

# Efficacy and Safety of Second-Generation Antidepressants in the Treatment of Major Depressive Disorder

Richard A. Hansen, PhD; Gerald Gartlehner, MD, MPH; Kathleen N. Lohr, PhD; Bradley N. Gaynes, MD, MPH; and Timothy S. Carey, MD, MPH

**Background:** Reviews have compared the efficacy and tolerability of newer second-generation antidepressants with those of placebo or older treatments, but comparative evidence for use of second-generation antidepressants to treat major depressive disorder has not been evaluated.

**Purpose:** To systematically evaluate comparative data on the efficacy, effectiveness, and tolerability of commonly prescribed second-generation antidepressants (selective serotonin reuptake inhibitors, bupropion, duloxetine, mirtazapine, and venlafaxine) in the treatment of major depressive disorder.

**Data Sources:** MEDLINE, EMBASE, and PsychLit; the Cochrane Library; and the International Pharmaceutical Abstracts were searched from January 1980 through February 2005 for reviews; randomized, controlled trials; meta-analyses; and observational studies.

**Study Selection:** The authors reviewed 46 head-to-head randomized, controlled trials comparing one second-generation antidepressant with another. Twenty-four observational studies and placebo-controlled trials were also included for assessment of safety and tolerability.

**Data Extraction:** Two researchers independently reviewed titles and abstracts. Trained reviewers abstracted data from each study and assigned an initial quality rating. A second reviewer verified the data extraction and quality rating.

**Data Synthesis:** According to fair to good evidence, the second-generation antidepressants that were compared had only minimal differences in efficacy, and 88% of comparative efficacy studies reported no statistically significant difference in any outcome measure at the end of the study. One effectiveness trial rated good and 2 effectiveness trials rated fair reported no statistically significant differences in primary outcome measures for compared drugs. Meta-analyses showed a modest but statistically significant additional treatment effect for sertraline and venlafaxine compared with fluoxetine. About 96% of comparative trials were sponsored by or had at least 1 author affiliated with a pharmaceutical company; the remaining trials did not report funding sources. Adverse event profiles differed among drugs; however, the degree and quality of adverse event assessment varied and only 13% of trials used a standardized scale to assess adverse events.

**Limitations:** Quantitative analyses could not be done for many drug comparisons because the quantity and quality of the evidence were inadequate. Most published evidence was from trials sponsored by pharmaceutical companies, and publication bias may have occurred.

**Conclusions:** Overall, second-generation antidepressants probably do not differ substantially for treatment of major depressive disorder. Choosing the agent that is most appropriate for a given patient is difficult.

*Ann Intern Med.* 2005;143:415-426.

For author affiliations, see end of text.

www.annals.org

**M**ajor depressive disorder is a serious disabling illness that affects more than 16% of adults in the United States at some point in their lifetime (1). In 2000, the economic burden of depressive disorders in the United States was estimated to be \$83.1 billion (2). Current practice guidelines for the treatment of major depressive disorder recommend pharmacotherapy, psychotherapy, psychotherapy plus pharmacotherapy, or electroconvulsive therapy. In most cases, pharmacotherapy is first-line treatment for major depressive disorder. Moreover, it is a practical tool for primary care physicians, who prescribe the majority of antidepressants in the United States (3).

Pharmacologic treatment for major depressive disorder includes first-generation antidepressants (tricyclic antidepressants and monoamine oxidase inhibitors) and second-generation antidepressants. Second-generation medications include selective serotonin reuptake inhibitors (SSRIs); selective norepinephrine reuptake inhibitors; and other drugs that selectively affect the activity of neurotransmitters, such as serotonin, norepinephrine, and dopamine. In general, the efficacy of first- and second-generation antidepressant medications is similar (4–6). However, first-generation antidepres-

sants often cause multiple side effects that many patients find intolerable (7–9), and the risk for harm when taken in overdose or in combination with certain medications is high. Because of their relatively favorable side effect profile, the second-generation antidepressants play a prominent role in the management of patients with major depressive disorder.

Reviews have compared the efficacy and tolerability of newer second-generation antidepressants with those of placebo or older treatments (6, 10, 11) but did not evaluate comparative evidence for second-generation antidepres-

See also:

## Print

Editors' Notes . . . . . 416

## Web-Only

Appendix Table

Appendix Figure

CME quiz

Conversion of figures and tables into slides

**Context**

The relative benefits and harms of newer, second-generation antidepressants are sometimes confusing.

**Contribution**

In this review of 46 head-to-head randomized trials, the authors generally found no major differences in the numbers of adults with major depression who responded to second-generation antidepressants, such as selective serotonin reuptake inhibitors, bupropion, duloxetine, mirtazapine, and venlafaxine. The overall incidence of adverse events appeared similar across drugs, although types of adverse events varied.

**Cautions**

Trials were funded by industry and had variable quality and follow-up duration.

**Implications**

Second-generation antidepressants generally have similar benefits but different possible harms for adults with major depression.

—The Editors

sants. We therefore sought to systematically evaluate comparative data on the efficacy, effectiveness, and tolerability of commonly prescribed second-generation antidepressants. Specifically, we conducted a systematic review and meta-analysis of comparative evidence for 6 SSRIs (citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, and sertraline) and 4 other second-generation antidepressants (bupropion, duloxetine, mirtazapine, and venlafaxine). From here on, we refer to these agents collectively as *antidepressants*. We examine the role of these agents in the initial treatment of ambulatory adult patients with major depressive disorder.

**METHODS****Key Questions**

Key questions designed to address the comparative efficacy, effectiveness, safety, and tolerability of antidepressants guided our research. A consortium of 12 state Medicaid programs, the Canadian Coordinating Office for Health Technology Assessment, the California HealthCare Foundation, and key experts formulated the questions and provided funding for this research.

**Literature Search**

To identify articles relevant to each key question, we searched MEDLINE, EMBASE, and PsychLit; the Cochrane Library; and the International Pharmaceutical Abstracts. To capture articles relevant to the scope of our topic, our searches covered 1980 through 28 February

2005. We manually searched reference lists of relevant review articles and letters to the editor. Pharmaceutical manufacturers were invited to submit dossiers, including citations, as outlined by the Drug Effectiveness Review Project (12). We requested unpublished studies from the U.S. Food and Drug Administration, but this agency did not release unpublished data.

**Study Selection**

Two persons independently reviewed titles and abstracts. If both reviewers agreed that a trial did not meet preestablished eligibility criteria (**Appendix Table**, available at [www.annals.org](http://www.annals.org)), we excluded it. To assess efficacy and effectiveness, we included head-to-head trials comparing one antidepressant with another. We defined effectiveness trials as those that were conducted in primary care settings, had an adequate duration of follow-up ( $\geq 3$  months), had minimal inclusion and exclusion criteria (so that participants represented the general population), assessed health outcomes rather than intermediate outcomes, and had an adequate sample size to determine a minimally important difference (from a patient's perspective) on a health-related quality of life instrument (13).

To assess safety and tolerability, we included head-to-head trials, placebo-controlled trials, and observational studies with large samples ( $> 100$  patients) lasting at least 1 year. We required a larger sample size for observational studies because we wanted primarily to detect adverse events that were not frequent enough to be apparent in smaller trials.

**Data Abstraction and Quality Assessment**

Trained reviewers abstracted data from each study, and a senior reviewer read each abstracted article and evaluated completeness of data extraction. We recorded intention-to-treat results if they were available. We assessed the internal validity (quality) of trials by using predefined criteria from the U.S. Preventive Services Task Force (ratings of good, fair, or poor) (14) and the National Health Service Centre for Reviews and Dissemination (15). Elements of internal validity assessment included randomization, allocation concealment, similarity of compared groups at baseline, use of intention-to-treat analysis, and overall and differential loss to follow-up. We defined loss to follow-up as the number of persons who underwent randomization but did not complete the study (16), independent of the reason and whether intention-to-treat analysis was used. We rated studies as poor if they had more than 40% overall loss to follow-up or more than 15 percentage points of differential loss to follow-up between study groups.

**Data Synthesis**

We first qualitatively summarized the studies. When more than 3 head-to-head trials compared the same treatments, we did quantitative analyses. In these, the primary outcome measure was treatment response, defined as 50%

**Table 1. Characteristics of Efficacy or Effectiveness Trials Comparing One Selective Serotonin Reuptake Inhibitor with Another\***

Author, Year (Reference)	Patients, n	Mean Age, y	Women, %	Duration, wk	Drug Comparison	Main Results†	Quality Rating
Burke et al., 2002 (25)	491	40	64	8	Citalopram‡, 40 mg/d, vs. escitalopram‡, 10 mg/d, and 20 mg/d	No differences	Fair
Lepola et al., 2003 (26)	471	43	72	8	Citalopram‡, 20–40 mg/d, vs. escitalopram‡§, 10–20 mg/d	Significant differences	Fair
Patris et al., 1996 (24)	357	43	77	8	Citalopram‡, 20 mg/d, vs. fluoxetine, 20 mg/d	No differences	Fair
Ekselius et al., 1997 (21)	400	47	72	24	Citalopram, 20–60 mg/d, vs. sertraline‡, 50–150 mg/d	No differences	Good
Dalery and Honig, 2003 (27)	184	42	63	6	Fluoxetine, 20 mg/d, vs. fluvoxamine‡, 100 mg/d	No differences	Fair
Rapaport et al., 1996 (41)	100	39	62	7	Fluoxetine, 20–80 mg/d, vs. fluvoxamine‡, 100–150 mg/d	No differences	Fair
Cassano et al., 2002 (28)¶	242	75	56	52	Fluoxetine, 20–60 mg/d, vs. paroxetine‡, 20–40 mg/d	No differences	Fair
Chouinard et al., 1999 (29)**	203	41	62	12	Fluoxetine, 20–80 mg/d, vs. paroxetine‡, 20–50 mg/d	No differences	Fair
De Wilde et al., 1993 (30)**	100	44	48	6	Fluoxetine, 40–60 mg/d, vs. paroxetine‡, 30–40 mg/d	No differences	Fair
Schone and Ludwig, 1993 (31)**	108	74	87	6	Fluoxetine, 20–60 mg/d, vs. paroxetine‡§¶, 20–40 mg/d	Significant differences	Fair
Fava et al., 1998 (32)**	128	41	51	12	Fluoxetine, 20–80 mg/d, vs. paroxetine‡, 20–50 mg/d	No differences	Fair
Fava et al., 2002 (36)**††	284	43	60	10–16	Fluoxetine‡, 20–60 mg/d, vs. paroxetine, 20–60 mg/d	No differences	Fair
Gagliano, 1993 (40)**	90	38	80	6	Fluoxetine, 20–60 mg/d, vs. paroxetine, 20–40 mg/d	No differences	Fair
Kroenke et al., 2001 (23)  ¶††	601	46	79	39	Fluoxetine‡, 23 mg/d‡‡, vs. paroxetine, 24 mg/d‡‡	No differences	Fair
Bennie et al., 1995 (33)**	286	50	61	6	Fluoxetine, 20–40 mg/d, vs. sertraline‡, 50–100 mg/d	No differences	Fair
Boyer et al., 1998 (34)**	242	43	78	26	Fluoxetine, 20–60 mg/d, vs. sertraline‡, 50–150 mg/d	No differences	Fair
Sechter et al., 1999 (22)  **	238	43	67	24	Fluoxetine, 20–60 mg/d, vs. sertraline‡, 50–150 mg/d	No differences	Fair
Newhouse et al., 2000 (35)**	236	68	57	12	Fluoxetine, 20–40 mg/d, vs. sertraline‡, 50–100 mg/d	No differences	Fair
Fava et al., 2002 (36)**††	284	43	60	10–16	Fluoxetine‡, 20–40 mg/d, vs. sertraline, 50–200 mg/d	No differences	Fair
Kroenke et al., 2001 (23)  ¶††	601	46	79	39	Fluoxetine‡, 23 mg/d‡‡, vs. sertraline, 73 mg/d‡‡	No differences	Fair
Kiev and Feiger, 1997 (42)	60	41	53	7	Fluvoxamine‡, 50–150 mg/d, vs. paroxetine, 20–50 mg/d	No differences	Fair
Nemeroff et al., 1995 (38)	95	40	61	7	Fluvoxamine‡, 50–150 mg/d, vs. sertraline, 50–200 mg/d	No differences	Fair
Franchini et al., 1997 (39)	64	48	77	104	Fluvoxamine, 200–300 mg/d, vs. sertraline, 100–200 mg/d	No differences	Fair
Aberg-Wistedt et al., 2000 (37)	353	NR	NR	24	Paroxetine, 20–40 mg/d, vs. sertraline‡, 50–150 mg/d	No differences	Fair
Fava et al., 2002 (36)††	284	43	60	10–16	Paroxetine, 20–60 mg/d, vs. sertraline, 50–200 mg/d	No differences	Fair
Kroenke et al., 2001 (23)  ¶††	601	46	79	39	Paroxetine, 24 mg/d‡‡, vs. sertraline, 73 mg/d‡‡	No differences	Fair

\* NR = not reported.  
 † Statistically significant differences in any measure of efficacy. The assessment of statistical significance is influenced by sample size, population characteristics, and study design.  
 ‡ The trial was sponsored by, or at least 1 author was affiliated with, the manufacturer of the drug.  
 § Statistically significant differences favor this drug.  
 || Effectiveness trial.  
 ¶ Although a meta-analysis was conducted, this trial was excluded.  
 \*\* This trial was included in the meta-analysis.  
 †† This trial had multiple arms; we report each comparison separately.  
 ‡‡ Mean dosage.

or greater improvement on the Hamilton Rating Scale for Depression (HAM-D) or the Montgomery–Asberg Depression Rating Scale from baseline to study end. The relative benefit reflects the ratio of benefits or risks in one

treatment group compared with another. When treatment effects differed between studies, we explored potential reasons for these differences. For each meta-analysis, we tested for heterogeneity of treatment effects by using  $I^2$  statistics.

If no heterogeneity was detected, we applied both a random-effects and a fixed-effects model. We report the results of the more conservative random-effects models (17) because the validity of tests of heterogeneity can be limited with a small number of component studies. To estimate possible publication bias caused by the tendency of published studies to be positive, we used funnel plots, the Begg adjusted rank correlation test (18), and the Egger regression approach (19). However, because these tests have low statistical power when the number of trials is small (20), undetected bias may still be present. All statistical analyses were conducted by using StatsDirect Statistical Software, version 2.3.8 (StatsDirect, Ltd., Sale, United Kingdom).

We calculated the mean incidence and 95% CIs for specific adverse events reported in included trials. Because assessment and reporting of adverse events varied greatly among trials, this evidence should be interpreted with caution.

### Role of the Funding Sources

The funding sources contributed to the development of the key questions but had no role in the conduct or reporting of the study or in the decision to submit the manuscript for publication.

## RESULTS

We found 820 unduplicated citations. Manual review of the reference lists of pertinent review articles produced another 74 articles. Although 6 pharmaceutical companies submitted dossiers, no included studies stemmed from the dossiers. Therefore, 894 citations were included in our database (**Appendix Figure**, available at [www.annals.org](http://www.annals.org)).

Forty-six randomized, controlled trials compared the effectiveness or efficacy of one SSRI or other antidepressant with that of another in the treatment of major depressive disorder (**Tables 1 and 2**). (Complete evidence tables are available at [www.ohsu.edu/drugeffectiveness/reports/final.cfm](http://www.ohsu.edu/drugeffectiveness/reports/final.cfm) or from the authors.) Most studies were efficacy trials and received a rating of fair for internal validity; some trials rated fair may have fulfilled all quality criteria but did not report methods to an extent that answered all of our questions. We considered 2 trials from Europe (21, 22) and 1 trial from the United States (23) conducted in primary care settings to be effectiveness trials. Sixty percent of included trials were less than 12 weeks in duration. The samples consisted mostly of persons younger than 60 years of age; samples consisted of persons 60 years of age or older in 6 trials (13%) and children or adolescents younger than 18 years of age in 3 trials (7%). Although sponsorship did not influence our quality rating, it may have influenced reporting. About 85% of trials in our review were sponsored by a pharmaceutical company, and an additional 11% had at least 1 author affiliated with a pharmaceutical company. The remaining 4% of included studies did not report funding sources.

We excluded 118 full-text articles that did not meet our eligibility criteria (**Appendix Figure**, available at [www.annals.org](http://www.annals.org)). The majority of studies were excluded because of study design. We excluded 10 studies because of poor quality. Four of the poor-quality studies were comparative trials that were excluded because of high overall or differential loss to follow-up (68–71); the remaining 6 were pooled analyses that lacked a systematic literature search or excluded important trials (72–77).

### Comparisons of SSRIs with SSRIs

#### Efficacy or Effectiveness

Twenty-one studies rated fair (22–42) and 1 study rated good (21) compared one SSRI with another SSRI (**Table 1**). Of these studies, we classified 3 (14%) as effectiveness trials (21–23).

Overall, these trials reported similar outcomes among the 6 SSRIs; moreover, 20 trials reported no statistically significant differences in any efficacy measure at study end (21–25, 27–30, 32–42). Two trials rated fair reported statistically significant differences between SSRIs on at least 1 outcome measure (26, 31). Of these, one 8-week trial (26) comparing escitalopram with citalopram reported more responders in the escitalopram group ( $\geq 50\%$  improvement in total score on the Montgomery–Asberg Depression Rating Scale; 64% vs. 53% [ $P = 0.021$ ]). A similarly designed trial did not find statistically significant differences between escitalopram and citalopram (25). Of the 8 trials that compared paroxetine with fluoxetine, 7 reported no differences on any efficacy measure (23, 28–30, 32, 36, 40). One study (31) reported statistically significantly more responders in the paroxetine group ( $\geq 50\%$  improvement in HAM-D total score; 38% vs. 17% [ $P < 0.05$ ]); it was conducted in an older (mean age, 74 years), primarily female (87%) sample, was relatively short (6 weeks), and was small (108 patients). A larger study of patients 65 years of age or older reported no differences between treatment groups at 1 year (28).

The most likely explanation for the fact that some trials reported statistically significant differences in certain efficacy measures but similarly designed trials did not is the play of chance in these relatively small studies. Furthermore, most trials were sponsored by pharmaceutical companies, and statistically significant differences frequently favored the sponsor's antidepressant. On average, response rates were 5 percentage points (95% CI, 2 to 9 percentage points) higher for sponsored drugs. No effectiveness trial reported statistically significant differences between antidepressants. We did not identify systematic differences between findings of trials that were related to their quality rating, sample size, noninferiority design, or duration.

#### Speed of Response

Most trials reported no difference among SSRIs in speed of patient response. Evidence of a faster response to citalopram (24), fluvoxamine (27), and paroxetine (28, 30,

**Table 2. Characteristics of Efficacy or Effectiveness Trials Comparing a Selective Serotonin Reuptake Inhibitor with Another Second-Generation Antidepressant**

Author, Year (Reference)	Patients, n	Mean Age, y	Women, %	Duration, wk	Drug Comparison*	Main Results†	Quality Rating
Bielski et al., 2004 (62)	198	37	58	8	Escitalopram‡, 20 mg/d, vs. venlafaxine XR, 225 mg/d	No differences	Fair
Montgomery et al., 2004 (63)	293	48	71	8	Escitalopram‡, 10–20 mg/d, vs. venlafaxine XR, 75–150 mg/d	No differences	Fair
Feighner et al., 1991 (43)	123	42	62	7	Fluoxetine, 20–80 mg/d, vs. bupropion‡, 225–450 mg/d	No differences	Fair
Coleman et al., 2001 (44)	456	37	63	8	Fluoxetine, 20–60 mg/d, vs. bupropion SR‡, 150–400 mg/d	No differences	Fair
Goldstein et al., 2002 (60)	173	41	64	8	Fluoxetine‡, 20 mg/d, vs. duloxetine‡, 120 mg/d	No differences	Fair
Hong et al., 2003 (45)	132	47	63	6	Fluoxetine, 20–40 mg/d, vs. mirtazapine‡, 15–45 mg/d	No differences	Fair
Costa e Silva, 1998 (64)§	382	40	79	8	Fluoxetine, 20–40 mg/d, vs. venlafaxine‡, 75–150 mg/d	No differences	Good
Alves et al., 1999 (46)	87	44	92	12	Fluoxetine, 20–40 mg/d, vs. venlafaxine‡§, 75–150 mg/d	No differences	Fair
Tylee et al., 1997 (47)	341	45	71	12	Fluoxetine, 20 mg/d, vs. venlafaxine‡¶, 75–150 mg/d	No differences	Fair
Dierick et al., 1996 (48)	314	43	65	8	Fluoxetine, 20 mg/d, vs. venlafaxine‡, 75–150 mg	Significant differences	Fair
De Nayer et al., 2002 (49)	146	43	68	12	Fluoxetine, 20–40 mg/d, vs. venlafaxine‡¶, 75–150 mg/d	Significant differences	Fair
Rudolph and Feiger, 1999 (50)	301	40	68	8	Fluoxetine, 20–60 mg/d, vs. venlafaxine XR‡, 75–225 mg/d	No differences	Fair
Silverstone and Ravindran, 1999 (51)	368	42	59	12	Fluoxetine, 20–60 mg/d, vs. venlafaxine XR‡¶, 75–225 mg/d	No differences	Fair
Weihls et al., 2000 (65); Doraiswamy et al., 2001 (67)	100	71	57	6	Paroxetine‡, 20–40 mg/d, vs. bupropion SR‡, 100–300 mg/d	No differences	Good
Detke et al., 2004 (61)	367	43	73	8	Paroxetine, 20 mg/d, vs. duloxetine‡, 80 mg/d, and 120 mg/d	No differences	Fair
Schatzberg et al., 2002 (52)	254	72	50	8	Paroxetine, 20–40 mg/d, vs. mirtazapine‡, 15–45 mg/d	No differences	Fair
Benkert et al., 2000 (53)	275	47	63	6	Paroxetine, 20–40 mg/d, vs. mirtazapine‡, 15–45 mg/d	No differences	Fair
Ballus et al., 2000 (54)	84	45	88	24	Paroxetine, 20–40 mg/d, vs. venlafaxine‡, 75–150 mg/d	No differences	Fair
McPartlin et al., 1998 (55)	361	45	68	12	Paroxetine, 20 mg/d, vs. venlafaxine XR‡, 75 mg/d	No differences	Fair
Coleman et al., 1999 (56)	364	38	56	8	Sertraline, 50–200 mg/d, vs. bupropion SR‡, 150–400 mg/d	No differences	Fair
Croft et al., 1999 (57)	360	36	50	8	Sertraline, 50–200 mg/d, vs. bupropion SR‡, 150–400 mg/d	No differences	Fair
Kavoussi et al., 1997 (58)	248	40	48	16	Sertraline, 50–200 mg/d, vs. bupropion SR‡, 100–300 mg/d	No differences	Fair
Behnke et al., 2003 (59)	346	42	58	8	Sertraline, 50–150 mg/d, vs. mirtazapine‡, 30–45 mg/d	No differences	Fair
Mehntonon et al., 2000 (66)	147	43	66	8	Sertraline, 50–100 mg/d, vs. venlafaxine‡¶, 75–150 mg/d	Significant differences	Good

\* XR and SR refer to extended-release formulations.  
 † Differences reported here reflect statistically significant differences in any measure of efficacy; this assessment of statistical significance is influenced by sample size, study design, and population characteristics.  
 ‡ The trial was sponsored by, or at least 1 author was affiliated with, the manufacturer of the drug.  
 § Although a meta-analysis was conducted, this trial was excluded.  
 || The trial was included in the meta-analysis.  
 ¶ Statistically significant differences favor this drug.

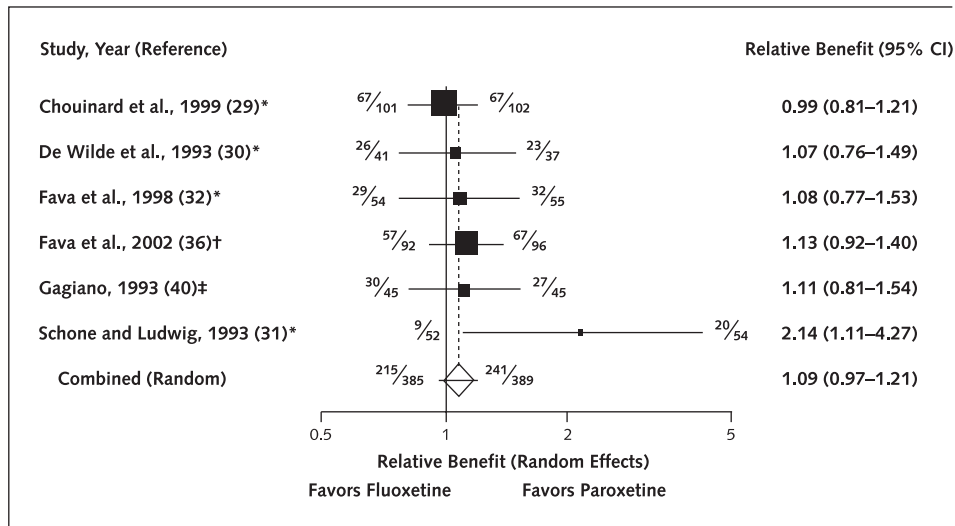
31) than to fluoxetine is based on results of 1 trial only (24) or is inconsistent with other evidence (27–31).

**Quality of Life**

Comparative evidence for health-related quality of life derives from 9 studies (22, 23, 25, 27, 28, 31, 34, 35, 37). All

studies comparing one SSRI with another provided good to fair evidence that SSRIs do not greatly differ in improving health-related quality of life. Although sleep quality was rarely assessed, 1 trial comparing fluvoxamine with fluoxetine (27) reported better sleep quality among fluvoxamine-treated patients.

Figure 1. Meta-analysis of fluoxetine compared with paroxetine.



The numbers on each side of the 95% CI are the number of responders over the total number of participants who were randomly allocated to receive that drug. The total number of responders does not always match Table 1 because of postrandomization exclusions or use of observed cases analysis. \*Trial sponsored by or authors affiliated with GlaxoSmithKline (Philadelphia, Pennsylvania) or SmithKline Beecham (Research Triangle Park, North Carolina), the manufacturer of paroxetine. †Trial sponsored by or authors affiliated with Eli Lilly (Indianapolis, Indiana), the manufacturer of fluoxetine. ‡Funding source not reported.

## Meta-Analyses of SSRIs

### Fluoxetine versus Paroxetine

We pooled 6 studies (774 patients) that compared paroxetine with fluoxetine and classified treatment response on the HAM-D scale (27–32, 36, 40). We excluded 1 study because of lack of data (28) and 1 study because it used a different outcome measure (23). The rate of being a responder at study end (Figure 1) did not differ significantly between fluoxetine and paroxetine (relative benefit, 1.09 [CI, 0.97 to 1.21]). The Begg adjusted rank correlation test (Kendall  $\tau$ , 0.3;  $P = 0.47$ ) and the Egger regression approach (intercept, 2.107 [CI,  $-0.237$  to 4.450]) suggested no major publication bias. One small study of short duration (31) reported a treatment effect that appeared to be inconsistent with the body of evidence; exclusion of this study had little effect on results.

### Fluoxetine versus Sertraline

We pooled 5 studies (1190 patients) that compared fluoxetine with sertraline (22, 33–36). We excluded 1 study because it used a different outcome measure (23). Although no individual trial reported statistically significant findings, pooled results suggested a modest additional treatment effect (relative benefit, 1.10 [CI, 1.01 to 1.22]) for sertraline compared with fluoxetine (Figure 2). The Begg adjusted rank order correlation test (Kendall  $\tau$ , 0.2;  $P = 0.82$ ) and the Egger regression approach (intercept, 0.799 [CI,  $-6.069$  to 7.668]) suggested no publication bias.

## SSRIs Compared with Other Second-Generation Antidepressants

### Efficacy

No trial comparing an SSRI with another antidepressant could be characterized as an effectiveness trial. Twenty-one efficacy trials rated fair (43–63) and 3 efficacy trials rated good (64–66) compared an SSRI with another antidepressant (Table 2).

Most trials showed evidence that SSRIs and other antidepressants have similar efficacy (43–47, 53–65). In several trials, rates of response to venlafaxine were better than those to SSRIs, but differences reached statistical significance in only 3 trials (48, 49, 66). One such trial was a Scandinavian study rated good that compared venlafaxine with sertraline (66). More venlafaxine-treated patients than sertraline-treated patients were responders ( $\geq 50\%$  improvement in HAM-D total score; 83% vs. 68%;  $P = 0.05$ ), although no statistically significant differences were noted for rates of response or remission on the Montgomery–Asberg Depression Rating Scale or Clinical Global Impression scales. Of 7 studies that compared venlafaxine with fluoxetine (46–51, 64), 4 trials rated fair and 1 trial rated good reported no statistically significant differences in response rate on the HAM-D scale at study end (46, 47, 50, 51, 64). Two trials rated fair found statistically significant differences: For venlafaxine-treated patients, 1 study reported significantly higher response rates on the HAM-D scale (48) and 1 reported significantly higher response rates on the Montgomery–Asberg Depression Rating Scale (49).

We explored possible reasons for the conflicting evidence. As with comparisons of SSRIs with other SSRIs,

trials reporting significant differences frequently were funded by a pharmaceutical company, and the company's drug was favored over the comparator. Also, as with the SSRI comparisons, we could not determine whether the statistically nonsignificant findings were consistently related to quality ratings, trial duration, sample size, or non-inferiority design.

**Speed of Response**

Twenty trials assessed evidence of speed of response to SSRIs or other antidepressants (43–45, 47–51, 54–58, 60–66). On average, clinical response occurred 4 to 6 weeks after the start of therapy. Mirtazapine consistently had a faster onset of action than did fluoxetine, sertraline, or paroxetine (45, 52, 53, 59). However, no statistically significant differences in response rates existed at study end in any of these trials. The evidence regarding a faster speed of response to venlafaxine compared with SSRIs is mixed. Three of 12 trials comparing venlafaxine with an SSRI reported significantly more responders to venlafaxine at 1 to 4 weeks (46, 49, 66).

**Quality of Life**

Five trials in 6 published articles assessed quality of life or health-related functional capacity (47, 55, 61, 62, 65, 67). No differences in overall quality of life were reported for comparisons of an SSRI with another antidepressant. In 1 trial comparing mirtazapine with sertraline (59), sleep quality was better among mirtazapine-treated patients. However, this outcome was rarely reported (and possibly not assessed) in other trials.

**Meta-Analysis of Venlafaxine versus Fluoxetine**

We pooled 6 studies (1340 patients) comparing venlafaxine with fluoxetine (46–51). We excluded 1 study be-

cause we could not extract all necessary data (64). More venlafaxine-treated patients than fluoxetine-treated patients were responders, according to the HAM-D scale, at study end (Figure 3). The Begg adjusted rank order correlation test (Kendall  $\tau$ , 0.067;  $P > 0.99$ ) and the Egger regression approach (intercept, 0.141 [CI, -4.158 to 4.440]) did not suggest publication bias.

**Tolerability and Adverse Events**

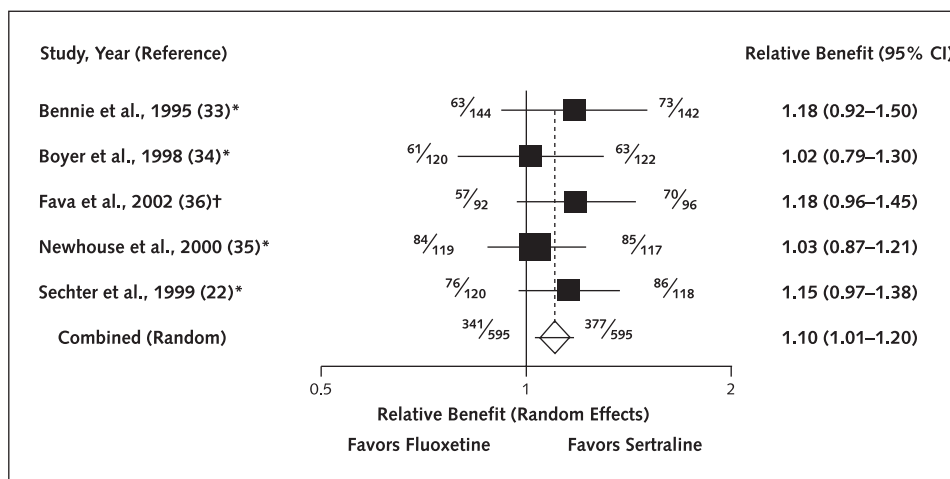
Methods used to assess adverse events and the quality of reporting of specific events differed among studies. In most trials, open-ended questioning was used to elicit adverse events, leading to great variability in the quantity and quality of reporting. An assessment scale for adverse events was used in only 6 trials (13%). Most observational studies did not assess specific adverse events, such as nausea, vomiting, dizziness, headache, or insomnia.

The overall incidence of adverse events was similar among antidepressants. Discontinuation rates attributed to adverse events were similar (78), but good to fair evidence suggested that the profiles of side effects differed among drugs. Table 3 shows the mean incidence and 95% CIs for specific adverse events that were commonly reported in trials. Statistics are descriptive only, and comparisons across drugs should be made with caution, given differences in assessment and reporting of adverse events across trials.

**Nausea and Vomiting**

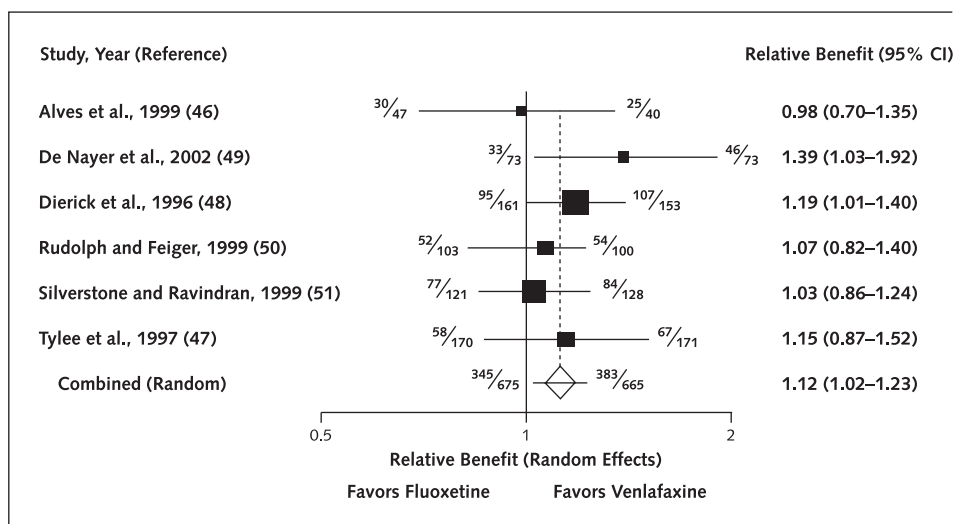
Rates of nausea and vomiting were consistently higher for venlafaxine than for SSRIs. This trend was statistically nonsignificant in 6 trials (46, 49, 51, 55, 64, 66) and significant in 6 trials (47, 48, 50, 54, 62, 63). Across venlafaxine trials, the mean incidence of nausea was 31% (CI, 27% to 34%).

Figure 2. Meta-analysis of fluoxetine compared with sertraline.



The numbers on each side of the 95% CI are the number of responders over the total number of participants who were randomly allocated to receive that drug. The total number of responders does not always match Table 1 because of postrandomization exclusions or use of observed cases analysis. \*Trial sponsored by or authors affiliated with Pfizer, Inc. (New York, New York), the manufacturer of sertraline. †Trial sponsored by or authors affiliated with Eli Lilly (Indianapolis, Indiana), the manufacturer of fluoxetine.

Figure 3. Meta-analysis of fluoxetine compared with venlafaxine.



The numbers on each side of the 95% CI are the number of responders over the total number of participants who were randomly allocated to receive that drug. The total number of responders does not always match Table 2 because of postrandomization exclusions or use of observed cases analysis. All trials were sponsored by or had authors affiliated with Wyeth-Ayerst (Madison, New Jersey), the manufacturer of venlafaxine.

**Sexual Side Effects**

We did not pool data for sexual side effects because of great variability in assessment methods and reporting. Only bupropion trials used validated scales to assess sexual side effects.

In 3 trials, the rate of sexual side effects was significantly lower for bupropion than for sertraline (44, 56, 79). One additional study reported fewer sexual side effects in bupropion-treated patients than in fluoxetine-treated patients (58). Cross-sectional evidence suggested that bupropion had one of the lowest rates of sexual side effects (80). Multiple trials provided fair evidence that paroxetine, sertraline, and mirtazapine tend to have higher rates of sexual side effects than do other antidepressants (36–38, 58, 59, 80–82).

**Weight Gain**

We did not pool data for changes in weight because of large variability in assessment methods and reporting. In

general, most trials did not report statistically or clinically ( $\geq 7\%$  change in body weight) significant changes in weight. Of trials that did report changes in body weight (21, 23, 29, 30, 36, 37, 45, 52–54, 59), those of mirtazapine, paroxetine, fluoxetine, and sertraline differed in the percentage of patients who reported such changes. On average, the most consistent percentage of patients who reported weight gain were taking mirtazapine; these patients had an estimated increase in body weight of 2 kg over 8 weeks (52, 53, 59). In the trials of SSRIs that reported changes in body weight, the percentage of patients who reported weight gain was highest in the paroxetine group and lowest in the fluoxetine group (29, 30, 36, 37); we could not estimate specific changes in weight on the basis of these data. In 1 placebo-controlled study, bupropion treatment appeared to be associated with moderate loss of body weight: Patients with a baseline body mass index of 30 kg/m<sup>2</sup> or more lost an average of 2.4 kg after 52 weeks of treatment (83).

Table 3. Mean Incidence of Specific Adverse Events across Comparative Trials\*

Drug	Mean Incidence (95% CI), %†				
	Diarrhea	Dizziness	Headache	Insomnia	Nausea
Bupropion	8.7 (1.2–16.1)	12.5 (3.4–21.6)	27.2 (18.4–36.0)	16.0 (13.3–18.7)	14.8 (8.9–20.6)
Citalopram	6.8 (1.8–11.8)	NR	5 (0–24.1)	6.4 (1.6–11.2)	11.9 (0–24.8)
Duloxetine	NR	NR	NR	NR	10.9 (0–35.6)
Escitalopram	8.9 (1.6–16.1)	NR	14.1 (0–29.9)	8.7 (1.3–16.2)	14.8 (6.1–23.5)
Fluoxetine	11.7 (6.8–16.6)	7.2 (4.3–10.0)	16.6 (10.2–23.0)	13.7 (10.0–17.4)	18.6 (15.1–22.1)
Fluvoxamine	NR	NR	14.5 (0–41.5)	NR	22.2 (0–46.8)
Mirtazapine	8.8 (0–22.4)	12.0 (2.9–21.2)	12.1 (6.3–17.9)	8 (0–49.2)	4.3 (0–8.9)
Paroxetine	9.2 (5.6–12.9)	10.6 (7.5–13.7)	21.2 (11.1–31.3)	14.3 (8.6–20.1)	18.3 (11.1–25.6)
Sertraline	15.4 (10.2–20.6)	7.5 (4.6–10.4)	20.2 (12.8–27.6)	15.0 (8.7–21.3)	19.5 (14.4–24.6)
Venlafaxine	5.5 (1.0–10.1)	15.7 (7.0–24.4)	12.8 (8.0–17.6)	11.2 (3.4–19.0)	31.0 (27.4–34.0)

\* NR = not reported.

† Calculated from data from randomized, controlled trials. The method and extent of assessment of adverse events varied among studies, and the pooled incidence should be interpreted with caution.

### Other Adverse Events

Among comparative trials, dizziness was most commonly reported among venlafaxine-treated patients (mean incidence, 16% [CI, 7% to 24%]). Headache (mean incidence, 27% [CI, 18% to 36%]) and insomnia (mean incidence, 16% [CI, 13% to 19%]) were most commonly reported among bupropion-treated patients.

In a retrospective study, venlafaxine appeared to be associated with a statistically significant increase in diastolic blood pressure (84). Case reports and secondary data analyses highlighted the potential risk for hyponatremia (85), seizures (86, 87), liver toxicity (88), and suicidality (89–95) with certain antidepressants, but evidence from randomized, controlled trials or prospective observational studies was insufficient to draw firm conclusions about these risks.

### DISCUSSION

Overall, second-generation antidepressants did not differ substantially for treating patients with major depressive disorder. Individual drugs may differ with respect to onset of action, specific adverse events, and response rates on certain health measurement scales. We underscore, however, that response rates differed only minimally and that overall rates of adverse events and discontinuation of therapy were similar across agents (78).

With rare exceptions, identifying which agent is most appropriate for any given patient is difficult. As the American Psychiatric Association suggests, when therapy with antidepressants is indicated, clinicians should make their initial selection largely on the basis of the individual patient, expected side effects, patient preference, and cost (96). In the absence of individual patient preferences in terms of expected side effects or a product-specific dosing schedule, selection of initial treatment might be based on cost.

Effectiveness trials (sometimes referred to as *practical clinical trials*) provide information on clinically relevant alternatives, cover diverse groups in heterogeneous practice settings, and encompass a broad range of health outcomes (97). We applied predefined criteria to distinguish effectiveness trials for this review and identified 3 such trials (21–23), whose conclusions were consistent with the body of evidence from efficacy trials. Although our methods have face validity, little work has been done to validate these criteria or assess the extent to which effectiveness trials differ from efficacy trials.

Systematic reviews help to eliminate some sources of bias, yet they still have important limitations. Although we found a reasonable number of studies comparing one antidepressant with another, existing evidence is insufficient to draw conclusions on all possible comparisons of individual agents. Older agents (such as fluoxetine) were studied more frequently than newer products (such as duloxetine) and commonly served as the comparator. Investigators often

compared bupropion, mirtazapine, and venlafaxine with an SSRI but not with each other. Another limitation is incomplete trial reporting. Although our analytic methods did not indicate substantial publication bias (that is, higher rates of publication among studies that show a statistically significant effect of treatment [20]), these methods had low statistical power because few trials were published for each comparison. We requested unpublished data from the U.S. Food and Drug Administration and manufacturers of included drugs but excluded unpublished studies because abstracts or summaries did not provide enough information to allow critical appraisal of the study.

Future research on pharmacotherapies for patients with major depressive disorder should include large, high-quality effectiveness trials. Measures of treatment effectiveness should focus on clinically relevant health outcomes. Investigators should define potential adverse events a priori and use consistent methods to assess them. Specifically, to assess risk fairly, investigators need to use standardized assessment and reporting techniques, especially for such events as sexual side effects and suicidal thoughts or behaviors. More, and more timely, head-to-head comparisons are crucial for clarifying remaining questions about responses to and risks of therapy.

Finally, evidence-based medicine is only as powerful as the available evidence. Because inconclusive trials funded by pharmaceutical companies have been less likely to be published, industry sponsorship for antidepressant trials is prevalent, and sponsorship was associated with a 5–percentage point difference in treatment response (favoring the sponsor's drug), publication bias probably does influence the evidence base. Evidence from clinical trials of antidepressants should be made public, regardless of the sponsor or direction of the findings.

From University of North Carolina, Chapel Hill, North Carolina.

**Acknowledgments:** The authors thank Laura Morgan for support with study retrieval; Leah Randolph for her contribution to study coordination and reference management; and Leila Kahwati, Heather Himburg, and Melissa Butler for assistance with data abstraction. In addition, they thank Mark Helfand (Center for Evidence-based Policy at the Oregon Health and Science University) and other peer reviewers, who provided constructive comments on earlier versions of the article.

**Grant Support:** Funding for this research was provided to the Cecil G. Sheps Center for Health Services Research through a subcontract with the Center for Evidence-Based Policy, Oregon Health and Science University. Dr. Gaynes was supported in part by a National Institute of Mental Health K23 Career Development Award (MH01951-03); he has also received grant and research support from the Robert Wood Johnson Foundation and the Agency for Healthcare Research and Quality.

**Potential Financial Conflicts of Interest:** *Consultancies:* B.N. Gaynes (Pfizer, Inc., Wyeth-Ayerst); *Honoraria:* B.N. Gaynes (GlaxoSmithKline); *Grants received:* B.N. Gaynes (Pfizer, Inc., Ovation Pharmaceuticals).

**Requests for Single Reprints:** Richard A. Hansen, PhD, Division of Pharmaceutical Policy and Evaluative Sciences, University of North Carolina at Chapel Hill, Room 205M, Beard Hall, Campus Box 7360, Chapel Hill, NC 27599; e-mail, rahansen@unc.edu.

Current author addresses and author contributions are available at [www.annals.org](http://www.annals.org).

## References

- Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *National Comorbidity Survey Replication*. *JAMA*. 2003;289:3095-105. [PMID: 12813115]
- Greenberg PE, Kessler RC, Birnbaum HG, Leong SA, Lowe SW, Berglund PA, et al. The economic burden of depression in the United States: how did it change between 1990 and 2000? *J Clin Psychiatry*. 2003;64:1465-75. [PMID: 14728109]
- Olfson M, Marcus SC, Druss B, Elinson L, Tanielian T, Pincus HA. National trends in the outpatient treatment of depression. *JAMA*. 2002;287:203-9. [PMID: 11779262]
- Song F, Freemantle N, Sheldon TA, House A, Watson P, Long A, et al. Selective serotonin reuptake inhibitors: meta-analysis of efficacy and acceptability. *BMJ*. 1993;306:683-7. [PMID: 8471919]
- Geddes JR, Freemantle N, Mason J, Eccles MP, Boynton J. SSRIs versus other antidepressants for depressive disorder. *Cochrane Database Syst Rev*. 2000; CD001851. [PMID: 10796826]
- Williams JW Jr, Mulrow CD, Chiquette E, Noel PH, Aguilar C, Cornell J. A systematic review of newer pharmacotherapies for depression in adults: evidence report summary. *Ann Intern Med*. 2000;132:743-56. [PMID: 10787370]
- Anderson IM. Meta-analytical studies on new antidepressants. *Br Med Bull*. 2001;57:161-78. [PMID: 11719915]
- Mulrow CD, Williams JW Jr, Chiquette E, Aguilar C, Hitchcock-Noel P, Lee S, et al. Efficacy of newer medications for treating depression in primary care patients. *Am J Med*. 2000;108:54-64. [PMID: 11059441]
- Montgomery SA, Henry J, McDonald G, Dinan T, Lader M, Hindmarch I, et al. Selective serotonin reuptake inhibitors: meta-analysis of discontinuation rates. *Int Clin Psychopharmacol*. 1994;9:47-53. [PMID: 8195583]
- Nieuwstraten CE, Dolovich LR. Bupropion versus selective serotonin-reuptake inhibitors for treatment of depression. *Ann Pharmacother*. 2001;35:1608-13. [PMID: 11793630]
- Depression Guideline Panel. *Depression in Primary Care. Vol. 2. Treatment of Major Depression. Clinical Practice Guideline No. 5*. Rockville, MD: U.S. Department of Health and Human Services, Agency for Health Care Policy and Research; 1993. AHCPR publication no. 93-0551.
- Fox DM. Evidence of evidence-based health policy: the politics of systematic reviews in coverage decisions. *Health Aff (Millwood)*. 2005;24:114-22. [PMID: 15647221]
- Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41:582-92. [PMID: 12719681]
- Harris RP, Helfand M, Woolf SH, Lohr KN, Mulrow CD, Teutsch SM, et al. Current methods of the US Preventive Services Task Force: a review of the process. Methods Work Group, Third US Preventive Services Task Force. *Am J Prev Med*. 2001;20(3 Suppl):21-35. [PMID: 11306229]
- Centre for Reviews and Dissemination. *Undertaking Systematic Reviews of Research on Effectiveness: CRD's Guidance for Those Carrying Out or Commissioning Reviews*. CRD Report No. 4. 2nd ed. York, United Kingdom: Centre for Reviews and Dissemination; 2001.
- Egger M, Smith GD, Altman DG. *Systematic Reviews in Health Care*. 2nd ed. London: BMJ; 2001.
- Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. *BMJ*. 1997;315:1533-7. [PMID: 9432252]
- Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics*. 1994;50:1088-101. [PMID: 7786990]
- Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315:629-34. [PMID: 9310563]
- Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *J Clin Epidemiol*. 2000;53:1119-29. [PMID: 11106885]
- Ekselius L, von Knorring L, Eberhard G. A double-blind multicenter trial comparing sertraline and citalopram in patients with major depression treated in general practice. *Int Clin Psychopharmacol*. 1997;12:323-31. [PMID: 9547134]
- Sechter D, Troy S, Paternetti S, Boyer P. A double-blind comparison of sertraline and fluoxetine in the treatment of major depressive episode in outpatients. *Eur Psychiatry*. 1999;14:41-8. [PMID: 10572324]
- Kroenke K, West SL, Swindle R, Gilseman A, Eckert GJ, Dolor R, et al. Similar effectiveness of paroxetine, fluoxetine, and sertraline in primary care: a randomized trial. *JAMA*. 2001;286:2947-55. [PMID: 11743835]
- Patris M, Bouchard JM, Bougerol T, Charbonnier JF, Chevalier JF, Clerc G, et al. Citalopram versus fluoxetine: a double-blind, controlled, multicentre, phase III trial in patients with unipolar major depression treated in general practice. *Int Clin Psychopharmacol*. 1996;11:129-36. [PMID: 8803650]
- Burke WJ, Gergel I, Bose A. Fixed-dose trial of the single isomer SSRI escitalopram in depressed outpatients. *J Clin Psychiatry*. 2002;63:331-6. [PMID: 12000207]
- Lepola UM, Loft H, Reines EH. Escitalopram (10-20 mg/day) is effective and well tolerated in a placebo-controlled study in depression in primary care. *Int Clin Psychopharmacol*. 2003;18:211-7. [PMID: 12817155]
- Dalery J, Honig A. Fluvoxamine versus fluoxetine in major depressive episode: a double-blind randomised comparison. *Hum Psychopharmacol*. 2003;18:379-84. [PMID: 12858325]
- Cassano GB, Puca F, Scapicchio PL, Trabucchi M. Paroxetine and fluoxetine effects on mood and cognitive functions in depressed nondemented elderly patients. Italian Study Group on Depression in Elderly Patients. *J Clin Psychiatry*. 2002;63:396-402. [PMID: 12019663]
- Chouinard G, Saxena B, Belanger MC, Ravindran A, Bakish D, Beauclair L, et al. A Canadian multicenter, double-blind study of paroxetine and fluoxetine in major depressive disorder. *J Affect Disord*. 1999;54:39-48. [PMID: 10403145]
- De Wilde J, Spiers R, Mertens C, Bartholome F, Schotte G, Leyman S. A double-blind, comparative, multicentre study comparing paroxetine with fluoxetine in depressed patients. *Acta Psychiatr Scand*. 1993;87:141-5. [PMID: 8447241]
- Schone W, Ludwig M. A double-blind study of paroxetine compared with fluoxetine in geriatric patients with major depression. *J Clin Psychopharmacol*. 1993;13(6 Suppl 2):34S-39S. [PMID: 8106654]
- Fava M, Amsterdam JD, Deltito JA, Salzman C, Schwaller M, Dunner DL. A double-blind study of paroxetine, fluoxetine, and placebo in outpatients with major depression. *Ann Clin Psychiatry*. 1998;10:145-50. [PMID: 9988054]
- Bennie EH, Mullin JM, Martindale JJ. A double-blind multicenter trial comparing sertraline and fluoxetine in outpatients with major depression. *J Clin Psychiatry*. 1995;56:229-37. [PMID: 7775364]
- Boyer P, Danion JM, Bissler JC, Hotton JM, Troy S. Clinical and economic comparison of sertraline and fluoxetine in the treatment of depression. A 6-month double-blind study in a primary-care setting in France. *Pharmacoeconomics*. 1998;13(1 Pt 2):157-69. [PMID: 10184835]
- Newhouse PA, Krishnan KR, Doraiswamy PM, Richter EM, Batar ED, Clary CM. A double-blind comparison of sertraline and fluoxetine in depressed elderly outpatients. *J Clin Psychiatry*. 2000;61:559-68. [PMID: 10982198]
- Fava M, Hoog SL, Judge RA, Kopp JB, Nilsson ME, Gonzales JS. Acute efficacy of fluoxetine versus sertraline and paroxetine in major depressive disorder including effects of baseline insomnia. *J Clin Psychopharmacol*. 2002;22:137-47. [PMID: 11910258]
- Aberg-Wistedt A, Agren H, Ekselius L, Bengtsson F, Akerblad AC. Sertraline versus paroxetine in major depression: clinical outcome after six months of continuous therapy. *J Clin Psychopharmacol*. 2000;20:645-52. [PMID: 11106136]
- Nemeroff CB, Ninan PT, Ballenger J, Lydiard RB, Feighner J, Patterson WM, et al. Double-blind multicenter comparison of fluvoxamine versus sertraline in the treatment of depressed outpatients. *Depression*. 1995;3:163-9.
- Franchini L, Gasperini M, Perez J, Smeraldi E, Zanardi R. A double-blind study of long-term treatment with sertraline or fluvoxamine for prevention of highly recurrent unipolar depression. *J Clin Psychiatry*. 1997;58:104-7. [PMID: 9108811]
- Gagliano CA. A double blind comparison of paroxetine and fluoxetine in patients with major depression. *Br J Clin Res*. 1993;4:145-52.
- Rapaport M, Coccaro E, Sheline Y, Perse T, Holland P, Fabre L, et al. A comparison of fluvoxamine and fluoxetine in the treatment of major depression.

- J Clin Psychopharmacol. 1996;16:373-8. [PMID: 8889909]
42. Kiev A, Feiger A. A double-blind comparison of fluvoxamine and paroxetine in the treatment of depressed outpatients. *J Clin Psychiatry*. 1997;58:146-52. [PMID: 9164424]
  43. Feighner JP, Gardner EA, Johnston JA, Batey SR, Khayrallah MA, Ascher JA, et al. Double-blind comparison of bupropion and fluoxetine in depressed outpatients. *J Clin Psychiatry*. 1991;52:329-35. [PMID: 1907963]
  44. Coleman CC, King BR, Bolden-Watson C, Book MJ, Segraves RT, Richard N, et al. A placebo-controlled comparison of the effects on sexual functioning of bupropion sustained release and fluoxetine. *Clin Ther*. 2001;23:1040-58. [PMID: 11519769]
  45. Hong CJ, Hu WH, Chen CC, Hsiao CC, Tsai SJ, Ruwe FJ. A double-blind, randomized, group-comparative study of the tolerability and efficacy of 6 weeks' treatment with mirtazapine or fluoxetine in depressed Chinese patients. *J Clin Psychiatry*. 2003;64:921-6. [PMID: 12927007]
  46. Alves C, Cachola I, Brandao J. Efficacy and tolerability of venlafaxine and fluoxetine in outpatients with major depression. *Primary Care Psychiatry*. 1999; 5:57-63.
  47. Tylee A, Beaumont G, Bowden MW, Reynolds A. A double-blind, randomized, 12-week comparison study of the safety and efficacy of venlafaxine and fluoxetine in moderate to severe depression in general practice. *Primary Care Psychiatry*. 1997;3:51-8.
  48. Dierick M, Ravizza L, Realini R, Martin A. A double-blind comparison of venlafaxine and fluoxetine for treatment of major depression in outpatients. *Prog Neuropsychopharmacol Biol Psychiatry*. 1996;20:57-71. [PMID: 8861177]
  49. De Nayer A, Geerts S, Ruelens L, Schittecatte M, De Bleeker E, Van Eeckhoutte I, et al. Venlafaxine compared with fluoxetine in outpatients with depression and concomitant anxiety. *Int J Neuropsychopharmacol*. 2002;5:115-20. [PMID: 12135535]
  50. Rudolph RL, Feiger AD. A double-blind, randomized, placebo-controlled trial of once-daily venlafaxine extended release (XR) and fluoxetine for the treatment of depression. *J Affect Disord*. 1999;56:171-81. [PMID: 10701474]
  51. Silverstone PH, Ravindran A. Once-daily venlafaxine extended release (XR) compared with fluoxetine in outpatients with depression and anxiety. Venlafaxine XR 360 Study Group. *J Clin Psychiatry*. 1999;60:22-8. [PMID: 10074873]
  52. Schatzberg AF, Kremer C, Rodrigues HE, Murphy GM Jr. Double-blind, randomized comparison of mirtazapine and paroxetine in elderly depressed patients. Mirtazapine vs. Paroxetine Study Group. *Am J Geriatr Psychiatry*. 2002; 10:541-50. [PMID: 12213688]
  53. Benkert O, Szegedi A, Kohlen R. Mirtazapine compared with paroxetine in major depression. *J Clin Psychiatry*. 2000;61:656-63. [PMID: 11030486]
  54. Ballus C, Quiros G, De Flores T, de la Torre J, Palao D, Rojo L, et al. The efficacy and tolerability of venlafaxine and paroxetine in outpatients with depressive disorder or dysthymia. *Int Clin Psychopharmacol*. 2000;15:43-8. [PMID: 10836286]
  55. McPartlin GM, Reynolds A, Anderson C, Casoy J. A comparison of once-daily venlafaxine XR and paroxetine in depressed outpatients treated in general practice. *Primary Care Psychiatry*. 1998;4:127-32.
  56. Coleman CC, Cunningham IA, Foster VJ, Batey SR, Donahue RM, Houser TL, et al. Sexual dysfunction associated with the treatment of depression: a placebo-controlled comparison of bupropion sustained release and sertraline treatment. *Ann Clin Psychiatry*. 1999;11:205-15. [PMID: 10596735]
  57. Croft H, Settle E Jr, Houser T, Batey SR, Donahue RM, Ascher JA. A placebo-controlled comparison of the antidepressant efficacy and effects on sexual functioning of sustained-release bupropion and sertraline. *Clin Ther*. 1999;21: 643-58. [PMID: 10363731]
  58. Kavoussi RJ, Segraves RT, Hughes AR, Ascher JA, Johnston JA. Double-blind comparison of bupropion sustained release and sertraline in depressed outpatients. *J Clin Psychiatry*. 1997;58:532-7. [PMID: 9448656]
  59. Behnke K, Sogaard J, Martin S, Bauml J, Ravindran AV, Agren H, et al. Mirtazapine orally disintegrating tablet versus sertraline: a prospective onset of action study. *J Clin Psychopharmacol*. 2003;23:358-64. [PMID: 12920411]
  60. Goldstein DJ, Mallinckrodt C, Lu Y, Demitrack MA. Duloxetine in the treatment of major depressive disorder: a double-blind clinical trial. *J Clin Psychiatry*. 2002;63:225-31. [PMID: 11926722]
  61. Detke MJ, Wiltse CG, Mallinckrodt CH, McNamara RK, Demitrack MA, Bitter I. Duloxetine in the acute and long-term treatment of major depressive disorder: a placebo- and paroxetine-controlled trial. *Eur Neuropsychopharmacol*. 2004;14:457-70. [PMID: 15589385]
  62. Bielski RJ, Ventura D, Chang CC. A double-blind comparison of escitalo-  
pram and venlafaxine extended release in the treatment of major depressive disorder. *J Clin Psychiatry*. 2004;65:1190-6. [PMID: 15367045]
  63. Montgomery SA, Huusom AK, Bothmer J. A randomised study comparing escitalopram with venlafaxine XR in primary care patients with major depressive disorder. *Neuropsychobiology*. 2004;50:57-64. [PMID: 15179022]
  64. Costa e Silva J. Randomized, double-blind comparison of venlafaxine and fluoxetine in outpatients with major depression. *J Clin Psychiatry*. 1998;59: 352-7. [PMID: 9714263]
  65. Weihs KL, Settle EC Jr, Batey SR, Houser TL, Donahue RM, Ascher JA. Bupropion sustained release versus paroxetine for the treatment of depression in the elderly. *J Clin Psychiatry*. 2000;61:196-202. [PMID: 10817105]
  66. Mehtonen OP, Sogaard J, Roponen P, Behnke K. Randomized, double-blind comparison of venlafaxine and sertraline in outpatients with major depressive disorder. Venlafaxine 631 Study Group. *J Clin Psychiatry*. 2000;61:95-100. [PMID: 10732656]
  67. Doraiswamy PM, Khan ZM, Donahue RM, Richard NE. Quality of life in geriatric depression: a comparison of remitters, partial responders, and non-responders. *Am J Geriatr Psychiatry*. 2001;9:423-8. [PMID: 11739069]
  68. Aguglia E, Casacchia M, Cassano GB, Faravelli C, Ferrari G, Giordano P, et al. Double-blind study of the efficacy and safety of sertraline versus fluoxetine in major depression. *Int Clin Psychopharmacol*. 1993;8:197-202. [PMID: 8263318]
  69. Oslin DW, Ten Have TR, Streim JE, Datto CJ, Weintraub D, DiFilippo S, et al. Probing the safety of medications in the frail elderly: evidence from a randomized clinical trial of sertraline and venlafaxine in depressed nursing home residents. *J Clin Psychiatry*. 2003;64:875-82. [PMID: 12927001]
  70. Stahl SM. Placebo-controlled comparison of the selective serotonin reuptake inhibitors citalopram and sertraline. *Biol Psychiatry*. 2000;48:894-901. [PMID: 11074227]
  71. Wade A, Crawford GM, Angus M, Wilson R, Hamilton L. A randomized, double-blind, 24-week study comparing the efficacy and tolerability of mirtazapine and paroxetine in depressed patients in primary care. *Int Clin Psychopharmacol*. 2003;18:133-41. [PMID: 12702891]
  72. Davidson JR, Meoni P, Haudiquet V, Cantillon M, Hackett D. Achieving remission with venlafaxine and fluoxetine in major depression: its relationship to anxiety symptoms. *Depress Anxiety*. 2002;16:4-13. [PMID: 12203668]
  73. Entsuah AR, Huang H, Thase ME. Response and remission rates in different subpopulations with major depressive disorder administered venlafaxine, selective serotonin reuptake inhibitors, or placebo. *J Clin Psychiatry*. 2001;62:869-77. [PMID: 11775046]
  74. Feiger AD, Flament MF, Boyer P, Gillespie JA. Sertraline versus fluoxetine in the treatment of major depression: a combined analysis of five double-blind comparator studies. *Int Clin Psychopharmacol*. 2003;18:203-10. [PMID: 12817154]
  75. Gorman JM, Korotzer A, Su G. Efficacy comparison of escitalopram and citalopram in the treatment of major depressive disorder: pooled analysis of placebo-controlled trials. *CNS Spectr*. 2002;7(4 Suppl 1):40-4. [PMID: 15131492]
  76. Stahl SM, Entsuah R, Rudolph RL. Comparative efficacy between venlafaxine and SSRIs: a pooled analysis of patients with depression. *Biol Psychiatry*. 2002;52:1166-74. [PMID: 12488062]
  77. Thase ME, Entsuah AR, Rudolph RL. Remission rates during treatment with venlafaxine or selective reuptake inhibitors. *Br J Psychiatry*. 2001; 178:234-41. [PMID: 11230034]
  78. Gartlehner G, Hansen RA, Carey TS, Lohr KN, Gaynes BN, Randolph LC. Discontinuation rates for selective serotonin reuptake inhibitors and other second-generation antidepressants in outpatients with major depressive disorder: a systematic review and meta-analysis. *Int Clin Psychopharmacol*. 2005;20:59-69. [PMID: 15729080]
  79. Segraves RT, Kavoussi R, Hughes AR, Batey SR, Johnston JA, Donahue R, et al. Evaluation of sexual functioning in depressed outpatients: a double-blind comparison of sustained-release bupropion and sertraline treatment. *J Clin Psychopharmacol*. 2000;20:122-8. [PMID: 10770448]
  80. Clayton AH, Pradko JF, Croft HA, Montano CB, Leadbetter RA, Bolden-Watson C, et al. Prevalence of sexual dysfunction among newer antidepressants. *J Clin Psychiatry*. 2002;63:357-66. [PMID: 12000211]
  81. Fava M, Judge R, Hoog SL, Nilsson ME, Koke SC. Fluoxetine versus sertraline and paroxetine in major depressive disorder: changes in weight with long-term treatment. *J Clin Psychiatry*. 2000;61:863-7. [PMID: 11105740]
  82. Feiger A, Kiev A, Shrivastava RK, Wisselink PG, Wilcox CS. Nefazodone versus sertraline in outpatients with major depression: focus on efficacy, tolerabil-

- ity, and effects on sexual function and satisfaction. *J Clin Psychiatry*. 1996;57 Suppl 2:53-62. [PMID: 8626364]
83. Croft H, Houser TL, Jamerson BD, Leadbetter R, Bolden-Watson C, Donahue R, et al. Effect on body weight of bupropion sustained-release in patients with major depression treated for 52 weeks. *Clin Ther*. 2002;24:662-72. [PMID: 1201741]
84. Thase ME. Effects of venlafaxine on blood pressure: a meta-analysis of original data from 3744 depressed patients. *J Clin Psychiatry*. 1998;59:502-8. [PMID: 9818630]
85. Liu BA, Mittmann N, Knowles SR, Shear NH. Hyponatremia and the syndrome of inappropriate secretion of antidiuretic hormone associated with the use of selective serotonin reuptake inhibitors: a review of spontaneous reports. *CMAJ*. 1996 1;155:519-27. [PMID: 8804257]
86. Johnston JA, Lineberry CG, Ascher JA, Davidson J, Khayrallah MA, Feighner JP, et al. A 102-center prospective study of seizure in association with bupropion. *J Clin Psychiatry*. 1991;52:450-6. [PMID: 1744061]
87. Dunner DL, Zisook S, Billow AA, Batey SR, Johnston JA, Ascher JA. A prospective safety surveillance study for bupropion sustained-release in the treatment of depression. *J Clin Psychiatry*. 1998;59:366-73. [PMID: 9714265]
88. Stewart DE. Hepatic adverse reactions associated with nefazodone. *Can J Psychiatry*. 2002;47:375-7. [PMID: 12025437]
89. Beasley CM Jr, Dornseif BE, Bosomworth JC, Saylor ME, Rampey AH Jr, Heiligenstein JH, et al. Fluoxetine and suicide: a meta-analysis of controlled trials of treatment for depression. *BMJ*. 1991;303:685-92. [PMID: 1833012]
90. Tollefson GD, Rampey AH Jr, Beasley CM Jr, Enas GG, Potvin JH. Absence of a relationship between adverse events and suicidality during pharmacotherapy for depression. *J Clin Psychopharmacol*. 1994;14:163-9. [PMID: 8027412]
91. Jick SS, Dean AD, Jick H. Antidepressants and suicide. *BMJ*. 1995;310:215-8. [PMID: 7677826]
92. Khan A, Khan S, Kolts R, Brown WA. Suicide rates in clinical trials of SSRIs, other antidepressants, and placebo: analysis of FDA reports. *Am J Psychiatry*. 2003;160:790-2. [PMID: 12668373]
93. Fergusson D, Doucette S, Glass KC, Shapiro S, Healy D, Hebert P, et al. Association between suicide attempts and selective serotonin reuptake inhibitors: systematic review of randomised controlled trials. *BMJ*. 2005;330:396. [PMID: 15718539]
94. Martinez C, Rietbrock S, Wise L, Ashby D, Chick J, Moseley J, et al. Antidepressant treatment and the risk of fatal and non-fatal self harm in first episode depression: nested case-control study. *BMJ*. 2005;330:389. [PMID: 15718538]
95. Gunnell D, Saperia J, Ashby D. Selective serotonin reuptake inhibitors (SSRIs) and suicide in adults: meta-analysis of drug company data from placebo controlled, randomised controlled trials submitted to the MHRA's safety review. *BMJ*. 2005;330:385. [PMID: 15718537]
96. Karasu TB, Gelenberg A, Merriam AE, Wang P. American Psychiatric Association practice guideline for the treatment of patients with major depressive disorder. *Am J Psychiatry*. 2000;157(4 Suppl):1-45.
97. Tunis SR, Stryer DB, Clancy CM. Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *JAMA*. 2003;290:1624-32. [PMID: 14506122]

**Current Author Addresses:** Dr. Hansen: Division of Pharmaceutical Policy and Evaluative Sciences, University of North Carolina at Chapel Hill, Room 205M, Beard Hall, Campus Box 7360, Chapel Hill, NC 27599.

Drs. Gartlehner and Carey: Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, 725 Airport Road, Campus Box 7590, Chapel Hill, NC 27599.

Dr. Lohr: Department of Health Policy and Administration, School of Public Health, University of North Carolina at Chapel Hill, McGavran-Greenberg, Campus Box 7411, Chapel Hill, NC 27599.

Dr. Gaynes: Department of Psychiatry, School of Medicine, University of North Carolina at Chapel Hill, 240 Med School Wing C, Campus Box 7160, Chapel Hill, NC 27599.

**Author Contributions:** Conception and design: R.A. Hansen, G. Gartlehner, B.N. Gaynes.

Analysis and interpretation of the data: R.A. Hansen, G. Gartlehner, B.N. Gaynes.

Drafting of the article: R.A. Hansen, K.N. Lohr, B.N. Gaynes.

Critical revision of the article for important intellectual content: R.A. Hansen, G. Gartlehner, K.N. Lohr, B.N. Gaynes, T.S. Carey.

Final approval of the article: R.A. Hansen, K.N. Lohr, B.N. Gaynes.

Provision of study materials or patients: R.A. Hansen, G. Gartlehner.

Statistical expertise: R.A. Hansen, B.N. Gaynes.

Obtaining of funding: B.N. Gaynes, T.S. Carey.

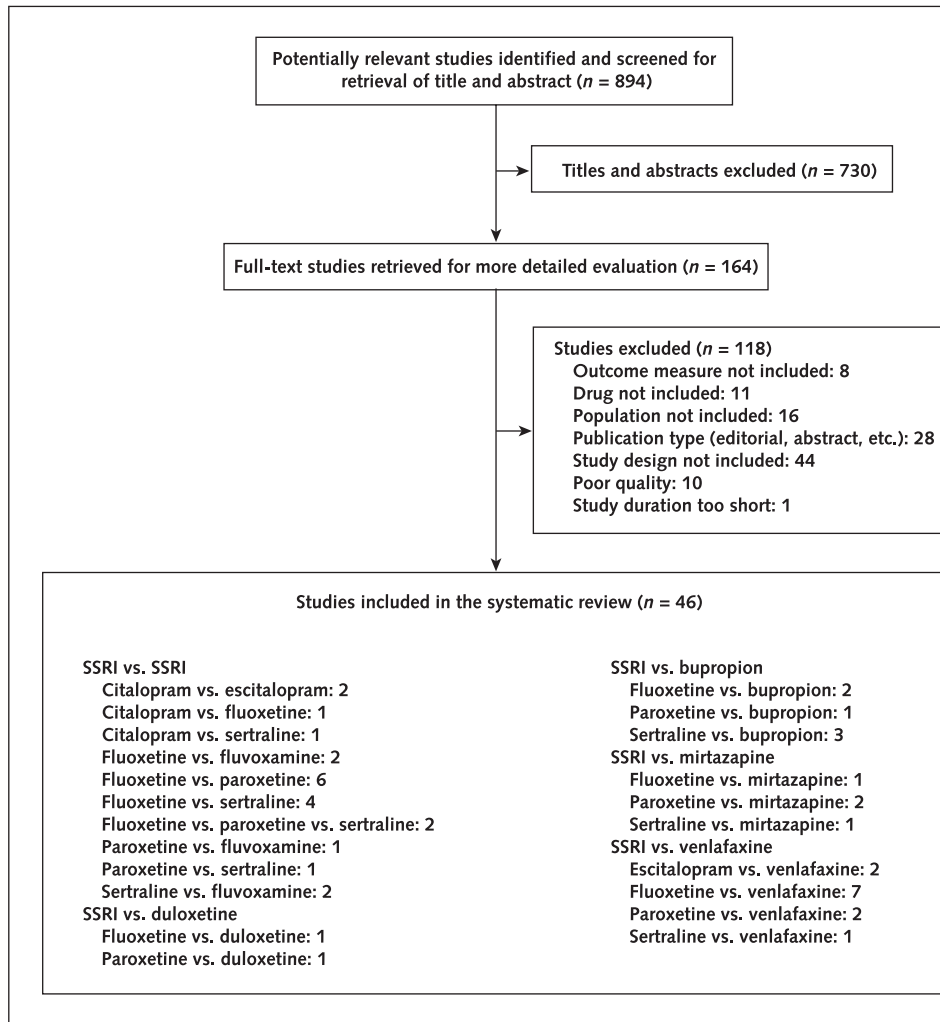
Administrative, technical, or logistic support: G. Gartlehner, T.S. Carey.

Collection and assembly of data: R.A. Hansen, G. Gartlehner, B.N. Gaynes, T.S. Carey.

*Appendix Table. Inclusion and Exclusion Criteria*

Outcome	Inclusion	Exclusion Criteria
Efficacy or effectiveness	Randomized, head-to-head trials comparing one second-generation antidepressant with another	Overall loss to follow-up > 40%
	Study duration $\geq$ 6 wk	Differential loss to follow-up > 15 percentage points
	Pediatric or adult outpatients	Statistically significant differences between treatment groups that were deemed to affect outcomes (e.g., baseline severity of illness)
	Outcomes include response, speed of response, remission, relapse, functional capacity, quality of life, or hospitalization	Fatal flaws in study design or data analysis that contribute to a quality rating of poor for internal validity
Safety or tolerability	Randomized, head-to-head trials comparing one second-generation antidepressant with another	Overall loss to follow-up > 40% in controlled trials
	Placebo-controlled trials designed specifically to assess adverse events	Differential loss to follow-up > 15 percentage points in comparative trials
	Study duration $\geq$ 6 wk (trials)	Statistically significant differences between treatment groups that were deemed to affect outcomes (e.g., baseline severity of illness)
	Observational studies covering $\geq$ 1 year	Fatal flaws in study design or data analysis that contribute to a quality rating of poor for internal validity
	Sample size $\geq$ 100 participants for observational studies	

Appendix Figure. Flow diagram of the systematic review.



SSRI = selective serotonin reuptake inhibitor.