidence	

Table A. Summary of findings on treatment of adult depression with strength of evidence (continued)

Key question, disorder, and outcome of interest	Strength of evidence[1]	Findings[2]
Key Question 3b: Comparative efficacy and effectiveness of second-generation antidepressants for treatment of symptom clusters in patients with depression (continued)		
Insomnia		
Comparative efficacy	Low	Evidence from 6 fair head-to-head trials is insufficient to draw conclusions about treating insomnia in depressed patients. Results are limited by study design; differences in outcomes are of unknown clinical significance.
Comparative effectiveness	No evidence	
Melancholia		
Comparative efficacy	No evidence	
Comparative effectiveness	No evidence	
Pain		
Comparative efficacy	Low	Evidence from 2 head-to-head trials (1 fair, 1 poor) and 3 placebo-controlled trials is insufficient to draw conclusions about treating coexisting pain in depressed patients. Results indicate no difference in efficacy but are limited by study design.
Comparative effectiveness	No evidence	
Psychomotor change		
Comparative efficacy	No evidence	
Comparative effectiveness	No evidence	
Somatization		
Comparative efficacy	No evidence	
Comparative effectiveness	Low	Evidence from 1 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.
Key Question 4: Comparative risk of harms		
General tolerability		
Adverse events profiles	High	Adverse events profiles are similar among second-generation antidepressants. Incidence rates of specific adverse events differ.
Nausea and vomiting	High	Meta-analysis of 15 fair studies indicates that venlafaxine has a higher rate of nausea and vomiting than SSRIs as a class.

Table A. Summary of findings on treatment of adult depression with strength of evidence (continued)

Key question, disorder, and outcome of interest	Strength of evidence[1]	Findings[2]
Key Question 4: Comparative risk of harms (continued)		
General tolerability (continued)		
Diarrhea	Moderate	Evidence from 15 fair studies indicates that sertraline has a higher incidence of diarrhea than bupropion, citalopram, fluoxetine, fluvoxamine, mirtazapine, nefazodone, paroxetine, and venlafaxine.
Weight change	Moderate	Seven fair trials indicate that mirtazapine leads to higher weight gains than citalopram, fluoxetine, paroxetine, and sertraline.
Somnolence	Moderate	Six fair studies provide evidence that trazodone has a higher rate of somnolence than bupropion, fluoxetine, mirtazapine, paroxetine, and venlafaxine.
Discontinuation syndrome	Moderate	A good systematic review provides evidence that paroxetine and venlafaxine have the highest rates of discontinuation syndrome; fluoxetine has the lowest.
Discontinuation rates	High	Meta-analyses of efficacy trials indicate that overall discontinuation rates are similar. Venlafaxine has a higher rate of discontinuations from adverse events and a lower rate of discontinuations from lack of efficacy than SSRIs as a class.
Severe adverse events		
Sexual dysfunction	Moderate	Evidence from 5 fair trials provides evidence that bupropion causes significantly less sexual dysfunction than fluoxetine, paroxetine, and sertraline. Among SSRIs, paroxetine has the highest rates of sexual dysfunction.
Suicidality	Low	Evidence from existing studies is insufficient to draw conclusions about the comparative risk of suicidality.
Seizures	Low	Evidence from existing studies is insufficient to draw conclusions about the comparative risk of seizures. Weak evidence indicates that bupropion may have an increased risk of seizures.
Cardiovascular events	Low	Evidence from existing studies is insufficient to draw conclusions about the comparative risk of cardiovascular adverse events. Weak evidence indicates that venlafaxine might have an increased risk of cardiovascular adverse events.
Hyponatremia	Low	The evidence is insufficient to draw conclusions about the comparative risk of hyponatremia.
Hepatotoxicity	Low	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.

Table A. Summary of findings on treatment of adult depression with strength of evidence (continued)

Key question, disorder, and outcome of interest	Strength of evidence[1]	Findings[2]
Key Question 4: Comparative risk of harms (continued)		
Severe adverse events (continued)		
Serotonin syndrome	Low	Evidence from existing studies is insufficient to draw conclusions about the comparative risk of serotonin syndrome. Observational studies indicate no differences in risk among second-generation antidepressants.
Adherence in efficacy studies	Moderate	Efficacy studies indicate no differences in adherence. One observational study suggests that extended-release formulations may have better adherence than immediate-release formulations.
Adherence in effectiveness studies	Low	Evidence from existing studies is insufficient to draw conclusions about adherence in "real-world" settings.
Key Question 5: Selected populations		
Age		
Comparative efficacy	Moderate	Results from numerous different types of studies indicate that no substantial differences exist in efficacy among second-generation antidepressants in the elderly or the very elderly.
Comparative effectiveness	Moderate	Based on findings from 1 fair head-to-head effectiveness trial, no substantial differences exist among second-generation antidepressants in the elderly compared with other age groups. A second trial in patients with dysthymia or minor depression provides mixed evidence.
Comparative harms	Low	Results from 2 fair studies indicate that adverse events may differ somewhat across second-generation antidepressants in the elderly or very elderly.
Sex		
Comparative efficacy	Low	Results from 1 fair pooled analysis of RCTs indicates that efficacy among second-generation antidepressants may not differ substantially between men and women.
Comparative effectiveness	No evidence	
Comparative harms	Low	One fair head-to-head trial suggests that harms (headache, nausea) may differ between men and women treated with venlafaxine vs. placebo and venlafaxine vs. SSRIs or placebo. Observational evidence (1 fair study) suggests that some sexual side effects may differ between men and women.

Table A. Summary of findings on treatment of adult depression with strength of evidence (continued)

Key question, disorder, and outcome of interest	Strength of evidence[1]	Findings[2]
Key Question 5: Selected populations (continued)		
Race or ethnicity		
Comparative efficacy	Low	Results from 1 poor RCT indicate that efficacy does not differ substantially among second-generation antidepressants in different racial subgroups.
Comparative effectiveness	No evidence	
Comparative harms	No evidence	
Comorbidities		
Comparative efficacy	Low	One poor head-to-head trial included patients with depression and HIV/AIDS; this study indicated that efficacy does not differ substantially among second-generation antidepressants. Findings from placebo-controlled trials were insufficient to draw conclusions about comparative efficacy.
Comparative effectiveness	No evidence	
Comparative harms	No evidence	

[1] Strength of evidence is based on a modified version of the GRADE system; see text above.
[2] Good, fair, or poor designations relate to quality grades given to each study.

Abbreviations: RCT = randomized controlled trial; SR = sustained release; SSRI = selective serotonin reuptake inhibitor; XR = extended release.

AHRQ Pub. No. 07-EHC007-1
January 2007