

Nonpharmacological Versus Pharmacological Treatments for Adult Patients With Major Depressive Disorder

(Reprinted with permission from the Agency for Healthcare Research)

BACKGROUND

Major depressive disorder (MDD)¹ is the most prevalent and disabling form of depression, affecting more than 16 percent of U.S. adults (lifetime).² MDD can be characterized as mild, moderate, or severe based on symptom severity, functional impairment, and level of patient distress;¹ in clinical trials, these distinctions are typically made by scores on a depressive rating instrument.³ Approximately one-third of patients with MDD are severely depressed,⁴ which is associated with depression that is harder to treat, as evidenced by more difficulty in achieving treatment response and remission.⁵

In any given year, nearly 7 percent of the U.S. adult population (approximately 17.5 million people in 2014) experience an episode of MDD that warrants treatment.² Most patients receiving care obtain treatment in primary care settings,⁶ where second-generation antidepressants (SGAs) are the most commonly prescribed agents.⁷ Nonetheless, patients and clinicians may prefer other options, or at least want to be able to consider them. These include psychological interventions, complementary and alternative medicine (CAM) options, and exercise.

The psychological interventions used to treat depressed patients include acceptance and commitment therapy, cognitive therapy (CT), cognitive behavioral therapy (CBT), interpersonal therapy, and psychodynamic therapies. Commonly used CAM interventions for the treatment of patients with MDD include acupuncture, meditation, omega-3 fatty acids, S-adenosyl-L-methionine (SAME), St. John's wort, and yoga. While acupuncture requires a licensed professional for treatment, the other options may be used in conjunction with a trained provider or be self-administered.

Exercise covers a broad range of activities; they can be done over varying durations of time and singly, in classes, or in informal groups.

About 40 percent of patients treated with SGAs do not respond to initial treatment; approximately 70 percent do not achieve remission during the first-step treatment.⁸ Those who do not achieve remission following initial pharmacological treatment require a different treatment strategy. Accordingly, various other interventions—such as medication combinations, psychotherapy, or CAM treatments—are options for patients and clinicians.

SCOPE AND KEY QUESTIONS

This review for the Effective Health Care Program of the Agency for Healthcare Research and Quality (AHRQ) examines the evidence base for primary care management of MDD for the

first two treatment attempts, after which primary care clinicians would consider referral to or consultation by a mental health professional. The specific Key Questions (KQs) are listed below:

- **KQ 1a.** In adult patients with major depressive disorder (MDD) who are undergoing an initial treatment attempt, what is the effectiveness of second-generation antidepressant (SGA) monotherapy compared with the effectiveness of either nonpharmacological monotherapy or combination therapy (involving nonpharmacological treatments with or without an SGA)?
- **KQ 1b.** Does comparative treatment effectiveness vary by MDD severity?
- **KQ 2a.** In adult patients with MDD who did not achieve remission following an initial adequate trial with one SGA, what is the comparative effectiveness of second-step therapies?^a
- **KQ 2b.** Does comparative treatment effectiveness vary by MDD severity?
- **KQ 3a.** In adult patients with MDD, what are the comparative risks of harms of these treatment options—
 1. For those undergoing an initial treatment attempt?
 2. For those who did not achieve remission following an initial adequate trial with an SGA?
- **KQ 3b.** Do the comparative risks of treatment harms vary by MDD severity?
- **KQ 4.** Do the benefits and risks of harms of these treatment options differ by subgroups of patients with MDD defined by common accompanying psychiatric symptoms (coexisting anxiety, insomnia, low energy, or somatization) or demographic characteristics (age, sex, race, or ethnicity)?

METHODS

Literature Search Strategy

We searched MEDLINE® (via PubMed®), Embase®, the Cochrane Library, AMED (Allied and Complementary Medicine Database), PsycINFO®, and CINAHL (Cumulative Index to Nursing and Allied Health Literature) from January 1, 1990, through January 13, 2015. We used a combination of medical subject headings and title and abstract keywords, focusing on terms to describe the relevant population and interventions of interest. We limited the electronic searches to English-, German-, and Italian-language and human-only studies.

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.

In addition, we manually searched reference lists of pertinent reviews, included trials, and background articles, and searched for gray literature relevant to this review following guidance from the *Methods Guide for Effectiveness and Comparative Effectiveness Reviews* for these steps.⁹

Inclusion and exclusion criteria are presented in Table 1.

Two trained research team members independently reviewed all titles, abstracts, and eligible full-text articles. We designed, pilot tested, and used a structured data abstraction form to ensure consistency of data abstraction. Trained reviewers initially abstracted data from each study. A senior reviewer then read each abstracted article and evaluated the completeness and accuracy of the data abstraction. We resolved discrepancies by consensus or by involving a third, senior reviewer.

Risk-of-Bias Assessment of Individual Studies

To assess the risk of bias of studies, we used definitions based on AHRQ guidance.¹⁰ We rated the risk of bias for each relevant outcome of a study as low, moderate, or high. To determine risk of bias in a standardized way, we used the Cochrane Risk of Bias tool to appraise randomized controlled trials (RCTs).¹¹ Two independent reviewers assigned risk-of-bias ratings. They resolved any disagreements by discussion and consensus or by consulting a third, independent party.

Data Synthesis

Throughout this review we synthesized the literature qualitatively. When data were sufficient, we augmented findings with quantitative analyses.

For meta-analyses, we used random-effects (DerSimonian-Laird) and fixed-effects models to estimate comparative effects. We assessed statistical heterogeneity in effects between

studies by calculating the chi-squared statistic and Cochran's Q . We used the I^2 statistic to estimate the magnitude of heterogeneity. We examined potential sources of heterogeneity using sensitivity analysis or analysis of subgroups. We assessed publication bias by checking study registries and using funnel plots and Kendall's tests. However, given the small number of component studies in our meta-analyses, these tests have low sensitivity to detect publication bias.

Because of the dearth of studies directly comparing interventions of interest, we planned network meta-analyses *a priori*. Our outcome measure of choice was the rate of response on the Hamilton Depression Rating Scale (HAM-D), defined as at least a 50-percent improvement of scores from baseline. We included all placebo- and active-controlled RCTs detected through our searches that were homogeneous in study populations and outcome assessments and were part of a connected network. We employed a hierarchical frequentist approach using random-effects models.^{12,13}

Strength of the Body of Evidence

We graded the strength of evidence (SOE) based on AHRQ guidance established for the Evidence-based Practice Centers.¹⁴ This approach incorporates five key domains: risk of bias, consistency, directness, precision, and reporting bias. Grades (high, moderate, low, insufficient) reflect the strength of the body of evidence for a specific outcome on the comparative benefits and harms of the interventions in this review. During the protocol development, we asked the Technical Expert Panel and the Key Informants to rank the relative importance of outcomes following a process proposed by the GRADE (Grading of Recommendations Assessment, Development and Evaluation) Working Group.¹⁵ We graded only those outcomes that Technical Expert Panel members and Key Informants deemed as important or critical for decisionmaking.

Applicability

We assessed applicability of the evidence following guidance from the *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*.¹⁶ We used the PICOTS (populations, interventions, comparators, outcomes, timing, settings) framework to explore factors that may affect applicability.

RESULTS

We documented the outputs of our literature searches and then described included trials in general terms. We also summarized findings by KQ, dealing with KQ 1 (benefits) and KQ 3 (harms) together, and organized the findings by intervention comparisons.

Results of Literature Searches

Our search strategies identified 7,813 possible articles. We excluded 7,368 references following independent dual title and abstract review, and another 390 references at the

full-text review stage. Reasons for exclusion were based on eligibility criteria. Overall, we included 44 trials reported in 55 published articles. Of these, 42 trials pertained to KQ 1a and 5 to KQ 1b. Two trials pertained to KQ 2a, and no trials were identified for KQ 2b. In addition, of the 44 trials, 43 trials pertained to KQ 3a and 1 to KQ 3b; 3 pertained to KQ 4.

For network meta-analyses, we included data from 85 additional published trials and 27 unpublished trials. These trials addressed comparisons of interventions of interest that did not meet eligibility criteria for this report; they did, however, provide common comparators that we could use for network meta-analyses.

Effectiveness and Harms of Treatment Options for Initial Treatment of Patients With Major Depressive Disorder

In all, 42 trials comparing SGAs with nonpharmacological treatment options for MDD provided direct evidence on acute-phase outcomes. Study durations ranged from 4 to 96 weeks. Most patients suffered from moderate to severe major depression. Many of the available trials had serious methodological limitations. Additionally, few trials adequately assessed harms or reported information on quality of life or functional capacity. The figures provide graphical overviews of response rates (Figure A) and discontinuation rates because of adverse events (Figure B) of SGAs compared with psychological interventions, CAM therapies, and exercise.

Second-Generation Antidepressants Compared With Psychological Interventions.

Second-Generation Antidepressants Versus Cognitive Behavioral Therapy. We identified 11 trials (1,566 participants) of interventions categorized by the Cochrane Collaboration Depression, Anxiety and Neurosis (CCDAN) Group Topic List as cognitive behavioral therapies. (Note that numbers do not sum to 11 because of studies with multiple CBT arms. The CCDAN Topic List is shown in Appendix B of the full report.) Six trials employed CBT, four used CT, and one each used problem-solving therapy and rational emotive behavior therapy. Three trials included a combination SGA plus CBT arm. Overall, SGAs and CBT monotherapies led to similar rates of response to treatment (moderate SOE), remission (HAM-D-17 ≤ 7) (low SOE), and overall discontinuation in patients with moderate to severe MDD after 8 to 16 weeks of followup (moderate SOE). After 24 weeks of followup, however, SGAs led to higher rates of overall discontinuation than CBT (low SOE). Rates of discontinuation because of adverse events following SGAs or CBT were not statistically different (low SOE).

Adding CBT to SGA did not show any benefit in remission or response, as defined previously, and led to similar rates of both overall discontinuation (low SOE) and discontinuation due to adverse events compared with SGA monotherapy (low SOE). The evidence was insufficient to draw conclusions about differences in functional capacity, quality of life,

overall risk of adverse events, suicidal ideas or behaviors, or overall risk of serious adverse events.

Second-Generation Antidepressants Versus Integrative Therapies. The only type of integrative therapy used in the included studies was interpersonal psychotherapy. We identified four trials (872 participants) that compared SGA monotherapy with interpersonal psychotherapy alone. One trial also examined the effect of adding interpersonal psychotherapy to the SGA regimen.

SGAs and interpersonal psychotherapy did not lead to statistically different response or remission rates (HAM-D-17 and HAM-D-21 ≤ 7) (low SOE). The evidence was insufficient to draw conclusions about differences in suicidal ideas or behaviors, overall risk of adverse events, overall risk of serious adverse events, rates of overall discontinuation, or rates of discontinuation because of adverse events. The combination of SGA and interpersonal psychotherapy had 25-percent higher remission rates than SGA monotherapy (low SOE).

Overall discontinuation rates were similar for SGA monotherapy and the combination of SGA and interpersonal therapy (low SOE). The evidence was insufficient to draw conclusions about differences in functional capacity, quality of life, overall risk of adverse events, suicidal ideas or behaviors, overall risk of serious adverse events, or discontinuation because of adverse events.

Second-Generation Antidepressants Versus Psychodynamic Therapies. Three trials (298 participants) compared SGA monotherapy with short-term (2 to 4 months) psychodynamic therapies (PSYD). One trial (272 participants) compared SGA monotherapy with long-term (24 months) PSYD; that study also examined the effect of adding long-term PSYD to the SGA regimen. SGA monotherapy and short-term PSYD monotherapy did not lead to statistically different rates of remission (HAM-D-17 ≤ 7) (low SOE) or improvements in functional capacity (low SOE). SGAs and PSYD also led to similar rates of overall discontinuation over 8 to 16 weeks (low SOE), 48 weeks (low SOE), and 96 weeks of followup (low SOE). The evidence was insufficient to draw conclusions about differences in quality of life, overall risk of adverse events, overall risk of serious adverse events, or discontinuation due to adverse events.

Adding long-term (96 weeks) PSYD to SGA treatment led to lower rates of overall discontinuation after 96 weeks of followup compared with SGA monotherapy (low SOE). Suicidal ideas or behaviors did not differ statistically for patients on SGAs, long-term PSYD, or a combination of the two (low SOE). The evidence was insufficient to draw conclusions about differences in functional capacity, quality of life, overall risk of adverse events, overall risk of serious adverse events, or discontinuation due to adverse events.

Second-Generation Antidepressants Versus Third-Wave Cognitive Behavioral Therapy. Two randomized trials (243 participants) compared treatment with an SGA versus treatment with behavioral activation, a type of third-wave cognitive behavioral therapy. Patients on SGAs had nearly three times higher rates of overall discontinuation (low SOE) and

TABLE 1. Inclusion/Exclusion Criteria

Criteria	Inclusion	Exclusion
Population	Adult (18 years or older) outpatients of all races and ethnicities with MDD during either an initial treatment attempt or a second treatment attempt in patients who did not have remission following an initial adequate trial with an SGA	<ul style="list-style-type: none"> • Children under age 18 • Patients with perinatal depression, seasonal affective disorder, psychotic depression, or treatment-resistant depression (i.e., 2 or more failures of treatment)
Interventions	<p>Second-generation antidepressants:^a</p> <ul style="list-style-type: none"> • Bupropion • Citalopram • Desvenlafaxine • Duloxetine • Fluoxetine • Escitalopram • Fluvoxamine • Levomilnacipran • Mirtazapine • Nefazodone • Paroxetine • Sertraline • Trazodone • Venlafaxine • Vilazodone • Vortioxetine <p>Common depression-focused psychotherapies:</p> <ul style="list-style-type: none"> • Behavioral therapies/behavior modification • Cognitive behavioral therapies • Integrative therapies (e.g., interpersonal therapy) • Psychodynamic therapies • Third-wave cognitive behavioral therapies <p>Complementary and alternative medicines:</p> <ul style="list-style-type: none"> • Acupuncture • Meditation (e.g., mindfulness-based stress reduction) • Omega-3 fatty acids • S-adenosyl-L-methionine (SAME) • St. John's wort (<i>Hypericum perforatum</i>) • Yoga <p>Exercise:</p> <ul style="list-style-type: none"> • Any formal exercise program <p>Other pharmacotherapies for combination or augmentation:</p> <ul style="list-style-type: none"> • Atypical antipsychotics (aripiprazole, asenapine maleate, clozapine, iloperidone, lurasidone, olanzapine, paliperidone, quetiapine, risperidone, ziprasidone) • Psychostimulants (amphetamine-dextroamphetamine, armodafinil, dexamethylphenidate, dextroamphetamine, lisdexamfetamine, methylphenidate, modafinil) • Buspirone • Levothyroxine (T4) • Lithium • Pindolol • Triiodothyronine (T3) 	<p>Ineligible interventions</p>
Control interventions	<p>For all populations of interest (i.e., KQ 1, KQ 3, and KQ 4):</p> <ul style="list-style-type: none"> • SGAs vs. psychotherapies • SGAs vs. CAM • SGAs vs. exercise • SGAs vs. SGA + psychotherapies 	<p>Ineligible interventions, such as placebo arms</p>

continued

TABLE 1, *continued*

Criteria	Inclusion	Exclusion
	<ul style="list-style-type: none"> • SGAs vs. SGA + CAM • SGAs vs. SGA + exercise • SGAs vs. combinations of eligible interventions <p>In addition, for populations who did not have remission following an initial adequate trial with an SGA (i.e., KQ 2, KQ 3, and KQ 4):</p> <ul style="list-style-type: none"> • SGA switch^b vs. SGA switch • SGA switch^b vs. nonpharmacological treatment • SGA switch^b vs. SGA augmentation^c • SGA augmentation^c vs. SGA augmentation • SGA augmentation^c vs. nonpharmacological treatment <p>In addition, for network meta-analyses:</p> <ul style="list-style-type: none"> • Placebo or other inactive control • Comparisons of eligible interventions without an SGA arm 	
Outcomes	<ul style="list-style-type: none"> • Benefits: response to treatment, remission, speed of response, speed of remission, relapse, quality of life, functional capacity, reduction of suicidal ideas or behaviors, reduction of hospitalization • Harms: overall adverse events, withdrawals because of adverse events, serious adverse events, specific adverse events (including hyponatremia, seizures, suicidal ideas or behaviors, hepatotoxicity, weight gain, gastrointestinal symptoms, sexual side effects), withdrawals because of specific adverse events, or drug interactions (pharmacological and complementary and alternative treatments) 	Studies that do not include at least 1 of the outcomes listed under the inclusion criteria
Timing of intervention	No limitations	Not applicable
Publication language	English, German, Italian	All other languages
Study design	<ul style="list-style-type: none"> • Original research • Eligible study designs include— • For efficacy/effectiveness: <ul style="list-style-type: none"> – RCTs – SRs and meta-analyses • In addition, for harms:^d <ul style="list-style-type: none"> – Nonrandomized controlled trials – Prospective controlled cohort studies – Retrospective controlled cohort studies – Case-control studies 	Case series <ul style="list-style-type: none"> • Case reports • Nonsystematic reviews • Studies without a control group • Nonrandomized studies with fewer than 500 participants • Post hoc or secondary analyses • Pooled studies
Publication type	Any publication reporting primary data	Publications not reporting primary data

CAM = complementary and alternative medicine; KQ = Key Question; MDD = major depressive disorder; RCT = randomized controlled trial;

SGA = second-generation antidepressant; SR = systematic review

^a SGAs approved for treatment of MDD by the Food and Drug Administration.

^b Switching to another SGA.

^c Augmenting with a second SGA, an additional non-SGA medication, or a nonpharmacological treatment.

^d Nonrandomized studies must have a minimum sample size of 500 participants.

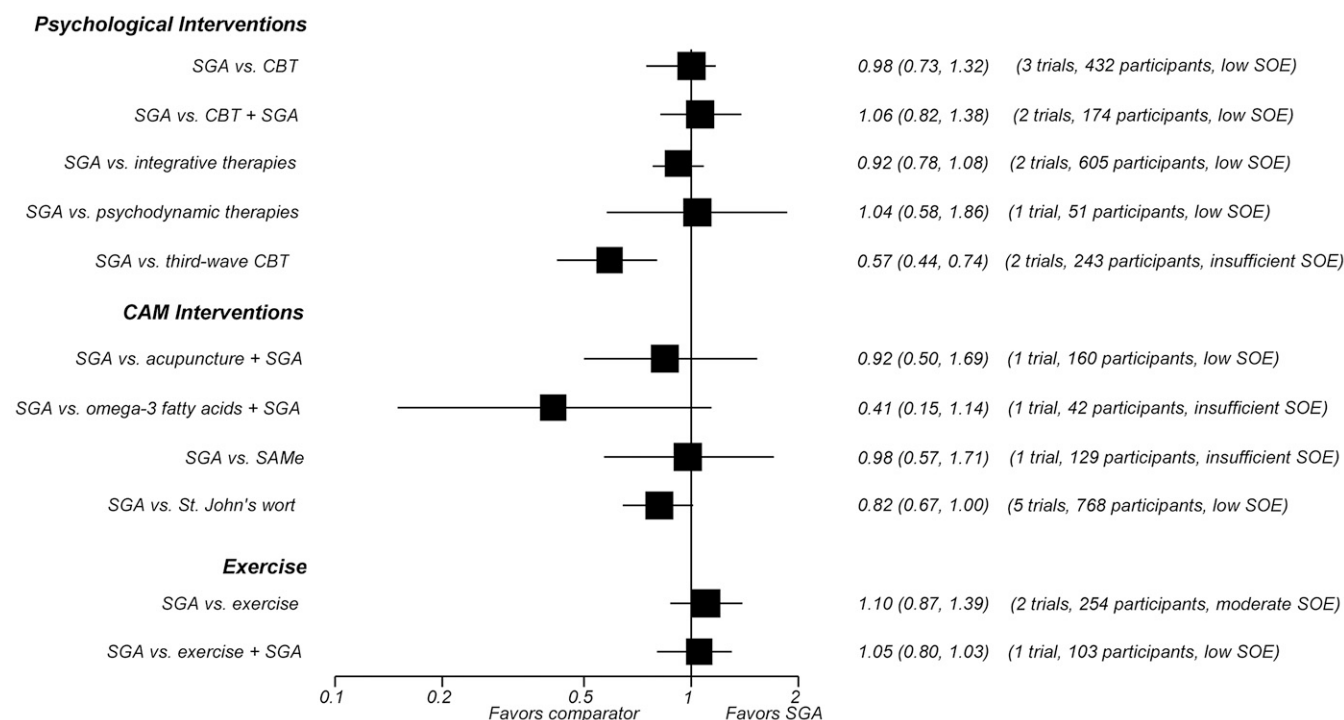
more than five times higher rates of discontinuation because of adverse events than those treated with behavioral activation (low SOE). The evidence was mixed with regard to response and remission, and was insufficient to draw conclusions about differences in response, remission, functional capacity, quality of life, overall risk of adverse events, overall risk of serious adverse events, or suicidal ideas or behaviors.

Severity as a Moderator of Comparative Treatment Effectiveness. Four trials yielded insufficient evidence to

determine whether the comparative effectiveness of SGAs versus any psychological treatment changes as a function of MDD severity.

Second-Generation Antidepressants Compared With Complementary and Alternative Medicine Interventions.

Second-Generation Antidepressants Versus Acupuncture. Three trials (263 participants), all conducted in China, compared an SGA with either full-body or scalp electroacupuncture.

FIGURE A. Comparison of Response Rates of SGAs Compared With Other Eligible Interventions (Relative Risks and 95% Confidence Intervals)

CAM = complementary and alternative medicine; CBT = cognitive behavioral therapy; NWMA = network meta-analysis; RCT = randomized controlled trial; SAMe = S-adenosyl-L-methionine; SGA = second-generation antidepressant; SOE = strength of evidence

For treatment response, pooled results from direct comparisons and network meta-analysis demonstrated no differences in benefits (low SOE). Two trials (237 participants) examined the effect of adding acupuncture to the SGA treatment regimen. Acupuncture in combination with an SGA had 37-percent higher response rates than SGAs alone (low SOE) but did not differ statistically in remission rates (low SOE).

Compared with SGA monotherapy, the combination of SGAs and acupuncture did not differ statistically in overall discontinuation rates (low SOE), overall rates of adverse events (low SOE), or discontinuation rates because of adverse events (low SOE).

The evidence was insufficient to conclude anything about differences in functional capacity, quality of life, or overall risk of harms. Evidence from meta-analyses of placebo-controlled trials, however, indicated lower overall adverse event rates for acupuncture than SGAs.

Second-Generation Antidepressants Versus Omega-3 Fatty Acids. One trial (40 participants) compared an SGA with omega-3 fatty acids. Network meta-analysis indicated a response rate that was twice as high for patients treated with SGAs as for those receiving omega-3 fatty acids (low SOE).

SGAs and omega-3 fatty acids did not lead to significantly different rates of overall discontinuation (low SOE) or discontinuation because of adverse events (insufficient SOE). Evidence was insufficient to draw conclusions about differences in remission, functional capacity, quality of life,

suicidal ideas or behaviors, overall risk of adverse events, or overall risk of serious adverse events.

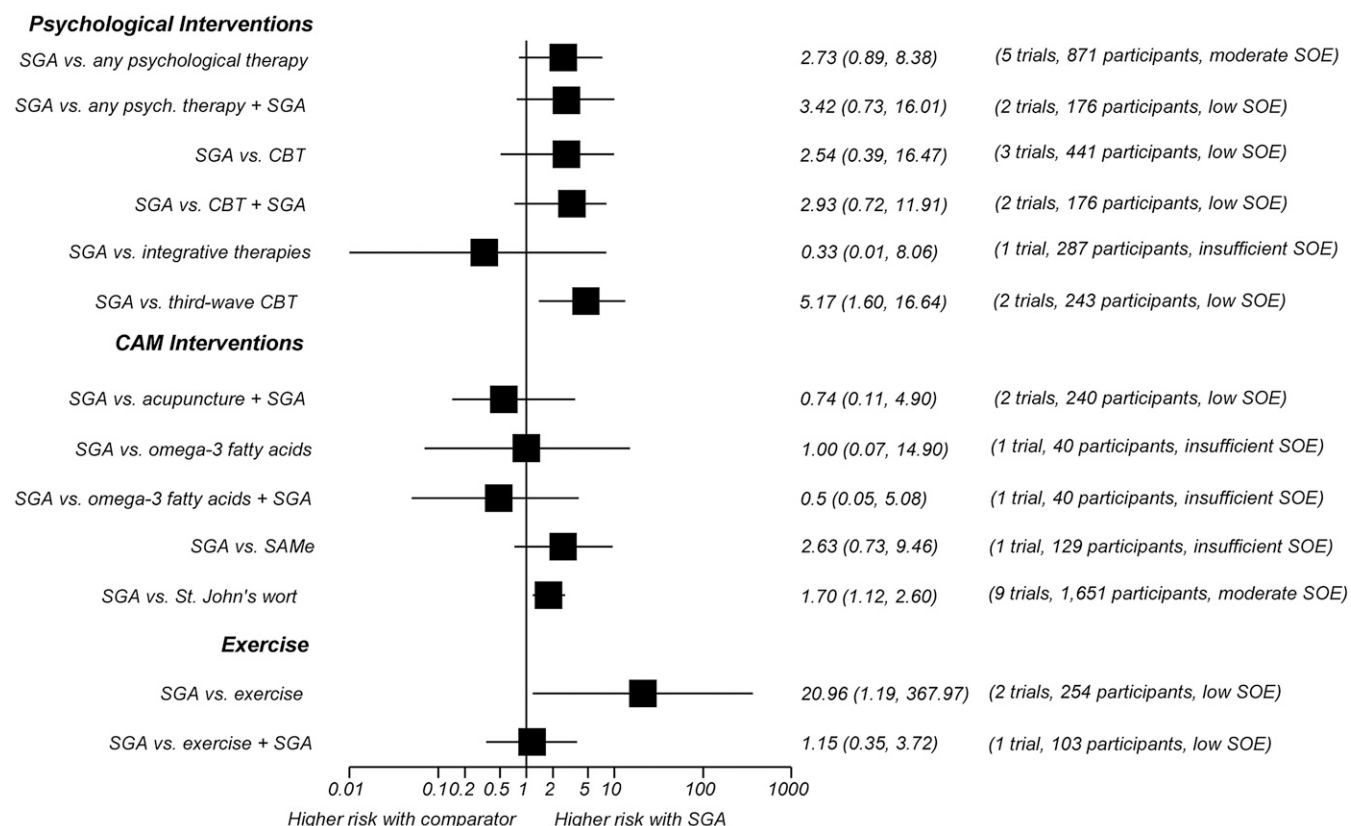
Two trials (72 participants) examined the effect of adding omega-3 fatty acids to the SGA regimen. Compared with SGA monotherapy, adding omega-3 fatty acids to the SGA regimen led to similar overall discontinuation rates (low SOE). Because of methodological shortcomings, the evidence was insufficient to draw any other conclusions.

Second-Generation Antidepressants Versus S-Adenosyl-L-Methionine. One trial (129 participants) compared an SGA with SAMe. Network meta-analysis indicated response rates that did not differ statistically for patients on SGAs or SAMe (low SOE).

Overall discontinuation rates were also similar between patients treated with SGAs or SAMe (low SOE).

The evidence was insufficient to draw conclusions about differences in remission, functional capacity, quality of life, discontinuation due to adverse events, or overall risk of adverse events.

Second-Generation Antidepressants Versus St. John's Wort. We identified 12 trials (1,806 participants) comparing SGAs with St. John's wort monotherapy. Meta-analysis of nine trials (1,513 participants) indicated similar response rates between SGAs and St. John's wort (low SOE). However, all trials compared St. John's wort with moderate- or low-dose SGA regimens, not fully using the approved range of SGA doses. Meta-analysis of five trials (768 participants) demonstrated similar remission rates for the two treatments (low SOE).

FIGURE B. Comparison of Rates of Discontinuation Because of Adverse Events From SGA With Other Eligible Interventions (Relative Risks and 95% Confidence Intervals)

CAM = complementary and alternative medicine; CBT = cognitive behavioral therapy; SAMe = S-adenosyl-L-methionine; SGA = second-generation antidepressant; SOE = strength of evidence

SGAs led to 28-percent higher rates of overall discontinuation (moderate SOE) and 70-percent higher rates of discontinuation because of adverse events (moderate SOE) as St. John's wort. The overall risk of adverse events was 17 percent higher among patients receiving SGAs than those receiving St. John's wort (moderate SOE). In contrast, the risk of serious adverse events did not differ significantly between patients receiving SGAs or St. John's wort (low SOE).

The evidence was insufficient to conclude anything about differences in functional capacity, quality of life, or suicidal ideas or behaviors.

Second-Generation Antidepressants Versus Yoga or Meditation. We identified no eligible trial that compared an SGA with yoga or meditation.

Severity as a Moderator of Comparative Treatment Effectiveness. One trial yielded insufficient evidence to determine whether the comparative effectiveness of SGAs versus SAMe changes as a function of MDD severity.

Second-Generation Antidepressants Compared With Exercise. Two trials (309 participants in active-treatment arms) compared an SGA with aerobic exercise. One trial also examined the effects of adding exercise to the SGA regimen. Rates of remission and discontinuation did not statistically differ for patients treated with SGAs and patients treated with exercise

monotherapy (low SOE). Estimates based on network meta-analysis indicated no significant difference in response for patients treated with SGAs and those treated with exercise (low SOE).

Although SGAs and exercise led to similar rates of overall discontinuation (low SOE), rates of discontinuation because of adverse events were 20 times as high for patients treated with SGAs as for those assigned to exercise (low SOE).

The combination treatment of SGAs and exercise led to remission, overall discontinuation rates, and rates of discontinuation because of adverse events that did not differ statistically from those among patients receiving SGA monotherapy (low SOE).

Second-Step Therapy: Effectiveness and Harms of Switching or Augmenting Treatment Options for Patients With Major Depressive Disorder

Switch: Second-Generation Antidepressant Versus Second-Generation Antidepressant. Results from two direct comparisons of second-step therapies involving 1,123 patients who were switched to different SGAs indicate no substantial differences in response rates between SGAs (moderate SOE). Results from one direct comparison involving 727 patients indicate no substantial difference in remission rates

or in the decrease in depressive severity between SGAs (low SOE).

Likewise, results from the same direct comparison of 727 patients indicate no significant difference in overall risk of adverse events (low SOE), rates of discontinuation because of adverse events (moderate SOE), overall risk of serious adverse events (low SOE), and suicidal ideas or behaviors (low SOE).

Switch: Second-Generation Antidepressant Versus Cognitive Therapy. Results from one direct comparison of second-step therapies involving 122 patients who were assigned to switch to a different SGA or to CT indicate no substantial differences in rates of response or remission or in the decrease in depressive severity (low SOE). In addition, rates of discontinuation because of adverse events (low SOE) were similar between SGAs and CT.

Switch: Second-Generation Antidepressant Versus Complementary and Alternative Medicine or Exercise. We did not find any eligible switch evidence comparing an SGA strategy with either CAM or exercise.

Augment: Second-Generation Antidepressant Versus Second-Generation Antidepressant. Results from one direct comparison of second-step therapies involving 565 patients indicate no substantial differences in rates of response or remission between SGAs (low SOE). However, results from one direct comparison involving 565 patients indicate a greater decrease in depressive severity after adding bupropion than buspirone (low SOE). In addition, adding bupropion led to lower rates of discontinuation because of adverse events (moderate SOE) but similar rates of serious adverse events (low SOE) and suicidal ideas or behaviors (low SOE) compared with adding buspirone.

Augment: Second-Generation Antidepressant Versus Cognitive Therapy. Results from one direct comparison of second-step therapies involving 182 patients whose treatment was augmented with a second medication versus augmented with CT indicate no substantial differences in rates of response or remission, or in the decrease in depressive severity (low SOE). The same results also indicate no significant differences in rates of discontinuation because of adverse events (low SOE) or overall risk of serious adverse events (low SOE).

Severity as a Moderator of Comparative Treatment Effectiveness of Second-Step Therapies. One industry-supported secondary analysis involving 396 patients found an insignificant trend toward differences in remission rates for those with severe depression (compared with moderate depression). In contrast, a second secondary analysis involving 727 patients, which was government funded, found that having mild or moderate rather than severe depression did not change the likelihood of remitting after treatment with one versus another SGA (insufficient evidence).

Comparative Benefits and Risks of Harms for Selected Subgroups

No trials were specifically designed to assess differences in our specified subgroups. Overall, only three trials addressing a subgroup of interest met the criteria for inclusion: one of subgroups defined by common accompanying psychiatric symptoms and two of subgroups defined by demographic characteristics. For common accompanying psychiatric symptoms, SGAs produced slightly higher remission rates than interpersonal psychotherapy in patients with a comorbid anxiety disorder but not in those without co-occurring anxiety (insufficient SOE). We had no evidence for any other common accompanying symptoms (insomnia, low energy, or somatization).

For subgroups defined by demographic characteristics, we included two trials. In one trial conducted in older adults, SGAs and St. John's wort led to similar response rates and discontinuation rates because of adverse events (low SOE). The other trial included only minority (predominantly black and Latina) women and showed similar reduction in depressive symptoms between SGAs and CBT (insufficient SOE). We did not identify any trials assessing differences between men and women in effectiveness or harms (insufficient SOE).

No trials at all addressed effectiveness or harms in selected subgroups of patients who did not achieve remission following an initial adequate trial with one SGA (insufficient SOE).

DISCUSSION

Key Findings and Strength of Evidence

Across all interventions, we graded the strength of evidence for benefits as moderate for only one comparison—namely, SGAs compared with CBT. Results from trials of this comparison indicate that SGAs and CBT have similar effectiveness regarding symptomatic relief in patients with mild to severe MDD. For risk of harms, we graded the strength of evidence as moderate for some outcomes of three comparisons—namely, SGAs compared with CBT, acupuncture, and St. John's wort. Patients treated with SGAs had a higher risk of experiencing adverse events or discontinuing treatment because of adverse events than patients treated with CBT, acupuncture, or St. John's wort. The evidence is insufficient to draw conclusions about differences in serious adverse events, such as suicidal ideas and behavior.

Our confidence in findings from the comparisons of remaining treatment options was low or insufficient, indicating that these bodies of evidence had major or unacceptable deficiencies. Nevertheless, for most comparisons the overall findings did not show statistically significant differences in benefits but indicated a lower risk of adverse events for nonpharmacological treatment options. Notable exceptions are omega-3-fatty acids, which appear to have lower effectiveness than SGAs; the combination of SGAs with acupuncture, which appears to have higher response rates than SGA

monotherapy; and the combination of SGAs with interpersonal psychotherapy, which appears to have better effectiveness than SGA monotherapy. Our confidence in these findings, however, is low, and results have to be interpreted cautiously. In addition, for many comparisons that are limited to single trials, determining whether similar treatment effects between SGAs and other interventions are based on similar effectiveness or high placebo response rates is impossible. Furthermore, we emphasize that detecting no statistically significant difference does not necessarily mean the treatments are equivalent.

The available data offer no conclusions on how selection of treatment strategies might differ based on a patient's severity of depression. Overall, data do not indicate differences in comparative effectiveness between SGAs and nonpharmacological interventions for patients with severe MDD. This important question concerning MDD severity, although raised by a few systematic reviews,¹⁷⁻¹⁹ remains without a clear answer.

Beyond the two articles identified comparing switching and augmentation strategies employing a limited number of medication options or CT, the absence of relevant comparative data about which treatment options are most effective for those needing second-step treatment (about 70% of patients with MDD)^{20,21} was striking.

Our findings are consistent with several prior systematic reviews and meta-analyses that compared SGAs with nonpharmacological interventions. Most of these reviews, however, included populations that were not eligible for our review, such as patients with minor depression, bipolar disorder, or dysthymia.

Our results are partially consistent with the recommendations of both the American Psychiatric Association²² and the Department of Veterans Affairs/Department of Defense.²³ These consider both pharmacotherapy and psychotherapy to be appropriate individual first-step treatments for patients with mild to moderate MDD, and state that the combination of pharmacotherapy and psychotherapy may be necessary in cases of moderate to severe depression.

In terms of clinical decisionmaking, the information in this review can be helpful to physicians because they can provide a summary of the available evidence base indicating the advantages and disadvantages of these options, and patients can identify which intervention they would prefer. Some options, such as medication and St. John's wort, would require physician supervision and monitoring, given potential side effects and drug interactions. Moreover, patients who would like to maintain or start an exercise regimen in addition to undergoing SGA therapy can be encouraged to do so. The enhanced potential for increasing physical well-being and expanding social interactions may be an added incentive to encourage an exercise regimen.

Applicability

The scope of this review was limited to trials that enrolled adult patients with MDD. We did not attempt to review literature on interventions for children with MDD or for

patients with subthreshold depression (depressive symptoms not severe enough to meet diagnostic criteria for a major depressive episode), dysthymia, psychotic depression, or perinatal depression. The included trials covered populations with mild, moderate, and severe MDD; the majority of participants were women. Most trial populations, however, excluded patients with medical comorbidities or suicidal ideas and behaviors; few trials included elderly patients. We did not find evidence to confirm or refute whether treatments are more or less efficacious for various subgroups (i.e., patients characterized by sex, race, or ethnicity, or individuals with coexisting psychiatric conditions).

With few exceptions, interventions in included trials were in line with clinical practice. Except for some CAM trials in which patients received SGA dosages at the lower end of the recommended range, prescribing patterns and doses in the SGA arms of our evidence base were consistent with clinical practice. Some newer SGAs, such as desvenlafaxine, levomilnacipran, vilazodone, or vortioxetine, have never been compared with psychological or CAM treatments or exercise. Nevertheless, reliable evidence indicates that the comparative effectiveness of SGAs is similar.²⁴ Consequently, we believe that our findings are applicable across the class of SGAs.

As noted previously, detecting no statistically significant difference does not necessarily mean that the treatments are equivalent. The studies involved were designed to test whether an outcome for one intervention was different from the outcome for another rather than to test equivalence, which would generally require a much larger sample size. This point is especially relevant for findings with a low SOE. While confidence intervals were relatively narrow and risk ratios were often close to 1 (findings consistent with equivalent outcomes), a conclusion of equivalence cannot be made. Further, while moderate-strength evidence at a group level did not detect a difference between SGAs and CBT, how best to tailor this information to an individual patient is still not clear. Indeed, other potentially relevant indicators (e.g., depressive severity, comorbid psychiatric illness) may favor one over another, but the current evidence base (as indicated in the KQ 1b and 2b findings) is quite limited.

Finally, many trials, particularly for CAM interventions, were conducted outside the United States. Whether and how differences in ethnic or cultural backgrounds and health systems affect the applicability of results to U.S. populations remain uninvestigated and unanswered.

Research Gaps

Across all comparisons of interventions, major research gaps pertain to information about patient-centered outcomes, such as functional capacity and quality of life, and the comparative risk of harms. Lack of information about harms can lead to a biased knowledge base and the potential for decisions that cause more harm than good.

We found no eligible studies that compared SGAs with behavior therapy or behavior modification, humanistic therapies, yoga, or mindfulness interventions. Given the wide use

of these types of psychotherapies in clinical practice, further research into their comparative effectiveness with SGAs in treating MDD patients is desirable. For many psychotherapies and all CAM therapies that have been evaluated against an SGA, the data were insufficient because trials did not report important outcomes, most notably quality of life and functional capacity. Future studies should assess remission, response to treatment, quality of life, functional capacity, suicidal ideas and behaviors, and adverse events using standardized measures to allow for more direct comparisons across studies using the same or similar SGAs and psychological interventions. These same deficiencies in the literature extend to the comparative effectiveness of SGAs and both psychological and CAM interventions for treating MDD as a function of depression severity. For CAM interventions, we found that most studies did not include the full range of SGA doses for comparison, and many studies made comparisons with only the very lowest SGA doses. To truly compare any CAM intervention for MDD treatment, future studies will need to incorporate SGA dosing strategies that use the entire SGA dosage range. Finally, a major gap in the evidence is the lack of studies addressing different treatment options for patients who have not achieved remission with first-step therapy. No second-step therapy data at all exist that compare SGA with CAM or exercise treatments. This void in the evidence base is a major one that will perplex and confound clinicians, patients, policymakers, and guideline developers alike.

Conclusions

Overall, the available evidence indicates that SGAs and CBT do not differ significantly in symptomatic relief as first-step treatments for adult outpatients with mild to severe MDD. The evidence is insufficient to draw conclusions about the comparative risk of serious adverse events, such as suicidal ideas and behaviors. Given comparable benefits among treatment options, the choice of the initial treatment of MDD should consider results of previous treatments, patient preferences, and feasibility (e.g., costs, likely adherence, and availability) following a discussion of the advantages and disadvantages of each treatment option, including risks of particular adverse effects and potential drug interactions.

Differences with respect to adverse events, personal engagement, and costs may be taken into consideration for the choice of a first-step treatment. Such shared and informed decisionmaking might enhance treatment adherence and improve treatment outcomes for patients with MDD, especially because treatment continuity is one of the main challenges in treating such patients. For second-step therapies, although evidence is limited, no clear benefit emerges to suggest that either switching to a particular SGA or CT, or augmenting with a particular medication or CT, is preferable. Available data suggest that switching to another SGA, switching to CT, or augmenting with a particular medication or CT are all reasonable options. The more important decision appears to be simply to try a different evidence-based approach.

Addendum

In the manuscript summarizing the findings of this report for journal submission, we employed a different statistical approach for random effects meta-analyses than in the AHRQ report. We followed journal policy and used restricted maximum likelihood models instead of DerSimonian and Laird methods. As a consequence, point estimates and the width of some confidence intervals for some effect estimates are slightly different between the AHRQ report and journal manuscript. Differences are minor and do not change conclusions.

The full report is available at <https://www.effectivehealthcare.ahrq.gov/ehc/products/568/2155/major-depressive-disorder-report-151202.pdf>.

Focus 2016; 14:283–293; doi: 10.1176/appi.focus.140203

REFERENCES

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th ed. Arlington, VA: American Psychiatric Publishing; 2013.
2. Kessler RC, Berglund P, Demler O, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA*. 2003 Jun 18;289(23):3095–105. PMID: 12813115.
3. Khan A, Sambunaris A, Edwards J, et al. Vilazodone in the treatment of major depressive disorder: efficacy across symptoms and severity of depression. *Int Clin Psychopharmacol*. 2014 Mar; 29(2):86–92. PMID: 24247740.
4. Nemeroff CB. The burden of severe depression: a review of diagnostic challenges and treatment alternatives. *J Psychiatr Res*. 2007 Apr-Jun;41(3–4):189–206. PMID: 16870212.
5. Thase ME. Treatment of severe depression. *J Clin Psychiatry*. 2000;61 Suppl 1:17–25. PMID: 10703759.
6. Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: The National Academies Press; 2001. www.nap.edu/openbook.php?record_id=10027.
7. Mojtabai R, Olfson M. National patterns in antidepressant treatment by psychiatrists and general medical providers: results from the National Comorbidity Survey Replication. *J Clin Psychiatry*. 2008 Jul;69(7):1064–74. PMID: 18399725.
8. Gartlehner G, Hansen RA, Morgan LC, et al. Second-Generation Antidepressants in the Pharmacologic Treatment of Adult Depression: An Update of the 2007 Comparative Effectiveness Review. (Prepared by the RTI International–University of North Carolina Evidence-based Practice Center, under 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. PMID: 22299185.
9. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; January 2014. Chapters available at www.effectivehealthcare.ahrq.gov.
10. Viswanathan M, Ansari MT, Berkman ND, et al. Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions; 2012. In: *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2014. Chapters available at www.effectivehealthcare.ahrq.gov.
11. Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928. PMID: 22008217.
12. Hong H, Carlin BP, Shamlilyan TA, et al. Comparing Bayesian and frequentist approaches for multiple outcome mixed treatment comparisons. *Med Decis Making*. 2013 Jul;33(5):702–14. PMID: 23549384.

13. Jones B, Roger J, Lane PW, et al. Statistical approaches for conducting network meta-analysis in drug development. *Pharm Stat*. 2011 Nov-Dec;10(6):523–31. PMID: 22213533.
14. Berkman ND, Lohr KN, Ansari M, et al. Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update, 2013. In: *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. AHRQ Publication No. 10(14)-EHC063-EF. Rockville MD: Agency for Healthcare Research and Quality; 2014. Chapters available at www.effectivehealthcare.ahrq.gov.
15. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol*. 2011 Apr;64(4):395–400. PMID: 21194891.
16. Atkins D, Chang S, Gartlehner G, et al. Assessing the Applicability of Studies When Comparing Medical Interventions; 2010. In: *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality; 2014. Chapters available at www.effectivehealthcare.ahrq.gov.
17. Fournier JC, DeRubeis RJ, Hollon SD, et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA*. 2010 Jan 6;303(1):47–53. PMID: 20051569.
18. Kirsch I, Deacon BJ, Huedo-Medina TB, et al. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med*. 2008 Feb;5(2):e45. PMID: 18303940.
19. Khan A, Leventhal RM, Khan SR, et al. Severity of depression and response to antidepressants and placebo: an analysis of the Food and Drug Administration database. *J Clin Psychopharmacol*. 2002 Feb;22(1):40–5. PMID: 11799341.
20. Gaynes BN, Rush AJ, Trivedi MH, et al. Primary versus specialty care outcomes for depressed outpatients managed with measurement-based care: results from STAR*D. *J Gen Intern Med*. 2008 May;23(5):551–60. PMID: 18247097.
21. Gaynes BN, Lux LJ, Lloyd SW, et al. Nonpharmacologic Interventions for Treatment-Resistant Depression in Adults. Comparative Effectiveness Review No. 33. (Prepared by RTI International-University of North Carolina Evidence-based Practice Center under Contract No. 290-02-00161.). AHRQ Publication No. 11-EHC056-EF. Rockville, MD: Agency for Healthcare Research and Quality; September 2011. www.effectivehealthcare.ahrq.gov/reports/final.cfm. PMID: 22091472.
22. American Psychiatric Association. Practice Guideline for the Treatment of Patients with Major Depressive Disorder. 3rd ed. Arlington, VA; Oct 2010. http://psychiatryonline.org/pb/assets/raw/sitewide/practice_guidelines/guidelines/mdd.pdf.
23. Department of Veteran Affairs, Department of Defense. VA/DoD Clinical Practice Guideline for Management of Major Depressive Disorder (MDD). May 2009. www.healthquality.va.gov/.
24. Gartlehner G, Hansen RA, Morgan LC, et al. Comparative benefits and harms of second-generation antidepressants for treating major depressive disorder: an updated meta-analysis. *Ann Intern Med*. 2011 Dec 6;155(11):772–85. PMID: 22147715.